Groupoi

1

G G [operation = function: G×G]

Vector Spaces

closed/closure

§ 1. **Binary operation on a set.** Let G be a non-empty set. Then $G \times G = \{(a, b) : a \in G, b \in G\}$. *If* $f : G \times G \to G$, *then* f *is said to be a binary operation on the set* G. The image of the ordered pair (a, b) under the function f is denoted by $f(a, b)$ or by $a f b$. Often we use the symbols $+, \times, \cdot, \circ, *$ etc. to denote binary operations on a set. Thus '+' will be a binary operation on G iff

$$a+b \in G \ \forall \ a, b \in G \text{ and } a+b \text{ is unique.}$$ (one-one)

Similarly '*' will be a binary operation on G iff

$$a * b \in G \ \forall \ a, b \in G \text{ and } a * b \text{ is unique.}$$

A **binary operation** on a set G is sometimes also called a **binary composition in the set** G. If '*' is a binary composition in G, then $\forall \ a, b \in G$, $a * b$ is a unique element of G. If $a * b \in G \ \forall \ a, b \in G$, then we also say that G is closed with respect to the composition denoted by $*$.

If there is a binary composition in a set G, the most convenient notation to denote this composition is the multiplicative notation. In this notation if $a, b \in G$, then ab represents the element obtained on multiplying a and b. Thus $ab \in G \ \forall \ a, b \in G$ if the binary composition in G has been denoted multiplicatively.

Examples. Addition is a binary operation on the set **N** of natural numbers. The sum of two natural numbers is also a natural number. Therefore **N** is closed with respect to addition

i.e., $$a+b \in \mathbf{N} \ \forall \ a, b \in \mathbf{N}.$$

Subtraction is not a binary operation on **N**. We have $4-7=-3 \notin \mathbf{N}$ whereas $4 \in \mathbf{N}$, $7 \in \mathbf{N}$. Thus **N** is not closed with respect to subtraction.

But subtraction is a binary operation in the set of integers **I**. We have $a-b \in \mathbf{I} \ \forall \ a, b \in \mathbf{I}$.

Division is not a binary operation in the set **R** of all real numbers. We have $0 \in \mathbf{R}$, $5 \in \mathbf{R}$ but $5 \div 0$ is not an element of **R**.

§ 2. **Algebraic Structure. Definition.** *A non-empty set G equipped with one or more binary operations is called an algebraic structure*

→ it must be closed

Scanned with CamScanner

ture. Suppose $*$ is a binary operation on G. Then $(G, *)$ is an algebraic structure. $(\mathbf{N}, +)$, $(\mathbf{I}, +)$, $(\mathbf{I}, -)$, $(\mathbf{R}, +, .)$ are all algebraic structures. Obviously addition and multiplication are both binary operations on the set $\mathbf{R}$ of real numbers. Therefore $(\mathbf{R}, +, .)$ is an algebraic structure equipped with two operations.

§ 3. Group. Definition.

Let G be a non-empty set equipped with a binary operation denoted by the symbol $*$ i.e., $a * b \in G \; \forall \; a, b \in G$. Then this algebraic structure is a group if the binary operation $*$ satisfies the following postulates :

1. **Associativity** *i.e., $(a * b) * c = a * (b * c) \; \forall \; a, b, c \in G$.*
2. **Existen[illegible] right [illegible]entity.** *There exists an element $e \in G$ such that a [illegible] $= a \; \forall \; a \in G$.*

The element e is called the right identity.

3. **Existence of right inverse.** *Each element of G possesses right inverse i.e., for each element $a \in G$, there exists an element $a^{-1} \in G$ such that $a * a^{-1} = e$.*

The element a^{-1} is then called the right inverse of a.

Abelian Group or Commutative Group.

Definition. *A group G is said to be abelian or commutative if in addition to the above three postulates the following postulate is also satisfied :*

4. **Commutativity** *i.e., $a * b = b * a \; \forall \; a, b \in G$.*

Note 1. In our definition of a group we have denoted the composition in G by the symbol $*$. However we can use any symbol like $\times$, o,., etc. to denote the composition. If we use the additive notation '$+$' to denote the composition in G, then the right inverse of an element $a \in G$ is denoted by the symbol $-a$ *i.e.*, we have

$$\boxed{a + (-a) = e.}$$ Inverse

Note 2. If we use multiplicative notation to denote the composition in G, then often we denote the right identity by the symbol '1'. Thus 1 is an element of G, such that $a1 = a \; \forall \; a \in G$.

There should be no confusion about 1. It is not the number 1 but it is an element of the set G whatever it may be.

In multiplicative notation the right inverse a^{-1} of a is often denoted by $1/a$.

In additive notation, we often denote the right identity by the symbol '0'. Thus 0 is an element of G such that

$$a + 0 = a \; \forall \; a \in G.$$

Note 3. In additive notation the element $a + (-b) \in G$ is denoted by $a-b$. In multiplicative notation the element $ab^{-1} \in G$ is denoted by a/b.

Ex. *Show that the set* **I** *of all integers*

$$\ldots, -4, -3, -2, -1, 0, 1, 2, 3, 4, \ldots$$

is an abelian group with respect to the operation of addition of integers.

§ 4. Some general properties of groups. Suppose G is a group with binary operation being denoted multiplicatively. Then we have the following properties in G.

1. Existence of right cancellation law *i.e., if* a, b, c *are in* G, *then*

$$\boxed{ba=ca \Rightarrow b=c.}$$

Proof. Since $a \in G$ therefore $\exists\, a^{-1} \in G$ such that $aa^{-1}=e$ where e is the right identity.

Now $\quad ba=ca$

$\Rightarrow (ba)\, a^{-1}=(ca)\, a^{-1}$

$\Rightarrow b\,(aa^{-1})=c\,(aa^{-1})$ [by associativity]

$\Rightarrow be=ce$ [$\because$ a^{-1} is right inverse of a]

$\Rightarrow b=c$ [$\because$ e is right identity]

2. *The right identity is also the left identity i.e., if e is the right identity then* $ea=a \;\forall\; a \in G$.

Proof. Let $a \in G$ and e be the right identity. Since a possesses right inverse, therefore there exists $a^{-1} \in G$ such that

$$aa^{-1}=e.$$

Now $\quad (ea)\, a^{-1}=e\,(aa^{-1})$ [by associativity]

$=ee$ [$\because$ $aa^{-1}=e$]

$=e$ [$\because$ e is right identity]

$=aa^{-1}$ [$\because$ $aa^{-1}=e$]

Now $\quad (ea)\, a^{-1}=aa^{-1}$

$\Rightarrow ea=a.$ [by right cancellation law]

$\therefore$ e is also the left identity.

Hence e is the identity *i.e.*, $ea=a=ae \;\forall\; a \in G$.

3. *The right inverse of an element is also its left inverse i.e., if* a^{-1} *is the right inverse of* a *then* $a^{-1}a=e$.

Proof. Let $a \in G$ and e be the identity element. Let a^{-1} be the right inverse of a *i.e.*, $aa^{-1}=e$. To prove that $a^{-1}\, a=e$.

We have

$(a^{-1}\, a)\, a^{-1}=a^{-1}\,(aa^{-1})$ [by associativity]

$=a^{-1}\, e$ [$\because$ $aa^{-1}=e$]

$=a^{-1}$ [$\because$ e is right identity]

$=ea^{-1}$ [$\because$ e is also left identity]

Now $(a^{-1}a)\,a^{-1}=ea^{-1}$ [by right cancellation law]

$\Rightarrow a^{-1}a=e.$

$\therefore$ a^{-1} is also the left inverse of a.

$\therefore$ a^{-1} is the inverse of a i.e., $a^{-1}a=e=aa^{-1}$.

4. Uniqueness of Identity. *The identity element in a group is unique.*

Proof. Suppose e and e' are two identity elements of a group G. We have

$ee'=e$ if e' is identity.

and $ee'=e'$ if e is identity.

But ee' is a unique element of G.

$\therefore$ $ee'=e$ and $ee'=e' \Rightarrow e=e'$.

Hence the identity element is unique.

5. Uniqueness of Inverse. *The inverse of each element of a group is unique.*

Proof. Let a be any element of a group G and let e be the identity element. Suppose b and c are two inverses of a

i.e., $ba=e=ab$

and $ca=e=ac$.

We have $b\,(ac)=be$

$=b$ [$\because$ e is identity]

Also $(ba)\,c=ec$ [$\because$ $ba=e$]

$=c$ [$\because$ e is identity]

But in a group the composition is associative.

$\therefore$ $b\,(ac)=(ba)\,c$

$\Rightarrow b=c$.

6. Existence of left cancellation law *i.e., if a, b, c are in G, then*

$$ab=ac \Rightarrow b=c.$$

Proof. We have $ab=ac$

$\Rightarrow a^{-1}(ab)=a^{-1}(ac)$

$\Rightarrow (a^{-1}a)\,b=(a^{-1}a)\,c$

$\Rightarrow eb=ec$ [$\because$ a^{-1} is left inverse of a]

$\Rightarrow b=c$ [$\because$ e is left identity]

§ 5. Field. Definition. *Suppose F is a non-empty set equipped with two binary operations called addition and multiplication and denoted by '+' and '.' respectively i.e., for all $a, b \in F$ we have $a+b \in F$ and $a.b \in F$. Then this algebraic structure $(F, +, .)$ is called a field, if the following postulates are satisfied :*

F_1. *Addition is commutative, i.e.,*

$$a+b=b+a \;\forall\; a, b \in F.$$

F_2. *Addition is associative, i.e.,*
$$(a+b)+c=a+(b+c) \ \forall \ a, b, c \in F.$$
F_3. *$\exists$ an element denoted by* 0 *(called zero) in F such that*
$$a+0=a \ \forall \ a \in F.$$
F_4. *To each element a in F there exists an element $-a$ in F such that* $a+(-a)=0$.

F_5. *Multiplication is commutative, i.e.,*
$$a.b=b.a \ \forall \ a, b \in F.$$
F_6. *Multiplication is associative, i.e.,*
$$a.(b.c)=(a.b).c \ \forall \ a, b, c \in F.$$
F_7. *$\exists$ a non-zero element denoted by* 1, *(called* **one**) *in F such that* $a.1=a \ \forall \ a \in F$.

F_8. *To every non-zero element a in F there corresponds an element a^{-1} (or $1/a$) in F, such that* $a.a^{-1}=1$.

F_9. *Multiplication is distributive with respect to addition i.e., for all a, b, c in F, we have* $a.(b+c)=a.b+a.c$.

The element 0 is identity element for addition composition in F. It is called the **zero element** of the field. Obviously $(F, +)$ is an abelian group. If $0 \neq a \in F$, then a is called a non-zero element of F. The element 1 is identity element for multiplication composition in F. It is called the **unity** of the field. In a field each non-zero element is **invertible** *i.e.*, possesses inverse for multiplication composition.

Note. In future we shall denote the multiplication composition in a field F not by the symbol '.' but by multiplication notation. Thus we shall omit '.' and we shall write ab in place of $a.b$.

Subfield. Definition. *Let F be a field. A non-empty subset of the set F is said to be a subfield of F if K is closed with respect to the operations of addition and multiplication in F and K itself is a field for these operations.*

Examples of Fields

Example 1. The set **Q** of all rational numbers is a field the addition and multiplication of rational numbers being the two field compositions. The rational number 0 is the **zero element** of this field and the rational number 1 is the **unity** of this field.

Example 2. The set **R** of all real numbers is a field, the addition and multiplication of real numbers being the two field compositions. Since $\mathbf{Q} \subset \mathbf{R}$, therefore the field of rational numbers is a subfield of the field of rational numbers.

Scanned with CamScanner

Example 3. The set C of all complex numbers is a field, the addition and multiplication of complex numbers being the two field compositions. Since $R \subset C$, therefore the field of real numbers is a subfield of the field of complex numbers.

Example 4. *The set of numbers of the form $a+b\sqrt{2}$, with a and b as rational numbers is a field.* We can easily show that all the field postulates are satisfied in this case.

§ 6. Elementary properties of a field.

Theorem. *If F is a field, then for all $a, b, c \in F$*

(i) $a+b=a+c \Rightarrow b=c$.

(ii) $a0=0a=0$.

(iii) $a(-b)=-(ab)$.

(iv) $-(-a)=a$.

(v) $(-a)(-b)=ab$.

(vi) $a(b-c)=ab-ac$.

(vii) $ab=0 \Rightarrow a=0$ *or* $b=0$ (*or both*).

(viii) $a\neq 0,\ ab=ac \Rightarrow b=c$.

Proof. (i) We have $a+b=a+c$

$\Rightarrow (-a)+(a+b)=(-a)+(a+c)$

$\Rightarrow [(-a)+a]+b=[(-a)+a]+c$ [by F_2]

$\Rightarrow [a+(-a)]+b=[a+(-a)]+c$ [by F_1]

$\Rightarrow 0+b=0+c$ [by F_4]

$\Rightarrow b+0=c+0$ [by F_1]

$\Rightarrow b=c$. [by F_3]

(ii) We have

$a0=a(0+0)$

$=a0+a0$. [$\because$ by F_3, $0+0=0$]

$\therefore a0+0=a0+a0$ [by F_9]

$\Rightarrow 0=a0$ [cancelling $a0$ by (i) which we have just proved]. [$\because$ $a0 \in F$ and $a0+0=a0$]

Since in a field multiplication is commutative, therefore

$0a=a0=0$.

(iii) We have $a[b+(-b)]=a0$

$\Rightarrow ab+a(-b)=0$ [$\because$ $b+(-b)=0$]

$\Rightarrow ab+a(-b)=(ab)+[-(ab)]$ [by F_9 and by (ii)]

$\Rightarrow a(-b)=-(ab)$ [cancelling ab by (i)]

(iv) We have $a+(-a)=0$ [by F_4]

$\Rightarrow (-a)+a=0$ [by F_1]

$\Rightarrow (-a)+a=(-a)+[-(-a)]$ [by F_4]

$\Rightarrow a=-(-a)$ [by (i)]

(v) $(-a)(-b) = -[(-a)\, b]$ [by (iii)]

$= -[b\,(-a)]$ [by F_5]

$= -[-(ba)]$ [by (iii)]

$= ba.$ [by (iv)]

$= ab.$ [by F_5]

(vi) $a\,(b-c) = a\,[b+(-c)]$

$= ab + a\,(-c)$ [by F_9]

$= ab + [-(ac)]$ [by (iii)]

$= ab - ac.$

(vii) Let $ab = 0.$

Suppose $a \neq 0.$ Then a^{-1} exists.

$\therefore\ ab=0 \Rightarrow a^{-1}(ab) = a^{-1}\,0 \Rightarrow (a^{-1}a)\,b = 0 \Rightarrow (aa^{-1})\,b = 0$

$\Rightarrow 1\,b = 0 \Rightarrow b\,1 = 0 \Rightarrow b = 0.$

Similarly we can prove that if $b \neq 0$, then a must be zero.

(viii) We have $a \neq 0$ and $ab = ac.$

$\because\ a \neq 0$, therefore a^{-1} exists.

$\therefore\ ab = ac \Rightarrow a^{-1}(ab) = a^{-1}(ac) \Rightarrow (a^{-1}a)\,b = (a^{-1}a)\,c$

$\Rightarrow (aa^{-1})\,b = (aa^{-1})\,c \Rightarrow 1b = 1c \Rightarrow b1 = c1 \Rightarrow b = c.$

§ 7. **Vector spaces.** So far we have studied groups and fields. Now we shall study another important algebraic structure known as **vector space** or (**linear space**). Before giving the definition of a vector space we shall make a distinction between **internal and external compositions.**

Let A be any set. If $a * b \in A \ \forall\ a, b \in A$, and $a * b$ is unique then $*$ is said to be an **internal composition** in the set A. Here a and b are both elements of the set A.

Let V and F be any two sets. If $a \circ \alpha \in V$ for all $a \in F$ and for all $\alpha \in V$ and $a \circ \alpha$ is unique, then $\circ$ is said to be an **external composition** in V over F. Here a is an element of the set F and α is an element of the set V and the resulting element $a \circ \alpha$ is an element of the set V

Vector space. Definition. **(Nagarjuna 1980; Kakatiya 91; Osmania 90; Marathwada 92; Meerut 89, 90; Madras 81; Kanpur 81; Poona 88)**

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

1. *There is defined an internal composition in* V *called* **addition of vectors** *and denoted by* '$+$'. *Also for this composition* V *is an abelian group i.e.,*

(i) $\alpha+\beta \in V$ *for all* $\alpha, \beta \in V$.
(ii) $\alpha+\beta=\beta+\alpha$ *for all* $\alpha, \beta \in V$.
(iii) $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$ *for all* $\alpha, \beta, \gamma \in V$.
(iv) $\exists$ *an element* $\mathbf{0} \in V$ *such that* $\alpha+\mathbf{0}=\alpha$ *for all* $\alpha \in V$.
This element $\mathbf{0} \in V$ *will be called the zero vector.*
(v) *To every vector* $\alpha \in V$ *there exists a vector* $-\alpha \in V$ *such that*

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

2. *There is an external composition in V over F called* **scalar multiplication** *and denoted multiplicatively i.e.,* $a\alpha \in V$ *for all* $a \in F$ *and for all* $\alpha \in V$. *In other words V is closed with respect to scalar multiplication.*

3. *The two compositions i.e., scalar multiplication and addition of vectors satisfy the following postulates :*

(i) $a(\alpha+\beta)=a\alpha+a\beta \ \forall \ a \in F$ *and* $\forall \ \alpha, \beta \in V$.
(ii) $(a+b)\alpha=a\alpha+b\alpha \ \forall \ a, b \in F$ *and* $\forall \ \alpha \in V$.
(iii) $(ab)\alpha=a(b\alpha) \ \forall \ a, b \in F$ *and* $\forall \ \alpha \in V$.
(iv) $1\alpha=\alpha \ \forall \ \alpha \in V$ *and* 1 *is the unity element of the field F.*

When V is a vector space over the field F, we shall say that $V(F)$ is a vector space. If the field F is understood we can simply say that V is a vector space. If F is the field **R** of real numbers, V is called a *real vector space*; similarly if F is **Q** or F is **C**, we speak of *rational vector spaces or complex vector spaces.*

In the above definition of a vector space V over the field F, we have denoted the addition of vectors by the symbol '+'. This symbol also denotes the addition composition of the field F i.e., addition of scalars. There should be no confusion about the two compositions though we have used the same symbol to denote each of them. If α $\beta \in V$, then $\alpha+\beta$ represents addition of V i.e., addition of vectors. If $a, b \in F$ then $a+b$ represents addition of scalars *i.e.*, addition in the field F. Similarly there should be no confusion in multiplication of scalars *i.e.*, multiplication of elements of F and in scalar multiplication *i.e.*, multiplication of an element of V by an element of F. If $a, b \in F$, then ab represents multiplication of F and $ab \in F$. If $a \in F$ and $\alpha \in V$, then $a\alpha$ represents scalar multiplication and $a\alpha \in V$. Since $1 \in F$ and $\alpha \in V$, therefore 1α represents scalar multiplication. Again $a\alpha \in V$, $a\beta \in V$, therefore $a\alpha+a\beta$ represents addition of vectors and thus $a\alpha+a\beta$ is an element of V. Further $a \in F$ and $\alpha+\beta \in V$, therefore $a(\alpha+\beta)$ represents scalar multiplication and we have $a(\alpha+\beta) \in V$.

Note 1. Since $(V, +)$ is an abelian group, therefore ...

properties of an abelian group will hold in V. A few of them are as follows :

(i) $\alpha+\beta=\alpha+\gamma \Rightarrow \beta=\gamma$ (left cancellation law)

(ii) $\beta+\alpha=\gamma+\alpha \Rightarrow \beta=\gamma$ (right cancellation law)

(iii) $\alpha+\beta=\mathbf{0} \Rightarrow \alpha=-\beta$ and $\beta=-\alpha$.

(iv) $-(\alpha+\beta)=-\alpha-\beta$ where by $\alpha-\beta$ we mean $\alpha+(-\beta)$.

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

(vi) $\alpha+\beta=\alpha \Rightarrow \beta=\mathbf{0}$.

(vii) $\alpha+(\beta-\alpha)=\beta$.

(viii) The additive identity $\mathbf{0}$ will be unique.

(ix) The additive inverse of each vector will be unique.

(x) If $\alpha+\beta=\gamma$, then $\alpha+\beta-\gamma=\mathbf{0}$.

Note 2. There should be no confusion about the use of the word vector. Here by vector we do not mean the vector quantity which we have defined in vector algebra as a directed line segment. Here we shall call the elements of the set V as vectors.

Note 3. In a vector space we shall be dealing with two types of zero elements. One is the zero vector and the other is the zero element of the field F *i.e.*, the 0 scalar. To distinguish between the two, we shall use the zero letter in bold type to represent the zero vector. Also we shall use the lower case Greek letters α, β, γ etc. to denote vectors *i.e.*, the elements of V and the lower case Latin letters a, b, c etc. to denote scalars *i.e.*, the elements of the field F.

Example 1. *A field K can be regarded as a vector space over any subfield F of K.* **(Kanpur 1981, Poona 85; Meerut 89)**

Here K is the set of vectors. Addition of vectors is the addition composition in the field K. Since K is a field, therefore $(K, +)$ is an abelian group. Further the elements of the subfield F constitute the set of scalars. The composition of scalar multiplication is the multiplication composition in the field K. K is a field, therefore $a\alpha \in K \; \forall \; a \in F$ and $\forall \; \alpha \in K$ because both a and α are elements of K. If 1 is the unity element of K, then 1 is also the unity element of the subfield F. We make the following observations :

(i) $a\,(\alpha+\beta)=a\alpha+a\beta \; \forall \; a \in F$ and $\forall \; \alpha, \beta \in K$. This result follows from the left distributive law in K.

(ii) $(a+b)\,\alpha=a\alpha+b\alpha \; \forall \; a, b \in F$ and $\forall \; \alpha \in K$. This result is a consequence of the right distributive law in K.

Scanned with CamScanner

(iii) $(ab)\,\alpha = a\,(b\alpha)$ $\forall$ $a, b \in F$ and $\forall$ $\alpha \in K$. This result is a consequence of associativity of multiplication in K.

(iv) $1\alpha = \alpha$ $\forall$ $\alpha \in K$ and 1 is the unity element of the subfield F. Since 1 is also the unity element of the field K, therefore $1\alpha = \alpha$ $\forall$ $\alpha \in K$. Hence $K(F)$ is a vector space.

Note 1. If F is any field, then F itself is a vector space over the field F.

Note 2. If **C** is the field of complex numbers and **R** is the field of real numbers, then **C** is a vector space over **R** because **R** is a subfield of **C**. But **R** is not a vector space over **C**. Here **R** is not closed with respect to scalar multiplication. For example $2 \in \mathbf{R}$ and $3+4i \in \mathbf{C}$ and $(3+4i)\,2 \notin \mathbf{R}$.

(Meerut 1988)

Example 2. *The set V of all $m \times n$ matrices with their elements as real numbers is a vector space over the field F of real numbers with respect to addition of matrices as addition of vectors and multiplication of a matrix by a scalar as scalar multiplication.*

(Meerut 1967)

We can easily prove that V is an abelian group with respect to addition of matrices. The null matrix O of the type $m \times n$ is the additive identity of this abelian group.

If $a \in F$ and $\alpha \in V$ (*i.e.*, α is a matrix of the type $m \times n$ with elements as real numbers), then $a\alpha \in V$ because $a\alpha$ is also a matrix of the type $m \times n$ with elements as real numbers. Therefore V is closed with respect to scalar multiplication. Also from our study of matrices we observe that

(i) $a(\alpha+\beta) = a\alpha + a\beta$ $\forall$ $a \in F$ and $\forall$ $\alpha, \beta \in V$.

(ii) $(a+b)\,\alpha = a\alpha + b\alpha$ $\forall$ $a, b \in F$ and $\forall$ $\alpha \in V$.

(iii) $(ab)\,\alpha = a\,(b\alpha)$ $\forall$ $a, b \in F$ and $\forall$ $\alpha \in V$.

(iv) $1\alpha = \alpha$ $\forall$ $\alpha \in V$ and 1 is the unity element of the field F of real numbers.

Hence $V(F)$ is a vector space.

Note. If V is the set of all $m \times n$ matrices with their elements as rational numbers and F is the field of real numbers, then V will not be closed with respect to scalar multiplication. For $\sqrt{7} \in F$ and if $\alpha \in V$, then $\sqrt{7}\alpha \notin V$ because the elements of the matrix $\sqrt{7}\alpha$ will not be rational numbers. Therefore $V(F)$ will not be a vector space.

Example 3. *The vector space of all ordered n-tuples over a field F.*

(Meerut 1983P; Gorakhpur 81)

Scanned with CamScanner

Let F be a field. An ordered set $\alpha=(a_1, a_2, a_3, ..., a_n)$ of n elements of F is called an *n-tuple* over F. Let V be the totality of all ordered *n-tuples over F i.e.*, let

$$V=\{(a_1, a_2, ..., a_n) : a_1, a_2, ..., a_n \in F\}.$$

Now we shall give a vector space structure to V over the field F. For this we define equality of two n-tuples, addition of two n-tuples and multiplication of an n-tuple by a scalar as follows :

Equality of two n-tuples. Two elements $\alpha=(a_1, a_2, ..., a_n)$ and $\beta=(b_1, b_2, ..., b_n)$ of V are said to be equal if and only if

$$a_i=b_i \text{ for each } i=1, 2, ..., n.$$

Addition composition in V. We write

$$\alpha+\beta=(a_1+b_1, a_2+b_2, ..., a_n+b_n)$$
$$\forall\ \alpha=(a_1, a_2, ..., a_n), \beta=(b_1, b_2, ..., b_n)\in V.$$

Since $a_1+b_1, a_2+b_2, ..., a_n+b_n$ are all elements of F, therefore $\alpha+\beta \in V$ and thus V is closed with respect to addition of n-tuples.

Scalar Multiplication composition in V over F. We define

$$a\alpha=(aa_1, aa_2, ..., aa_n)\ \forall\ a\in F, \alpha=(a_1, a_2, ..., a_n)\in V.$$

Since $aa_1, aa_2, ..., aa_n$ are all elements of F, therefore $a\alpha\in V$ and thus V is closed with respect to scalar multiplication.

Now we shall see that V is a vector space for these two compositions.

Associativity of addition in V. We have

$$(a_1, a_2,..., a_n)+[(b_1, b_2,..., b_n)+(c_1, c_2,..., c_n)]$$
$$=(a_1, a_2,..., a_n)+(b_1+c_1, b_2+c_2,..., b_n+c_n)$$
$$=(a_1+[b_1+c_1], a_2+[b_2+c_2],..., a_n+[b_n+c_n])$$
$$=([a_1+b_1]+c_1, [a_2+b_2]+c_2,..., [a_n+b_n]+c_n)$$
$$=(a_1+b_1, a_2+b_2,..., a_n+b_n)+(c_1, c_2,..., c_n)$$
$$=[(a_1, a_2,..., a_n)+(b_1, b_2,..., b_n)]+(c_1, c_2,..., c_n).$$

Commutativity of addition in V. We have

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

Existence of additive identity in V. We have $(0, 0,..., 0)\in V$. Also if $(a_1, a_2,..., a_n)\in V$, then

$$(a_1, a_2,..., a_n)+(0, 0,..., 0)=(a_1+0, a_2+0,..., a_n+0)$$
$$=(a_1, a_2,..., a_n).$$

$\therefore$ $(0, 0,..., 0)$ is the additive identity in V.

Existence of additive inverse of each element of V. If $(a_1, a_2,..., a_n)\in V$, then $(-a_1, -a_2,..., -a_n)\in V$.

Also we have

$$(-a_1, -a_2,..., -a_n)+(a_1, a_2,..., a_n)$$

Scanned with CamScanner

$=(-a_1+a_1, -a_2+a_2, \ldots, -a_n+a_n)=(0, 0, 0, \ldots, 0).$

$\therefore (-a_1, -a_2, \ldots, -a_n)$ is the additive inverse of $(a_1, a_2, \ldots, a_n)$.

Thus V is an abelian group with respect to addition. Further we observe that

1. If $a \in F$ and $\alpha=(a_1, a_2, \ldots, a_n)$, $\beta=(b_1, b_2, \ldots, b_n) \in V$, then

$$a(\alpha+\beta)=a\,(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$$
$$=(a\,[a_1+b_1], a[a_2+b_2], \ldots, a[a_n+b_n])$$
$$=(aa_1+ab_1, aa_2+ab_2, \ldots, aa_n+ab_n)$$
$$=(aa_1, aa_2, \ldots, aa_n)+(ab_1, ab_2, \ldots, ab_n)$$
$$=a\,(a_1, a_2, \ldots, a_n)+a\,(b_1, b_2, \ldots, b_n)=a\alpha+a\beta.$$

2. If $a, b \in F$ and $(a_1, a_2, \ldots, a_n) \in V$, then

$$(a+b)\,\alpha=([a+b]\,a_1, [a+b]\,a_2, \ldots, [a+b]\,a_n)$$
$$=(aa_1+ba_1, aa_2+ba_2, \ldots, aa_n+ba_n)$$
$$=(aa_1, aa_2, \ldots, aa_n)+(ba_1, ba_2, \ldots, ba_n)$$
$$=a\,(a_1, a_2, \ldots, a_n)+b\,(a_1, a_2, \ldots, a_n)=a\alpha+b\alpha.$$

3. If $a, b \in F$ and $\alpha=(a_1, a_2, \ldots, a_n) \in V$, then

$$(ab)\,\alpha=([ab]\,a_1, [ab]\,a_2, \ldots, [ab]\,a_n)=(a\,[ba_1], a\,[ba_2], \ldots, a\,[ba_n])$$
$$=a\,(ba_1, ba_2, \ldots, ba_n)=a\,[b(a_1, a_2, \ldots, a_n)]=a\,(b\alpha).$$

4. If 1 is the unity element of F and $\alpha=(a_1, a_2, \ldots, a_n) \in V$, then $\quad 1\alpha=(1a_1, 1a_2, \ldots, 1a_n)=(a_1, a_2, \ldots, a_n)=\alpha.$

Hence V is a vector space over F. *The vector space of all ordered n-tuples over F will be denoted by $V_n(F)$. Sometimes we also denote it by $F^{(n)}$ or by F^n.* Here the zero vector *i.e.*, 0 is the *n*-tuple $(0, 0, \ldots, 0)$.

Note. $V_2(F)=\{(a_1, a_2) : a_1, a_2 \in F\}$ is the vector space of all **ordered pairs** over F. Similarly $V_3(F)=\{(a_1, a_2, a_3) : a_1, a_2, a_3 \in F\}$ is the vector space of all **ordered triads** over F.

Example 4 *The vector space of all polynomials over a field F.*

(Meerut 1982; Andhra 92)

Sol. Let $F[x]$ denote the set of all polynomials in an indeterminate x over a field F. Then $F[x]$ is a vector space over the field F with respect to addition of two polynomials as addition of vectors and the product of a polynomial by a constant polynomial (*i.e.*, by an element of F) as scalar multiplication.

Let $f(x)=\Sigma\, a_i x^i=a_0+a_1x+a_2x^2+a_3x^3+\ldots$

$g(x)=\Sigma\, b_i x^i=b_0+b_1x+b_2x^2+b_3x^3+\ldots$

and $h(x)=\Sigma\, c_i x^i=c_0+c_1x+c_2x^2+c_3x^3+\ldots$

be any arbitrary members of $F[x]$.

Equality of two polynomials. We define $f(x)=g(x)$ if and only if $a_i=b_i$ for each $i=0, 1, 2, \ldots, \ldots$

Addition composition in $F[x]$. We define

$$f(x)+g(x)=(a_0+b_0)+(a_1+b_1)\,x+(a_2+b_2)\,x^2+\ldots$$
$$=\Sigma\,(a_i+b_i)\,x^i.$$

Since a_0+b_0, a_1+b_1, $a_2+b_2, \ldots$, are all elements of F, therefore $f(x)+g(x) \in F[x]$ and thus $F[x]$ is closed with respect to addition of polynomials.

Scalar multiplication in $F[x]$ over F. If k is any scalar *i.e.*, $k \in F$, we define

$$kf(x)=ka_0+(ka_1)x+(ka_2)x^2+(ka_3)\,x^3+\ldots$$
$$=\Sigma\,(ka_i)\,x^i.$$

Since ka_0, ka_1, $ka_2, \ldots$, are all elements of F, therefore $kf(x) \in F[x]$ and thus $F[x]$ is closed with respect to scalar multiplication.

Now we shall show that $F[x]$ is a vector space for these two compositions.

Commutativity of addition in $F[x]$. We have

$$f(x)+g(x)=(a_0+b_0)+(a_1+b_1)\,x+(a_2+b_2)\,x^2+\cdots$$
$$=(b_0+a_0)+(b_1+a_1)\,x+(b_2+a_2)\,x^2+\ldots$$
[$\because$ addition in the field F is commutative]
$$=g(x)+f(x).$$

Associativity of addition in $F[x]$. We have

$$[f(x)+g(x)]+h(x)=\Sigma\,(a_i+b_i)\,x^i+\Sigma\,c_ix^i$$
$$=\Sigma\,[(a_i+b_i)+c_i]\,x^i=\Sigma\,[a_i+(b_i+c_i)]\,x^i$$
$$=\Sigma\,a_ix^i+\Sigma\,(b_i+c_i)\,x^i=f(x)+[g(x)+h(x)].$$

Existence of additive identity in $F[x]$. Let 0 denote the zero polynomial over the field F *i.e.*,

$$0=0+0x+0x^2+0x^3+\ldots.$$

Then $0 \in F[x]$ and $0+f(x)=f(x)$.

$\therefore$ the zero polynomial 0 is the additive identity.

Existence of additive inverse of each member of $F[x]$. Let $-f(x)$ be the polynomial over the field F defined as

$$-f(x)=-a_0+(-a_1)\,x+(-a_2)\,x^2+\ldots$$
$$=-a_0-a_1x-a_2x^2-\ldots.$$

Then $-f(x) \in F[x]$ and we have

$-f(x)+f(x)=0$ *i.e.*, the zero polynomial.

$\therefore$ $-f(x)$ is the additive inverse of $f(x)$.

Scanned with CamScanner

Thus $F[x]$ is an abelian group with respect to addition of polynomials.

Now for the operation of scalar multiplication we make the following observations.

1. If $k \in F$, then

$$k[f(x)+g(x)]=k[(a_0+b_0)+(a_1+b_1)x+(a_2+b_2)x^2+...]$$
$$=k(a_0+b_0)+k(a_1+b_1)x+k(a_2+b_2)x^2+...$$
$$=(ka_0+kb_0)+(ka_1+kb_1)x+(ka_2+kb_2)x^2+...$$
$$=[ka_0+(ka_1)x+(ka_2)x^2+...]+[kb_0+(kb_1)x+(kb_2)x^2+...]$$
$$=k(a_0+a_1x+a_2x^2+...)+k(b_0+b_1x+b_2x^2+...)$$
$$=kf(x)+kg(x).$$

2. If $k_1, k_2 \in F$, then

$$(k_1+k_2)f(x)=(k_1+k_2)a_0+[(k_1+k_2)a_1]x+[(k_1+k_2)a_2]x^2+...$$
$$=(k_1a_0+k_2a_0)+(k_1a_1+k_2a_1)x+(k_1a_2+k_2a_2)x^2+...$$
$$=[k_1a_0+(k_1a_1)x+(k_1a_2)x^2+...]+[k_2a_0+(k_2a_1)x+(k_2a_2)x^2+...]$$
$$=k_1(a_0+a_1x+a_2x^2+...)+k_2(a_0+a_1x+a_2x^2+...)$$
$$=k_1f(x)+k_2f(x).$$

3. If $k_1, k_2 \in F$, then

$$(k_1k_2)f(x)=(k_1k_2)a_0+[(k_1k_2)a_1]x+[(k_1k_2)a_2]x^2+...$$
$$=k_1(k_2a_0)+[k_1(k_2a_1)]x+[k_1(k_2a_2)]x^2+...$$
$$=k_1[k_2a_0+(k_2a_1)x+(k_2a_2)x^2+...]$$
$$=k_1[k_2f(x)].$$

4. If 1 is the unity element of the field F, then

$$1f(x)=(1a_0)+(1a_1)x+(1a_2)x^2+...$$
$$=a_0+a_1x+a_2x^2+...=f(x).$$

Hence $F[x]$ is a vector space over the field F.

Example 5. *Let S be any non-empty set and let F be any field. Let V be the set of all functions from S to F i.e., let*

$$V=\{f : f : S \to F\}.$$

Let us define sum of two elements f and g in V as follows :—

$$(f+g)(x)=f(x)+g(x) \ \forall \ x \in S.$$

Also let us define scalar multiplication of an element f in V by an element c in F as follows :

$$(cf)(x)=cf(x) \ \forall \ x \in S.$$

Then $V(F)$ is a vector space. **(Marathwada 1971)**

Solution 1. We have $\forall \ x \in S$, $(f+g)(x)=f(x)+g(x)$. Since $f(x)$ and $g(x)$ are in F and F is a field, therefore $f(x)+g(x)$

is also in F. Thus $f+g$ is also a function from S to F. Therefore $f+g \in V$ for all $f, g \in V$.

Associativity of addition. We have

$$[(f+g)+h](x)=(f+g)(x)+h(x) \quad \text{[by def.]}$$
$$=[f(x)+g(x)]+h(x) \quad \text{[by def.]}$$
$$=f(x)+[g(x)+h(x)]$$

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

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

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

Commutativity of addition. We have

$$(f+g)(x)=f(x)+g(x)$$
$$=g(x)+f(x) \quad [\because \text{ addition in } F \text{ is commutative}]$$
$$=(g+f)(x).$$

$\therefore$ $f+g=g+f$.

Existence of additive identity. Let us define a function $\hat{0} : S \to F$ such that $\hat{0}(x)=0 \ \forall \ x \in S$.

Then $\hat{0} \in V$ and it is called **zero function.**

We have $(f+\hat{0})(x)=f(x)+\hat{0}(x)=f(x)+0=f(x)$.

$$\therefore \quad f+\hat{0}=f.$$

$\therefore$ the function $\hat{0}$ is the additive identity.

Existence of additive inverse. Let $f \in V$. Let us define a function $-f : S \to F$ by the formula

$$(-f)(x)=-[f(x)] \ \forall \ x \in S.$$

Then $-f \in V$ and we have

$$[f+(-f)](x)=f(x)+[(-f)(x)]=f(x)+[-f(x)]$$
$$=f(x)-f(x)=0=\hat{0}(x).$$

$\therefore$ $f+(-f)=\hat{0}$.

$\therefore$ the function $-f$ is the additive inverse of f.

Thus V is an abelian group with respect to addition composition.

2. If $c \in F$ and $f \in V$, then $\forall \ x \in S$, we have

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

Now $f(x) \in F$ and $c \in F$. Therefore $cf(x)$ is in F. Then cf

Scanned with CamScanner

is a function from S to F. Therefore

$$cf \in V \text{ for all } c \in F \text{ and for all } f \in V.$$

Thus V is closed with respect to scalar multiplication.

3. We observe that

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

$$[c\,(f+g)]\,(x)=c\,[(f+g)\,(x)]=c\,[f\,(x)+g\,(x)]=cf\,(x)+cg\,(x)$$
$$=(cf)\,(x)+(cg)\,(x)=(cf+cg)\,(x).$$

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

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

$$[(c_1+c_2)\,f\,]\,(x)=(c_1+c_2)\,f\,(x)=c_1\,f\,(x)+c_2\,f\,(x)$$
$$=(c_1\,f\,)\,(x)+(c_2\,f\,)\,(x)=(c_1\,f+c_2\,f\,)\,(x).$$

$$\therefore \quad (c_1+c_2)\,f=c_1\,f+c_2\,f.$$

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

$$[(c_1\,c_2)\,f\,]\,(x)=(c_1\,c_2)\,f\,(x)=c_1\,[c_2\,f\,(x)]=c_1\,[(c_2\,f\,)\,(x)]$$
$$=[c_1\,(c_2\,f\,)]\,(x).$$

$$\therefore \quad (c_1\,c_2)\,f=c_1\,(c_2\,f\,).$$

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

$$(1\,f\,)\,(x)=1\,f\,(x)=f\,(x).$$

$$\therefore \quad 1\,f=f.$$

Hence V is a vector space over F.

Example 6. *The vector space of all real valued continuous (differentiable or integrable) functions defined in some interval* [0, 1].

(Nagarjuna 1980)

Solution. If f is a real valued function in the interval $[0, 1]$, then we mean that $f\,(x)$ is a real number $\forall\; x \in [0, 1]$. Let V denote the set of all real valued continuous functions defined in the interval $[0, 1]$. Then V is a vector space over the field $\mathbf{R}$ of real numbers with vector addition and scalar multiplication defined as below :

$$(f+g)\,(x)=f\,(x)+g\,(x) \qquad \forall\; f, g \in V$$

and

$$(af\,)\,(x)=af\,(x) \quad \forall\; a \in \mathbf{R},\; \forall\; f \in V.$$

The sum of two continuous functions is also a continuous function. Therefore if $f, g \in V$, then $f+g \in V$. Thus V is closed with respect to addition of vectors. Further if $a \in \mathbf{R}$ and f is a real valued continuous function in the interval $[0, 1]$, then af is also a real valued continuous function in the interval $[0, 1]$. Therefore V is closed with respect to scalar multiplication. To verify the other postulates of a vector space proceed as in Example 5.

Example 7. *The set of all convergent sequences is a vector space over the field of real numbers.*

Scanned with CamScanner

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

Let $\alpha=\{\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots\}=\{\alpha_n\}$, $\beta=\{\beta_1, \beta_2, \ldots, \beta_n, \ldots\}=\{\beta_n\}$ and $\gamma=\{\gamma_1, \gamma_2, \ldots, \gamma_n, \ldots\}=\{\gamma_n\}$ be any three convergent sequences.

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

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

(ii) Commutativity of addition. We have

$$\alpha+\beta=\{\alpha_n\}+\{\beta_n\}=\{\alpha_n+\beta_n\}=\{\beta_n+\alpha_n\}=\{\beta_n\}+\{\alpha_n\}=\beta+\alpha.$$

(iii) Associativity of addition. We have

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

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

(v) Existence of additive inverse. For every sequence $\{\alpha_n\}$ there exists a sequence $\{-\alpha_n\}$ such that

$$\{\alpha_n\}+\{-\alpha_n\}=\{\alpha_n-\alpha_n\}=\{0\}=\text{additive identity}.$$

$\therefore$ $(V, +)$ is an abelian group.

2. V is closed for scalar multiplication.

Let a be any scalar *i.e.*, a be any real number. Then $a\alpha=a\{\alpha_n\}=\{a\alpha_n\}$ which is also a convergent sequence because

$$\lim_{n\to\infty} a\alpha_n=a \lim_{n\to\infty} \alpha_n.$$

Thus V is closed for scalar multiplication.

3. Laws of scalar multiplication. Let $a, b \in \mathbf{R}$. We have

(i) $$\begin{aligned}a(\alpha+\beta)&=a[\{\alpha_n\}+\{\beta_n\}]=a\{\alpha_n+\beta_n\}\\&=\{a(\alpha_n+\beta_n)\}=\{a\alpha_n+a\beta_n\}=\{a\alpha_n\}+\{a\beta_n\}\\&=a\{\alpha_n\}+a\{\beta_n\}=a\alpha+a\beta.\end{aligned}$$

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

(iii) $$\begin{aligned}(ab)\alpha&=(ab)\{\alpha_n\}=\{(ab)\alpha_n\}=\{a(b\alpha_n)\}\\&=a\{b\alpha_n\}=a[b\{\alpha_n\}]=a(b\alpha).\end{aligned}$$

(iv) $1\alpha=1\{\alpha_n\}=\{1\alpha_n\}=\{\alpha_n\}=\alpha.$

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

Example 8. *Prove that the set of all vectors in a plane over the field of real numbers is a vector space.*

Solution. Let V be the set of all vectors in a plane defined as directed line segments. Let $\mathbf{R}$ be the field of real numbers whose elements will be scalars

Let $\alpha, \beta \in V$. If $\alpha=\overrightarrow{AB}$ and $\beta=\overrightarrow{BC}$, then we define $\alpha+\beta=\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC}$. Since $\overrightarrow{AC}$ is also a vector, therefore $\alpha, \beta \in V \Rightarrow \alpha+\beta \in V$ and thus V is closed for addition of vectors. Also from our knowledge of Vector Algebra we know that addition of vectors on the set V is commutative as well as associative. The zero vector $0=\overrightarrow{AA}$ is identity for addition of vectors If $\alpha=\overrightarrow{AB}$, then the vector $-\alpha=\overrightarrow{BA}$ is the additive inverse of α because $-\alpha+\alpha=\overrightarrow{BA}+\overrightarrow{AB}=\overrightarrow{BB}=$ the zero vector *i.e.*, the identity for addition of vectors.

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

If $\alpha \in V$ and $m \in \mathbf{R}$ *i.e.*, m is any scalar, then the scalar multiplication $m\alpha$ is defined as a vector whose direction is that of α or opposite to that of α according as m is $+$ive or $-$ive and $|m\alpha|=|m|.|\alpha|$.

Since $m \in \mathbf{R}, \alpha \in V \Rightarrow m\alpha \in V$, therefore V is closed for scalar multiplication.

Now if $a, b \in \mathbf{R}$ and $\alpha, \beta \in V$, then from our knowledge of Vector Algebra we know that

$a(\alpha+\beta)=a\alpha+a\beta$, $(a+b)\alpha=a\alpha+b\alpha$ and $(ab)\alpha=a(b\alpha)$.

Also if α is any vector and 1 is the multiplicative identity of the field $\mathbf{R}$, then by our definition of scalar multiplication the vector 1α is in the direction of the vector α and

$$|1\alpha|=|1|.|\alpha|=1.|\alpha|=|\alpha|.$$

$\therefore$ by our definition of equality of two vectors, we have $1\alpha=\alpha$.

Hence V is a vector space over the field $\mathbf{R}$.

Example 9. *Let V be the set of all pairs (x, y) of real numbers, and let F be the field of real numbers. Define*

$$(x, y)+(x_1, y_1)=(x+x_1, 0)$$
$$c(x, y)=(cx, 0).$$

Is V, with these operations, a vector space over the field of real numbers ?

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

$$(x, y)+(x_1, y_1)=(x, y) \ \forall \ x, y \in \mathbf{R}$$
$$\Rightarrow (x+x_1, 0)=(x, y) \ \forall \ x, y \in \mathbf{R}.$$

But if $y \neq 0$, then we cannot have $(x+x_1, 0)=(x, y)$. Thus there exists no element (x_1, y_1) of V such that

$$(x, y)+(x_1, y_1)=(x, y) \ \forall \ (x, y) \in V.$$

Therefore the identity element does not exist and V is not a vector space over the field **R**.

Example 10. *Let V be the set of all pairs (x, y) of real numbers, and let F be the field of real numbers. Examine in each of the following cases whether V is a vector space over the field of real numbers or not ?*

(*i*) $(x, y)+(x_1, y_1)=(x+x_1, y+y_1)$
$\quad c\,(x, y)=(\,|\,c\,|\,x, |\,c\,|\,y).$

(*ii*) $(x, y)+(x_1, y_1)=(x+x_1, y+y_1)$
$\quad c\,(x, y)=(0, cy).$

(*iii*) $(x, y)+(x_1, y_1)=(x+x_1, y+y_1)$
$\quad c\,(x, y)=(c^2x, c^2y).$

Solution. (i) We shall show that in this case the postulate $(a+b)\,\alpha=a\alpha+b\alpha \ \forall \ a, b \in F$ and $\alpha \in V$ fails.

Let $\alpha=(x, y)$ and $a, b \in \mathbf{R}$. We have

$$(a+b)\,\alpha=(a+b)\,(x, y)=(\,|\,a+b\,|\,x, |\,a+b\,|\,y), \quad \text{by def.} \qquad ...(1)$$

Also
$$a\alpha+b\alpha=a\,(x, y)+b\,(x, y)$$
$$=(\,|\,a\,|\,x, |\,a\,|\,y)+(\,|\,b\,|\,x, |\,b\,|\,y), \quad \text{by def. of scalar multiplication}$$
$$=(\,|\,a\,|\,x+|\,b\,|\,x, |\,a\,|\,y+|\,b\,|\,y), \quad \text{by def. of addition of vectors}$$
$$=(\{\,|\,a\,|+|\,b\,|\,\}\,x, \{\,|\,a\,|+|\,b\,|\,\}\,y). \qquad ...(2)$$

Since $|\,a+b\,| \leqslant |\,a\,|+|\,b\,|$, therefore from (1) and (2), we conclude that in general $(a+b)\,\alpha \neq a\alpha+b\alpha$. Hence $V(\mathbf{R})$ is not a vector space.

(ii) We shall show that in this case the postulate $1\alpha=\alpha \ \forall \ \alpha \in V$ fails. Let $\alpha=(x, y)$ where $x, y \in \mathbf{R}$. By definition of scalar multiplication we have $1\alpha=1\,(x, y)=(0, 1y)=(0, y)$.

Scanned with CamScanner

But $(0, y) \neq (x, y)$ if $x \neq 0$. Thus there exists $\alpha \in V$ such that $1\alpha \neq \alpha$. Hence $V(\mathbf{R})$ is not a vector space.

(iii) Show that in this case the postulate $(a+b)\alpha = a\alpha + b\alpha$ $\forall\, a, b \in F$ and $\alpha \in V$ fails. Note that in general

$$(a+b)^2 \neq a^2+b^2.$$

Ex. 11. *Let* $\mathbf{R}$ *be the field of real numbers and let* P_n *be the set of all polynomials (of degree at most n) over the field* $\mathbf{R}$. *Prove that* P_n *is a vector space over the field* $\mathbf{R}$.

Sol. Here P_n is the set of all polynomials of degree at most n over the field $\mathbf{R}$. The set P_n also includes the zero polynomial.

Thus $P_n = \{ f(x) : f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n,$
where $a_0, a_1, a_2, \ldots, a_n \in \mathbf{R}\}$.

If $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$
and $g(x) = b_0 + b_1 x + b_2 x^2 + \ldots + b_n x^n$
be any two members of P_n, then

$$f(x)+g(x) = (a_0+b_0)+(a_1+b_1)\,x+\ldots+(a_n+b_n)\,x^n$$

is also a member of P_n because it is also a polynomial of degree at most n over the field $\mathbf{R}$.

Thus P_n is closed for addition of polynomials.

Also we know that addition of polynomials is commutative as well as associative. The zero polynomial 0 is a member of P_n and is identity for addition of polynomials.

Also if $f(x) = a_0 + a_1 x + \ldots + a_n x^n \in P_n$,
then $-f(x) = -a_0 - a_1 x - \ldots - a_n x^n \in P_n$ because it is also a nomial of degree at most n over the field $\mathbf{R}$.

We have $-f(x)+f(x) =$ the zero polynomial.

$\therefore$ the polynomial $-f(x)$ is the inverse of $f(x)$ for addition of polynomials.

Hence P_n is an abelian group for addition of polynomials.

Now if $f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$ is any member of P_n and $c \in \mathbf{R}$, we define scalar multiplication $cf(x)$ by the relation

$$cf(x) = ca_0 + (ca_1)\,x + (ca_2)\,x^2 + \ldots + (ca_n)\,x^n.$$

Obviously $cf(x) \in P_n$ because it is also a polynomial of degree at most n over the field $\mathbf{R}$. Thus P_n is closed for scalar multiplication.

Now if $a, b \in \mathbf{R}$ and $f(x), g(x) \in P_n$, we have

$$a\,[f(x)+g(x)] = af(x)+ag(x),$$
$$(a+b)\,f(x) = af(x)+bg(x)$$

Scanned with CamScanner

and $(ab) f(x)=a[bf(x)]$ as can be easily shown.

Also $1f(x)=f(x) \ \forall \ f(x) \in P_n$.

Hence P_n is a vector space over the field **R**.

Ex. 12. *How many elements are there in the vector space of polynomials of degree at most n in which the coefficients are the elements of the field* **I** (p) *over the field* **I** (p), *p being a prime number ?* **(Meerut 1975)**

Sol. The field **I** (p) is the field

$(\{0, 1, 2, \ldots, p-1\}, +_p, \times_p)$.

The number of distinct elements in the field **I** (p) is p.

If $f(x)$ is a polynomial of degree at most n over the field **I** (p), then

$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_nx^n, \text{ where } a_0, a_1, a_2, \ldots, a_n \in \mathbf{I}(p).$$

Now in the polynomial $f(x)$, the coefficient of each of the $n+1$ terms $a_0, a_1x, a_2x^2, \ldots, a_nx^n$ can be filled in p ways because any of the p elements of the field **I** (p) can be filled there.

Thus we can have $p \times p \times p \times \ldots$ upto $(n+1)$ times *i.e.*, p^{n+1} distinct polynomials of degree at most n over the field **I** (p). Hence if P_n is the vector space of polynomials of degree at most n in which the coefficients are the elements of the field **I** (p) over the field **I** (p), then P_n has p^{n+1} distinct elements.

Exercises

1. In the axiom $(c_1+c_2)\,\alpha=c_1\alpha+c_2\alpha$ of a vector space what operation does each plus sign represent ? **(Meerut 1976)**
2. In the axiom $(c_1c_2)\,\alpha=c_1\,(c_2\alpha)$ of a vector space what operation does each product represent ?
3. What is the zero vector in the vector space $\mathbf{R}^4$?
4. Is the set of all polynomials in x of degree $\leqslant 2$ a vector space ?
   Ans. Yes.
5. Is the set of all non-zero polynomials in x of degree 2 a vector space ? **(Meerut 1977)**
   Ans. No.
6. Show that any field F may be considered as a vector space over F if scalar multiplication is identified with the field multiplication.
7. Show that the complex field **C** is a vector space over the real field **R**. **(Kumayon 1987)**
8. Prove that the set $V=\{(a, b) : a, b \in \mathbf{R}\}$ is a vector space over

Scanned with CamScanner

the field **R** for the compositions of addition and scalar multiplication defined as under :

$(a, b)+(c, d)=(a+c, b+d)$ **(Meerut 1987, Garhwal 86)**

$k(a, b)=(ka, kb)$.

9. Let V be the set of all pairs (x, y) of real numbers and let F be the field of real numbers. Define

$$(x, y)+(x_1, y_1)=(x+x_1, y+y_1)$$

$$c(x, y)=(cx, y).$$

Show that with these operations V is not a vector space over the field of real numbers. **(Meerut 1976)**

10. Let V be the set of all pairs (x, y) of real numbers and let F be the field of real numbers. Define

$$(x, y)+(x_1, y_1)=(3y+3y_1, -x-x_1)$$

$$c(x, y)=(3cy, -cx).$$

Verify that V, with these operations, is not a vector space over the field of real numbers. **(Meerut 1981, 83, 90)**

§ 8. General properties of vector spaces.

Theorem 1. *Let $V(F)$ be a vector space and* **0** *be the zero vector of V. Then*

(*i*) $a\mathbf{0}=\mathbf{0}\ \forall\ a \in F$.

(Andhra 1992; Nagarjuna 91; Meerut 89; I.A.S. 74)

(*ii*) $0\alpha=\mathbf{0}\ \forall\ \alpha \in V$. **(Nagarjuna 1991; Andhra 92; Meerut 90, Kanpur 80; I.A.S. 74)**

(*iii*) $a(-\alpha)=-(a\alpha)\ \forall\ a \in F,\ \forall\ \alpha \in V$.

(Andhra 1992; Meerut 90)

(*iv*) $(-a)\alpha=-(a\alpha)\ \forall\ a \in F,\ \forall\ \alpha \in V$.

(Andhra 1992; Meerut 86, Kanpur 80; I.A.S. 74)

(*v*) $a(\alpha-\beta)=a\alpha-a\beta\ \forall\ a \in F$ *and* $\forall\ \alpha, \beta \in V$.

(Meerut 1989)

(*vi*) $a\alpha=\mathbf{0} \Rightarrow a=0$ *or* $\alpha=\mathbf{0}$. **(Meerut 1989; I.A.S. 74)**

Proof. (i) We have $a\mathbf{0}=a(\mathbf{0}+\mathbf{0})$ $[\because\ \mathbf{0}=\mathbf{0}+\mathbf{0}]$

$=a\mathbf{0}+a\mathbf{0}$

$\therefore\ \mathbf{0}+a\mathbf{0}=a\mathbf{0}+a\mathbf{0}$ $[\because\ a\mathbf{0} \in V$ and $\mathbf{0}+a\mathbf{0}=a\mathbf{0}]$.

Now V is an abelian group with respect to addition.

Therefore by right cancellation law in V, we get $\mathbf{0}=a\mathbf{0}$.

(ii) We have $0\alpha=(0+0)\alpha$ $[\because\ 0=0+0]$

$=0\alpha+0\alpha$.

$\therefore\ \mathbf{0}+0\alpha=0\alpha+0\alpha$ $[\because\ 0\alpha \in V$ and $\mathbf{0}+0\alpha=0\alpha]$.

Now V is an abelian group with respect to addition of vectors. Therefore by right cancellation law in V, we get $\mathbf{0}=0\alpha$.

(iii) We have $a[\alpha+(-\alpha)]=a\alpha+a(-\alpha)$

$\Rightarrow a\mathbf{0}=a\alpha+a(-\alpha)$

$\Rightarrow \mathbf{0}=a\alpha+a(-\alpha)$ $\quad [\because \; a\mathbf{0}=\mathbf{0}]$

$\Rightarrow a(-\alpha)$ is the additive inverse of $a\alpha$

$\Rightarrow a(-\alpha)=-(a\alpha)$.

(iv) We have $[a+(-a)]\,\alpha=a\alpha+(-a)\,\alpha$

$\Rightarrow 0\alpha=a\alpha+(-a)\,\alpha$

$\Rightarrow \mathbf{0}=a\alpha+(-a)\,\alpha$

$\Rightarrow (-a)\,\alpha$ is the additive inverse of $a\alpha$

$\Rightarrow (-a)\,\alpha=-(a\alpha)$.

(v) We have $a(\alpha-\beta)=a\,[\alpha+(-\beta)]=a\alpha+a(-\beta)$

$=a\alpha+[-(a\beta)]$ $\quad [\because \; a(-\beta)=-(a\beta)]$

$=a\alpha-a\beta$.

(vi) Let $a\alpha=\mathbf{0}$ and $a\neq 0$. Then a^{-1} exists because a is a non-zero element of the field F.

$\therefore \; a\alpha=\mathbf{0} \Rightarrow a^{-1}(a\alpha)=a^{-1}\mathbf{0} \Rightarrow (a^{-1}a)\,\alpha=\mathbf{0} \Rightarrow 1\alpha=\mathbf{0} \Rightarrow \alpha=\mathbf{0}$.

Again let $a\alpha=\mathbf{0}$ and $\alpha\neq\mathbf{0}$. Then to prove that $a=0$. Suppose $a\neq 0$. Then a^{-1} exists.

$\therefore \; a\alpha=0 \Rightarrow a^{-1}(a\alpha)=a^{-1}\mathbf{0} \Rightarrow (a^{-1}a)\,\alpha=\mathbf{0} \Rightarrow 1\alpha=\mathbf{0} \Rightarrow \alpha=\mathbf{0}$.

Thus we get a contradiction that α must be a zero vector. Therefore a must be equal to 0. Hence $\alpha\neq\mathbf{0}$ and $a\alpha=\mathbf{0} \Rightarrow a=0$.

Theorem 2. *Let $V(F)$ be a vector space. Then*

(*i*) *If $a, b \in F$ and α is a non-zero vector of V, we have*

$$a\alpha=b\alpha \Rightarrow a=b. \qquad \textbf{(Poona 1972)}$$

(*ii*) *If $\alpha, \beta \in V$ and a is a non-zero element of F, we have*

$$a\alpha=a\beta \Rightarrow \alpha=\beta. \qquad \textbf{(Allahabad 1978)}$$

Proof. (i) We have $a\alpha=b\alpha$

$\Rightarrow a\alpha-b\alpha=\mathbf{0}$

$\Rightarrow (a-b)\,\alpha=\mathbf{0}$

$\Rightarrow a-b=0$ since $\alpha\neq\mathbf{0}$

$\Rightarrow a=b$

(ii) We have $a\alpha=a\beta$

$\Rightarrow a\alpha-a\beta=\mathbf{0}$

$\Rightarrow a(\alpha-\beta)=\mathbf{0}$

$\Rightarrow \alpha-\beta=\mathbf{0}$, since $a\neq 0$

$\Rightarrow \alpha=\beta$.

Scanned with CamScanner

§ 9. Vector Subspaces. Definition.

Let V be a vector space over the field F and let $W \subseteq V$. Then W is called a subspace of V if W itself is a vector space over F with respect to the operations of vector addition and scalar multiplication in V. (Meerut 1993P; S.V.U. Tirupati 93; Poona 72; Madras 81; Nagarjuna 90)

Theorem 1. *The necessary and sufficient condition for a non-empty subset W of a vector space $V(F)$ to be a subspace of V is that W is closed under vector addition and scalar multiplication in V.*

Proof. If W itself is a vector space over F with respect to vector addition and scalar multiplication in V, then W must be closed with respect to these two compositions. Hence the condition is necessary.

The condition is sufficient. Now suppose that W is a non-empty subset of V and W is closed under vector addition and scalar multiplication in V.

Let $\alpha \in W$. If 1 is the unity element of F, then $-1 \in F$. Now W is closed under scalar multiplication. Therefore

$$-1 \in F, \alpha \in W \Rightarrow (-1)\,\alpha \in W \Rightarrow -(1\alpha) \in W$$

$$\Rightarrow -\alpha \in W \qquad [\because \ \alpha \in W \Rightarrow \alpha \in V \text{ and } 1\alpha = \alpha \text{ in } V].$$

Thus the additive inverse of each element of W is also in W.

Now W is closed under vector addition.

Therefore $\alpha \in W, -\alpha \in W \Rightarrow \alpha + (-\alpha) \in W$

$\Rightarrow 0 \in W$ where 0 is the zero vector of V.

Hence the zero vector of V is also the zero vector of W. Since the elements of W are also the elements of V, therefore vector addition will be commutative as well as associative in W. Hence W is an abelian group with respect to vector addition. Also it is given that W is closed under scalar multiplication. The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset.

Hence W itself is a vector space for the two compositions.

$\therefore$ W is a subspace of V.

Theorem 2. *The necessary and sufficient conditions for a non-empty subset W of a vector space $V(F)$ to be a subspace of V are*

(*i*) $\alpha \in W, \beta \in W \Rightarrow \alpha - \beta \in W$,

(*ii*) $a \in F, \alpha \in W \Rightarrow a\alpha \in W$.

Scanned with CamScanner

Proof. The conditions are necessary. If W is a subspace of V, then W is an abelian group with respect to vector addition. Therefore $\alpha \in W, \beta \in W \Rightarrow \alpha-\beta \in W$. Also W must be closed under scalar multiplication. Therefore condition (ii) is also necessary.

The conditions are sufficient. Now suppose W is a non-empty subset of V satisfying the two given conditions. From condition (i) we have

$$\alpha \in W, \alpha \in W \Rightarrow \alpha-\alpha \in W \Rightarrow \mathbf{0} \in W.$$

Thus the zero vector of V belongs to W and it will also be the zero vector of W.

Now $\quad \mathbf{0} \in W, \alpha \in W \Rightarrow \mathbf{0}-\alpha \in W \Rightarrow -\alpha \in W.$

Thus the additive inverse of each element of W is also in W.

Again $\alpha \in W, \beta \in W \Rightarrow \alpha \in W, -\beta \in W$

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

Thus W is closed with respect to vector addition.

Since the elements of W are also the elements of V, therefore vector addition will be commutative as well as associative in W. Hence W is an abelian group under vector addition. Also from condition (ii), W is closed under scalar multiplication. The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset. Hence W is a subspace of V.

Theorem 3. *The necessary and sufficient condition for a non-empty subset W of a vector space $V(F)$ to be a subspace of V is*

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

(Nagarjuna 1990; S.V.U. Tirupati 93; Allahabad 76)

Proof. The condition is necessary. If W is a subspace of V, then W must be closed under scalar multiplication and vector addition.

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

and $\quad b \in F, \beta \in W \Rightarrow b\beta \in W.$

Now $a\alpha \in W, b\beta \in W \Rightarrow a\alpha+b\beta \in W$. Hence the condition is necessary.

The condition is sufficient. Now suppose W is a non-empty subset of V satisfying the given condition *i.e.*, $a, b \in F$ and $\alpha, \beta \in W \Rightarrow a\alpha+b\beta \in W$.

Taking $a=1, b=1$, we see that if $\alpha, \beta \in W$, then

$$1\alpha+1\beta \in W \Rightarrow \alpha+\beta \in W.$$

$$[\because \alpha \in W \Rightarrow \alpha \in V \text{ and } 1\alpha=\alpha \text{ in } V]$$

Scanned with CamScanner

Thus W is closed under vector addition.

Now taking $a=-1, b=0$, we see that if $\alpha \in W$ then

$(-1)\alpha+0\alpha \in W$ [In place of β we have taken

$\Rightarrow -(1\alpha)+0 \in W \Rightarrow -\alpha \in W$.

Thus the additive inverse of each element of W is also in W.

Taking $a=0, b=0$, we see that if $\alpha \in W$ then

$0\alpha+0\alpha \in W \Rightarrow 0+0 \in W \Rightarrow \mathbf{0} \in W$.

Thus the zero vector of V belongs to W. It will also be th zero vector of W.

Since the elements of W are also the elements of V, therefore vector addition will be associative as well as commutative in V. Thus W is an abelian group with respect to vector addition.

Now taking $\beta=0$, we see that if $a, b \in F$ and $\alpha \in W$, then

$a\alpha+b0 \in W$ i.e., $a\alpha+0 \in W$ i.e., $a\alpha \in W$.

Thus W is closed under scalar multiplication.

The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset.

Hence $W(F)$ is a subspace of $V(F)$.

Theorem 4. *A non-empty subset W of a vector space $V(F)$ is a subspace of V if and only if for each pair of vectors α, β in W and each scalar a in F the vector $a\alpha+\beta$ is again in W.*

Proof. The condition is necessary. If W is a subset of V, then W must be closed with respect to scalar multiplication and as well as with respect to vector addition. Therefore

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

Further

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

Hence the condition is necessary.

The condition is sufficient. It is given that W is a non-empty subset of V and $a \in F$, $\alpha, \beta \in W \Rightarrow a\alpha+\beta \in W$. We are to prove that W is a subspace of V.

(i) Since W is non-empty, therefore there is at least one vector in W, say γ. Now $1 \in F \Rightarrow -1 \in F$. Therefore taking $a=-1$, $\alpha=\gamma, \beta=\gamma$, we get from the given condition that

$$(-1)\gamma+\gamma=-(1\gamma)+\gamma=-\gamma+\gamma=\mathbf{0} \text{ is in } W.$$

(ii) Now let $a \in F$, $\alpha \in W$. Since $\mathbf{0}$ is in W, therefore taking $\beta=\mathbf{0}$ in the given condition, we get

$$a\alpha+\mathbf{0}=a\alpha \text{ is in } W.$$

Thus W is closed with respect to scalar multiplication.

Scanned with CamScanner

(iii) Let $\alpha \in W$. Since $-1 \in F$ and W is closed with respect to scalar multiplication, therefore $(-1)\,\alpha = -(1\alpha) = -\alpha$ is in W.

(iv) We have $1 \in F$. If $\alpha, \beta \in W$, then $1\alpha + \beta = \alpha + \beta$ is in W. Thus W is closed with respect to vector addition.

The remaining postulates of a vector space will hold in W since they hold in V of which W is a subset.

Hence W is a subspace of V.

Note. If we are to prove that a subset W of a vector space V is a subspace of V, then either it is sufficient to prove that

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

or it is sufficient to prove that

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

Illustrative Examples

Example 1. Let $V(F)$ be any vector space Then V itself and the subset of V consisting of zero vector only are always subspaces of V. These two are called improper subspaces. If V has any other subspace, then it is called a proper subspace. The subspace of V consisting of zero vector only is called the *zero subspace.*

Example 2. *The set W of ordered triads $(a_1, a_2, 0)$, where $a_1, a_2 \in F$ is a subspace of $V_3(F)$.* **(Meerut 1986)**

Solution. Let $\alpha = (a_1, a_2, 0)$ and $\beta = (b_1, b_2, 0)$ be any two elements of W. Then $a_1, a_2, b_1, b_2 \in F$. If a, b be any two elements of F, we have

$$\begin{aligned} a\alpha + b\beta &= a\,(a_1, a_2, 0) + b\,(b_1, b_2, 0) \\ &= (aa_1, aa_2, 0) + (bb_1, bb_2, 0) \\ &= (aa_1 + bb_1, aa_2 + bb_2, 0) \in W \end{aligned}$$

since $aa_1 + bb_1, aa_2 + bb_2 \in F$ and the last co-ordinate of this triad is zero.

Hence W is a subspace of $V_3(F)$.

Example 3. *Let V be the vector space of all polynomials in an indeterminate x over a field F. Let W be a subset of V consisting of all polynomials of degree $\leqslant n$. Then W is a subspace of V.*

Solution. Let α and β be any two elements of W. Then α, β are polynomials over F of degree $\leqslant n$. If a, b are any two elements of F, then $a\alpha + b\beta$ will also be a polynomial of degree $\leqslant n$. Therefore $a\alpha + b\beta \in W$. Hence W is a subspace of V.

Example 4. *If a_1, a_2, a_3 are fixed elements of a field F, then the set W of all ordered triads (x_1, x_2, x_3) of elements of F, such that*

$$a_1x_1 + a_2x_2 + a_3x_3 = 0,$$

is a subspace of $V_3(F)$.

(Poona 1972)

Scanned with CamScanner

Solution. Let $\alpha=(x_1, x_2, x_3)$ and $\beta=(y_1, y_2, y_3)$ be any two elements of W. Then $x_1, x_2, x_3, y_1, y_2, y_3$ are elements of F and are such that

$$a_1x_1+a_2x_2+a_3x_3=0 \quad \ldots(1)$$

and

$$a_1y_1+a_2y_2+a_3y_3=0 \quad \ldots(2)$$

If a, b be any two elements of F, we have

$$a\alpha+b\beta=a\,(x_1, x_2, x_3)+b\,(y_1, y_2, y_3)$$

$$=(ax_1, ax_2, ax_3)+(by_1, by_2, by_3)=(ax_1+by_1, ax_2+by_2, ax_3+by_3).$$

Now $a_1(ax_1+by_1)+a_2\,(ax_2+by_2)+a_3\,(ax_3+by_3)$

$$=a\,(a_1x_1+a_2x_2+a_3x_3)+b\,(a_1y_1+a_2y_2+a_3y_3)$$

$$=a0+b0=0 \qquad \text{[by (1) and (2)]}$$

$\therefore \quad a\alpha+b\beta=(ax_1+by_1, ax_2+by_2, ax_3+by_3)\in W.$

Hence W is a subspace of $V_3(F)$.

Example 5. *Prove that the set of all solutions (a, b, c) of the equation $a+b+2c=0$ is a subspace of the vector space V_3 (R).*

(Meerut 1989)

Sol. Let $W=\{(a, b, c) : a, b, c \in \mathbf{R} \text{ and } a+b+2c=0\}$.

To prove that W is a subspace of $V_3(\mathbf{R})$ or $\mathbf{R}^3$.

Let $\alpha=(a_1, b_1, c_1)$ and $\beta=(a_2, b_2, c_2)$ be any two elements of W. Then

$$a_1+b_1+2c_1=0 \quad \ldots(1)$$

and

$$a_2+b_2+2c_2=0. \quad \ldots(2)$$

If a, b be any two elements of $\mathbf{R}$, we have

$$a\alpha+b\beta=a\,(a_1, b_1, c_1)+b\,(a_2, b_2, c_2)$$

$$=(aa_1, ab_1, ac_1)+(ba_2, bb_2, bc_2)$$

$$=(aa_1+ba_2, ab_1+bb_2, ac_1+bc_2).$$

Now $(aa_1+ba_2)+(ab_1+bb_2)+2\,(ac_1+bc_2)$

$$=a\,(a_1+b_1+2c_1)+b\,(a_2+b_2+2c_2)$$

$$=a.0+b.0 \qquad \text{[from (1) and (2)]}$$

$$=0.$$

$\therefore \quad a\alpha+b\beta=(aa_1+ba_2, ab_1+bb_2, ac_1+bc_2) \in W.$

Thus $\alpha, \beta \in W$ and $a, b \in \mathbf{R} \Rightarrow a\alpha+b\beta \in W$.

Hence W is a subspace of $V_3(\mathbf{R})$.

Example 6. *Show that the set W of the elements of the vector space V_3 (R) of the form $(x+2y, y, -x+3y)$ where $x, y \in \mathbf{R}$ is a subspace of V_3 (R).*

(Meerut 1974)

Solution. Let $W=\{(x+2y, y, -x+3y) : x, y \in \mathbf{R}\}$.

To prove that W is a subspace of V_3 (R).

Let $\alpha=(x_1+2y_1, y_1, -x_1+3y_1)$ and $\beta=(x_2+2y_2, y_2, -x_2+3y_2)$ be any two elements of W.

If a, b be any two elements of $\mathbf{R}$, we have

$a\alpha+b\beta=a\,(x_1+2y_1,\ y_1,\ -x_1+3y_1)+b\,(x_2+2y_2,\ y_2,\ -x_2+3y_2)$

$=(ax_1+2ay_1,\ ay_1,\ -ax_1+3ay_1)+(bx_2+2by_2,\ by_2,\ -bx_2+3by_2)$

$=(ax_1+2ay_1+bx_2+2by_2,\ ay_1+by_2,\ -ax_1+3ay_1-bx_2+3by_2)$

$=([ax_1+bx_2]+2\,[ay_1+by_2],\ ay_1+by_2,\ -[ax_1+bx_2]+3\,[ay_1+by_2])$

which is in W because it is of the form $(x+2y,\ y,\ -x+3y)$.

Here in place of y we have ay_1+by_2 and in place of x we have ax_1+bx_2.

Thus $\alpha, \beta \in W$ and $a, b \in \mathbf{R} \Rightarrow a\alpha+b\beta \in W$.

Hence W is a subspace of $V_3\,(\mathbf{R})$.

Example 7. *Which of the following sets of vectors* $\alpha=(a_1, a_2, \ldots, a_n)$ *in* $\mathbf{R}^n$ *are subspaces of* $\mathbf{R}^n$ $(n \geqslant 3)$?

(*i*) *all* α *such that* $a_1 \leqslant 0$;

(*ii*) *all* α *such that* a_3 *is an integer* ;

(*iii*) *all* α *such that* $a_2+4a_3=0$;

(*iv*) *all* α *such that* $a_1+a_2+\ldots+a_n=k$ (*a given constant*).

Solution. (i) Let $W=\{\alpha : \alpha \in \mathbf{R}^n \text{ and } a_1 \leqslant 0\}$.

If we take $a_1=-3$, then $a_1 < 0$ and so

$$\alpha=(-3, a_2, \ldots, a_n) \in W.$$

Now if we take $a=-2$, then

$$a\alpha=(6, -2a_2, \ldots, -2a_n).$$

Since the first coordinate of $a\alpha$ is 6 which is > 0, therefore $a\alpha \notin W$.

Thus $\alpha \in W$, $a \in \mathbf{R}$ but $a\alpha \notin W$. Therefore W is not closed for scalar multiplication and so W is not a subspace of $\mathbf{R}^n$.

(ii) Let $W=\{\alpha : \alpha \in \mathbf{R}^n \text{ and } a_3 \text{ is an integer}\}$.

If we take $a_3=5$, then a_3 is an integer and so

$$\alpha=(a_1, a_2, 5, \ldots, a_n) \in W.$$

Now if we take $a=\frac{1}{2}$, then

$$a\alpha=(\tfrac{1}{2}a_1, \tfrac{1}{2}a_2, 5/2, \ldots, \tfrac{1}{2}a_n).$$

Since the third coordinate of $a\alpha$ is 5/2 which is not an integer therefore $a\alpha \notin W$.

Thus $\alpha \in W$, $a \in \mathbf{R}$ but $a\alpha \notin W$. Therefore W is not closed for scalar multiplication and so W is not a subspace of $\mathbf{R}^n$.

(iii) Let $W=\{\alpha : \alpha \in \mathbf{R}^n \text{ and } a_2+4a_3=0\}$.

Let $\alpha=(a_1, \ldots, a_n)$ and $\beta=(b_1, \ldots, b_n)$ be any two members of W. Then $a_2+4a_3=0$ and $b_2+4b_3=0$.

If $a, b \in \mathbf{R}$, then $a\alpha+b\beta=(aa_1+bb_1, \ldots, aa_n+bb_n)$.

We have $\quad (aa_2+bb_2)+4\,(aa_3+bb_3)$
$$=a\,(a_2+4a_3)+b\,(b_2+4b_3)=a.0+b.0=0.$$

Thus according to the definition of W, $a\alpha+b\beta \in W$.

In this way $\alpha, \beta \in W$ and $a, b \in \mathbf{R} \Rightarrow a\alpha+b\beta \in W$. Hence W is a subspace of $\mathbf{R}^n$.

(iv) A subspace if $k=0$ and not a subspace if $k \neq 0$.

Example 8. *Let* $\mathbf{R}$ *be the field of real numbers. Which of the following are subspaces of* $V_3(\mathbf{R})$:

(*i*) $\{(x, 2y, 3z) : x, y, z \in \mathbf{R}\}$. **(Meerut 1990)**

(*ii*) $\{(x, x, x) : x \in \mathbf{R}\}$.

(*iii*) $\{(x, y, z) : x, y, z$ *are rational numbers*$\}$. **(Meerut 1989 P)**

Solution. (i) Let $W=\{(x, 2y, 3z) : x, y, z \in \mathbf{R}\}$.

Let $\alpha=(x_1, 2y_1, 3z_1)$ and $\beta=(x_2, 2y_2, 3z_2)$ be any two elements of W. Then $x_1, y_1, z_1, x_2, y_2, z_2$ are all real numbers. If a, b are any two real numbers, then

$$a\alpha+b\beta=a\,(x_1, 2y_1, 3z_1)+b\,(x_2, 2y_2, 3z_2)$$
$$=(ax_1+bx_2, 2ay_1+2by_2, 3az_1+3bz_2)$$
$$=(ax_1+bx_2, 2\,[ay_1+by_2], 3\,[az_1+bz_2])$$

$\in W$ since $\quad ax_1+bx_2, ay_1+by_2, az_1+bz_2$ are real numbers.

Thus $a, b \in \mathbf{R}$ and $\alpha, \beta \in W \Rightarrow a\alpha+b\beta \in W$.

$\therefore$ W is a subspace of $V_3(\mathbf{R})$.

(ii) Let $W=\{(x, x, x) : x \in \mathbf{R}\}$.

Let $\alpha=(x_1, x_1, x_1)$ and $\beta=(x_2, x_2, x_2)$ be any two elements of W. Then x_1, x_2 are real numbers. If a, b are any real numbers, then

$$a\alpha+b\beta=a\,(x_1, x_1, x_1)+b\,(x_2, x_2, x_2)$$
$$=(ax_1+bx_2, ax_1+bx_2, ax_1+bx_2) \in W$$

since $ax_1+bx_2 \in \mathbf{R}$.

Thus W is a subspace of $V_3(\mathbf{R})$.

(iii) Let $W=\{(x, y, z) : x, y, z$ are rational numbers$\}$.

Now $\alpha=(3, 4, 5)$ is an element of W. Also $a=\sqrt{7}$ is an element of $\mathbf{R}$. But $a\alpha=\sqrt{7}\,(3, 4, 5)=(3\sqrt{7}, 4\sqrt{7}, 5\sqrt{7}) \notin W$ since $3\sqrt{7}, 4\sqrt{7}, 5\sqrt{7}$ are not rational numbers.

Therefore W is not closed under scalar multiplication. Hence W is not a subspace of $V_3(\mathbf{R})$.

Example 9. The solution of a system of homogeneous linear equations. *Let* $V(F)$ *be the vector space of all* $n \times 1$ *matrices over the field* F. *Let* A *be an* $m \times n$ *matrix over* F. *Then the set* W *of all*

$n\times 1$ matrices X over F such that $AX=O$ is a subspace of V. Here O is a null matrix of the type $m\times 1$.

Solution. Let $X, Y\in W$. Then X and Y are $n\times 1$ matrices over F such that
$$AX=O,\ AY=O.$$

Let $a\in F$. Then $aX+Y$ is also an $n\times 1$ matrix over F.

We have
$$A\,(aX+Y)=A\,(aX)+AY=a\,(AX)+AY$$
$$=aO+O=O+O=O.$$

Therefore $aX+Y\in W$. Thus
$$a\in F,\ X,\ Y\in W\Rightarrow aX+Y\in W.$$

Hence W is a subspace of V.

Example 10. *Which of the following sets of vectors $\alpha=(a_1,\ldots,a_n)$ in $\mathbf{R}^n$ are subspaces of $\mathbf{R}^n$? ($n\geqslant 3$).*

(i) all α such that $a_1\geqslant 0$;
(ii) all α such that $a_1+3a_2=a_3$;
(iii) all α such that $a_2=a_1^2$;
(iv) all α such that $a_1a_2=0$;
(v) all α such that a_2 is rational. **(Meerut 1972)**

Solution. (i) Let $W=\{\alpha:\alpha\in\mathbf{R}^n$ and $a_1\geqslant 0\}$. Let $\alpha=(a_1,\ldots,a_n)$ and $\beta=(b_1,\ldots,b_n)$ be any two members of W. Then $a_1\geqslant 0$ and $b_1\geqslant 0$. If $a, b\in\mathbf{R}$, then $a\alpha+b\beta=(aa_1+bb_1,\ldots,aa_n+bb_n)$. If a and b are any two real numbers, then aa_1+bb_1 will not necessarily be $\geqslant 0$. For example if we take $a_1=3$, $b_1=3$, $a=-2$ and $b=-2$, then $aa_1+bb_1=-6-6=-12$ which is <0. Thus $\alpha,\beta\in W$ and $a, b\in\mathbf{R}\Rightarrow a\alpha+b\beta\notin W$. Hence W is not a subspace of $\mathbf{R}^n$.

(ii) Let $W=\{\alpha:\alpha\in\mathbf{R}^n$ and $a_1+3a_2=a_3\}$. Let $\alpha=(a_1,\ldots,a_n)$ and $\beta=(b_1,\ldots,b_n)$ be any two members of W. Then $a_1+3a_2=a_3$ and $b_1+3b_2=b_3$. If $a, b\in\mathbf{R}$, then $a\alpha+b\beta=(aa_1+bb_1,\ldots,aa_n+bb_n)$.

We have $(aa_1+bb_1)+3\,(aa_2+bb_2)=a\,(a_1+3a_2)+b\,(b_1+3b_2)$ $=aa_3+bb_3$. Thus according to the definition of W, $a\alpha+b\beta\in W$. In this way $\alpha,\beta\in W$ and $a, b\in\mathbf{R}\Rightarrow a\alpha+b\beta\in W$. Hence W is a subspace of $\mathbf{R}^n$.

(iii) Let $W=\{\alpha:\alpha\in\mathbf{R}^n$ and $a_2=a_1^2\}$ Let $\alpha=(a_1,\ldots,a_n)$ and $\beta=(b_1,\ldots,b_n)$ be any two members of W. Then $a_2=a_1^2$ and $b_2=b_1^2$. If $a, b\in\mathbf{R}$, then $a\alpha+b\beta=(aa_1+bb_1,\ldots,aa_n+bb_n)$. Now aa_2+bb_2 $=aa_1^2+bb_1^2$ which is not necessarily equal to $(aa_1+bb_1)^2$. For example, take $a_1=2$, $a_2=4$, $b_1=3$, $b_2=9$, $a=2$, $b=3$. Then $a_2=a_1^2$ and $b_2=b_1^2$. Also $aa_2+bb_2=8+27=35$ and $(aa_1+bb_1)^2=(13)^2$.

Scanned with CamScanner

Thus $aa_2+bb_2 \neq (aa_1+bb_1)^2$. In this way $a\alpha+b\beta \notin W$. Hence W is not a subspace of $\mathbf{R}^n$.

(iv) Let $W=\{\alpha : \alpha \in \mathbf{R}^n$ and $a_1a_2=0\}$. Let $\alpha=(a_1,\ldots,a_n)$ and $\beta=(b_1,\ldots,b_n)$ be any two members of W. Then $a_1a_2=0$, $b_1b_2=0$. If $a, b \in \mathbf{R}$, then $a\alpha+b\beta=(aa_1+bb_1,\ldots,aa_n+bb_n)$. We have $(aa_1+bb_1)(aa_2+bb_2)=a^2a_1a_2+ab(a_1b_2+a_2b_1)+b^2b_1b_2=ab(a_1b_2+a_2b_1)$ which is not necessarily equal to zero In this way $a\alpha+b\beta$ is not necessarily a member of W. Hence W is not a subspace of $\mathbf{R}^n$.

(v) Let $W=\{\alpha : \alpha \in \mathbf{R}^n$ and a_2 is rational$\}$. Let $\alpha=(a_1,\ldots,a_n)$ and $\beta=(b_1,\ldots,b_n)$ be any two members of W. Then a_2 is rational and b_2 is rational. If $a, b \in \mathbf{R}$, then $a\alpha+b\beta=(aa_1+bb_1,\ldots,aa_n+bb_n)$. Now aa_2+bb_2 is not necessarily rational. For example if we take $a=\sqrt{3}$, $b=\sqrt{7}$, $a_2=3$, $b_2=4$, then aa_2+bb_2 is not rational. Thus in this case $a\alpha+b\beta \notin W$. Hence W is not a subspace of $\mathbf{R}^n$.

Example 11. *Let V be the (real) vector space of all functions f from $\mathbf{R}$ into $\mathbf{R}$. Which of the following sets of functions are subspaces of V?*

(i) *all f such that $f(x^2)=[f(x)]^2$;*
(ii) *all f such that $f(0)=f(1)$;*
(iii) *all f such that $f(3)=1+f(-5)$;*
(iv) *all f such that $f(-1)=0$;*
(v) *all f which are continuous.*

Solution. (i) Let $W=\{f : f \in V$ and $f(x^2)=[f(x)]^2\}$. Let f, g be any two members of W. Then $f(x^2)=[f(x)]^2$ and $g(x^2)=[g(x)]^2$. Let $a, b \in \mathbf{R}$. Then $(af+bg)(x^2)=(af)(x^2)+(bg)(x^2)=af(x^2)+bg(x^2)=a[f(x)]^2+b[g(x)]^2$. Also $[(af+bg)(x)]^2=[af(x)+bg(x)]^2=a^2[f(x)]^2+b^2[g(x)]^2+2ab\, f(x)\, g(x)$. Now $(af+bg)(x^2)$ is not necessarily equal to $[(af+bg)(x)]^2$. Thus $af+bg$ is not necessarily a member of W. Hence W is not a subspace of V.

(ii) Let $f, g \in W$ in this case. Then $f(0)=f(1)$ and $g(0)=g(1)$. Let $a, b \in \mathbf{R}$. Then $(af+bg)(0)=af(0)+bg(0)=af(1)+bg(1)=(af+bg)(1)$. Therefore by definition of W, $af+bg \in W$. Hence W is a subspace of V.

(iii) Let $f, g \in W$ in this case. Then $f(3)=1+f(-5)$ and $g(3)=1+g(-5)$. We have $(af+bg)(3)=af(3)+bg(3)=a[1+f(-5)]+b[1+g(-5)]=a+b+(af+bg)(-5)$ which is not necessarily equal to $1+(af+bg)(-5)$. Hence W is not a subspace of V.

(iv) Let $W=\{f : f \in V$ and $f(-1)=0\}$. Let $f, g \in W$. Then

$f(-1)=0$ and $g(-1)=0$. If $a, b \in \mathbf{R}$, then $(af+bg)(-1)$ $=(af)(-1)+(bg)(-1)=af(-1)+bg(-1)=a(0)+b(0)=0$. Therefore $af+bg \in W$. Hence W is a subspace of V.

(v) If f and g are continuous functions and $a, b \in \mathbf{R}$, then $af+bg$ is also a continuous function. Hence in this case W is a subspace of V.

Example 12. *If a vector space V is the set of all real valued continuous functions over the field of real numbers* $\mathbf{R}$, *then show that the set W of solutions of the differential equation*

$$2\frac{d^2y}{dx^2}-9\frac{dy}{dx}+2y=0$$

is a subspace of V. **(Meerut 1993P)**

Solution. We have $W=\left\{y : 2\dfrac{d^2y}{dx^2}-9\dfrac{dy}{dx}+2y=0\right\}$,

where $y=f(x)$.

Obviously $y=0$ satisfies the given differential equation and as such it belongs to W and thus $W \neq \varnothing$.

Now let $y_1, y_2 \in W$. Then

$$2\frac{d^2y_1}{dx^2}-9\frac{dy_1}{dx}+2y_1=0 \quad \text{...(1)}$$

and $$2\frac{d^2y_2}{dx^2}-9\frac{dy_2}{dx}+2y_2=0. \quad \text{...(2)}$$

Let $a, b \in \mathbf{R}$. If W is to be a subspace then we should show that ay_1+by_2 also belongs to W i.e., it is a solution of the given differential equation.

We have

$$2\frac{d^2}{dx^2}(ay_1+by_2)-9\frac{d}{dx}(ay_1+by_2)+2(ay_1+by_2)$$

$$=2a\frac{d^2y_1}{dx^2}+2b\frac{d^2y_2}{dx^2}-9a\frac{dy_1}{dx}-9b\frac{dy_2}{dx}+2ay_1+2by_2$$

$$=a\left(2\frac{d^2y_1}{dx^2}-9\frac{dy_1}{dx}+2y_1\right)+b\left(2\frac{d^2y_2}{dx^2}-9\frac{dy_2}{dx}+2y_2\right)$$

$=a.0+b.0$, by (1) and (2)

$=0$.

Thus ay_1+by_2 is a solution of the given differential equation and so it belongs to W.

Hence W is a subspace of V.

Scanned with CamScanner

Exercises

1. Let $V = \mathbf{R}^3$ and W be the set of all ordered triads (x, y, z) such that $x - 3y + 4z = 0$. Prove that W is a subspace of $\mathbf{R}^3$. **(Meerut 1992)**

2. Let C be the field of complex numbers and let n be a positive integer ($n \geqslant 2$). Let V be the vector space of all $n \times n$ matrices over C. Which of the following sets of matrices A in V are subspaces of V? **(Meerut 1981)**

(i) all invertible A;

(ii) all non-invertible A;

(iii) all A such that $AB = BA$, where B is some fixed matrix in V.

Ans. (i) not a subspace; (ii) not a subspace; (iii) a subspace.

3. Let V be a vector space of all real $n \times n$ matrices. Prove that the set W consisting of all $n \times n$ real matrices which commute with a given matrix T of V form a subspace of V. **(Meerut 1980)**

4. Let V be the vector space of all 2×2 matrices over the real field R. Show that the subset of V consisting of all matrices A for which $A^2 = A$ is not a subspace of V. **(Meerut 1976)**

5. State whether the following statements are true or false :—

(i) A subspace of $V_3(\mathbf{R})$, where $\mathbf{R}$ is the real field, must always contain the origin. **(Meerut 1977)**

(ii) The set of vectors $\alpha = (x, y) \in V_2(\mathbf{R})$ for which $x^2 = y^2$ is a subspace of $V_2(\mathbf{R})$. **(Meerut 1977)**

(iii) The set of ordered triads (x, y, z) of real numbers with $x > 0$ is a subspace of $V_3(\mathbf{R})$. **(Meerut 1977)**

(iv) The set of ordered triads (x, y, z) of real numbers with $x + y = 0$ is a subspace of $V_3(\mathbf{R})$.

Ans. (i) true; (ii) false; (iii) false; (iv) true.

§ 10. Algebra of subspaces.

Theorem 1. *The intersection of any two subspaces W_1 and W_2 of a vector space $V(F)$ is also a subspace of $V(F)$.*

(Meerut 1990; Andhra 92; Allahabad 78; Nagarjuna 74)

Proof. Since $\mathbf{0} \in W_1$ and W_2 both therefore $W_1 \cap W_2$ is not empty.

Let $\alpha, \beta \in W_1 \cap W_2$ and $a, b \in F$.

Now $\alpha \in W_1 \cap W_2 \Rightarrow \alpha \in W_1$ and $\alpha \in W_2$

and $\quad \beta \in W_1 \cap W_2 \Rightarrow \beta \in W_1$ and $\beta \in W_2$.

Since W_1 is a subspace, therefore

Scanned with CamScanner

$a, b \in F$ and $\alpha, \beta \in W_1 \Rightarrow a\alpha+b\beta \in W_1$.

Similarly $a, b \in F$ and $\alpha, \beta \in W_2 \Rightarrow a\alpha+b\beta \in W_2$.

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

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

Hence $W_1 \cap W_2$ is a subspace of $V(F)$.

Note. The union of two subspaces of $V(F)$ may not be a subspace of $V(F)$. For example if $\mathbf{R}$ be the field of real numbers, then $W_1=\{(0, 0, z) : z \in \mathbf{R}\}$ and $W_2=\{(0, y, 0) : y \in \mathbf{R}\}$ are two subspaces of $V_3(\mathbf{R})$. We have $(0, 0, 3) \in W_1$ and $(0, 5, 0) \in W_2$.

$\therefore$ $(0, 0, 3)$ and $(0, 5, 0)$ are both elements of $W_1 \cup W_2$

But $(0, 0, 3)+(0, 5, 0)=(0, 5, 3) \notin W_1 \cup W_2$ since neither $(0, 5, 3) \in W_1$ nor $(0, 5, 3) \in W_2$. Thus $W_1 \cup W_2$ is not closed under vector addition. Hence $W_1 \cup W_2$ is not a subspace of $V_3(\mathbf{R})$.

Theorem 2. *The union of two subspaces is a subspace if and only if one is contained in the other.*

(Meerut 1988; Poona 72; Allahabad 77)

Proof. Suppose W_1 and W_2 are two subspaces of a vector space V.

Let $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$. Then $W_1 \cup W_2 = W_2$ or W_1. But W_1, W_2 are subspaces and therefore, $W_1 \cup W_2$ is also a subspace.

Conversely, suppose $W_1 \cup W_2$ is a subspace.

To prove that $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Let us assume that W_1 is not a subset of W_2 and W_2 is also not a subset of W_1.

Now W_1 is not a subset of $W_2 \Rightarrow \exists\ \alpha \in W_1$ and $\alpha \notin W_2$ ...(1)

and W_2 is not a subset of $W_1 \Rightarrow \exists\ \beta \in W_2$ and $\beta \notin W_1$ ...(2)

From (1) and (2), we have

$$\alpha \in W_1 \cup W_2 \text{ and } \beta \in W_1 \cup W_2.$$

Since $W_1 \cup W_2$ is a subspace, therefore

$$\alpha+\beta \text{ is also in } W_1 \cup W_2.$$

But $\alpha+\beta \in W_1 \cup W_2 \Rightarrow \alpha+\beta \in W_1$ or W_2.

Suppose $\alpha+\beta \in W_1$. Since $\alpha \in W_1$ and W_1 is a subspace, therefore $(\alpha+\beta)-\alpha=\beta$ is in W_1.

But from (2), we have $\beta \notin W_1$. Thus we get a contradiction. Again suppose that $\alpha+\beta \in W_2$. Since $\beta \in W_2$ and W_2 is a subspace, therefore $(\alpha+\beta)-\beta=\alpha$ is in W_2. But from (1), we have $\alpha \notin W_2$. Thus here also we get a contradiction. Hence either $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Scanned with CamScanner

Theorem 3. *Arbitrary intersection af subspaces i.e., the intersection of any family af subspaces af a vector space is a subspace.* **(Meerut 1971, 73)**

Proof. Let $V(F)$ be a vector space and let $\{W_t : t \in T\}$ be any family of subspaces of V. Here T is an index set and is such that $\forall\ t \in T$, W_t is a subspace of V.

Let $\quad U = \bigcap_{t \in T} W_t = \{x \in V : x \in W_t \ \forall\ t \in T\}$

be the intersection of this family of subspaces of V. Then to prove that U is also a subspace of V.

Obviously $U \neq \varnothing$, since at least the zero vector **0** of V is in $W_t \ \forall\ t \in T$.

Now let $a, b \in F$ and α, β be any two elements of $\bigcap_{t \in T} W_t$.

Then $\alpha, \beta \in W_t \ \forall\ t \in T$. Since each W_t is a subspace of V, therefore $a\alpha + b\beta \in W_t \ \forall\ t \in T$. Thus $a\alpha + b\beta \in \bigcap_{t \in T} W_t$.

Thus $a, b \in F$ and $\alpha, \beta \in \bigcap_{t \in T} W_t \Rightarrow a\alpha + b\beta \in \bigcap_{t \in T} W_t$.

Hence $\bigcap_{t \in T} W_t$ is a subspace of $V(F)$.

Smallest subspace containing any subset of $V(F)$. Let $V(F)$ be a vector space and S be any subset of V. If U is a subspace of V containing S and is itself contained in every subspace of V containing S, then U is called the **smallest subspace** of V containing S. The smallest subspace of V containing S is also called the subspace of V generated or spanned by S and we shall denote it by the symbol $\{S\}$. It can be easily seen that the intersection of all the subspaces of $V(F)$ containing S is the subspace of $V(F)$ generated by S If $\{S\} = V$, then we say that V is spanned by S.

§ 11. Linear combination of vectors. Linear span of a set.

Linear combination. Definition. *Let $V(F)$ be a vector space. If $\alpha_1, \alpha_2, \ldots, \alpha_n \in V$, then any vector*

$$\boxed{\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \text{ where } a_1, a_2, \ldots, a_n \in F}$$

is called a linear combination of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$.

Linear span. Definition. **(Nagarjuna 1980)** *Let $V(F)$ be a vector space and S be any non-empty subset of V. Then the linear span of S is the set* *of all linear combinations of finite sets of elements of S and is denoted by L(S).* Thus we have

$\boxed{L(S)=\{a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n : \alpha_1, \alpha_2, \ldots, \alpha_n}$ is any arbitrary finite subset of S and $a_1, a_2, \ldots, a_n$ is any arbitrary finite subset of $F\}$.

Theorem 1. *The linear span $L(S)$ of any subset S of a vector space $V(F)$ is a subspace of V generated by S i.e., $L(S)=\{S\}$.*

(Andhra 1992; Kakatiya 91)

Proof. Let α, β be any two elements of $L(S)$.

Then $\quad \alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m$

and $\quad \beta=b_1\beta_1+b_2\beta_2+\ldots+b_n\beta_n$

where the a's and b's are elements of F and the α's and β's are elements of S.

If a, b be any two elements of F, then

$$a\alpha+b\beta=a\,(a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m)+b\,(b_1\beta_1+b_2\beta_2+\ldots+b_n\beta_n)$$

$$=a\,(a_1\alpha_1)+a\,(a_2\alpha_2)+\ldots+a\,(a_m\alpha_m)+b\,(b_1\beta_1)+b\,(b_2\beta_2)+\ldots+b(b_n\beta_n)$$

$$=(aa_1)\,\alpha_1+(aa_2)\,\alpha_2+\ldots+(aa_m)\,\alpha_m+(bb_1)\,\beta_1+(bb_2)\,\beta_2+\ldots+(bb_n)\,\beta_n.$$

Thus $a\alpha+b\beta$ has been expressed as a linear combination of a finite set $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n$ of the elements of S. Consequently $a\alpha+b\beta\in L(S)$.

Thus $a, b\in F$ and $\alpha, \beta\in L(S) \Rightarrow a\alpha+b\beta\in L(S)$.

Hence $L(S)$ is a subspace of $V(F)$.

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

Now if W is any subspace of V containing S, then each element of $L(S)$ must be in W because W is to be closed under vector addition and scalar multiplication. Therefore $L(S)$ will be contained in W.

Hence $L(S)=\{S\}$ *i.e.*, $L(S)$ is the smallest subspace of V containing S.

Note 1. Important. Suppose S is a non-empty subset of a vector space $V(F)$. Then a vector $\alpha\in V$ will be in the subspace of V generated by S if it can be expressed as a linear combination over F of a finite number of vectors belonging to S.

Note 2. If in any case we are to prove that $L(S)=V$, then we should prove that $V\subseteq L(S)$ because $L(S)\subseteq V$ since $L(S)$ is a

Scanned with CamScanner

subspace of V. In order to prove that $V \subseteq L(S)$, we should prove that each element of V can be expressed as a linear combination of a finite number of elements of S. Then each element of V will also be an element of $L(S)$ and we shall have $V \subseteq L(S)$.

Finally $V \subseteq L(S)$ and $L(S) \subseteq V \Rightarrow L(S)=V$.

Illustrative Examples

Example 1. The subset containing a single element (1, 0, 0) of the vector space $V_3(F)$ generates the subspace which is the totality of the elements of the form $(a, 0, 0)$.

Example 2. The subset $\{(1, 0, 0), (0, 1, 0)\}$ of $V_3(F)$ generates the subspace which is the totality of the elements of the form $(a, b, 0)$.

Example 3. The subset $S=\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ of $V_3(F)$ generates or spans the entire vector space $V_3(F)$ *i.e.*, $L(S)=V$.

If (a, b, c) be any element of V, then

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

Thus $(a, b, c) \in L(S)$. Hence $V \subseteq L(S)$. Also $L(S) \subseteq V$. Hence $L(S)=V$.

Example 4. Let V be the vector space of all polynomials over the field F. Let S be the subset of V consisting of the polynomials $f_0, f_1, f_2, \ldots$, defined by $f_n=x^n$, $n=0, 1, 2, \ldots\ldots$

Then $\quad V=L(S)$.

§ 12. Linear sum of two subspaces. Definition. *Let W_1 and W_2 be two subspaces of the vector space $V(F)$. Then the linear sum of the subspaces W_1 and W_2, denoted by W_1+W_2, is the set of all sums $\alpha_1+\alpha_2$ such that $\alpha_1 \in W_1$, $\alpha_2 \in W_2$.*

Thus $\quad W_1+W_2=\{\alpha_1+\alpha_2 : \alpha_1 \in W_1, \alpha_2 \in W_2\}$.

Theorem. *If W_1 and W_2 are subspaces of the vector space $V(F)$, then*

(*i*) *W_1+W_2 is a subspace of $V(F)$.*

(Marathwada 1971; Madras 81; Nagarjuna 90)

(*ii*) *$W_1+W_2=\{W_1 \cup W_2\}$ i.e., $L(W_1 \cup W_2)=W_1+W_2$.*

Proof. (i) Let α, β be any two elements of W_1+W_2. Then $\alpha=\alpha_1+\alpha_2$ and $\beta=\beta_1+\beta_2$ where $\alpha_1, \beta_1 \in W_1$ and $\alpha_2, \beta_2 \in W_2$. If $a, b \in F$, we have

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

Scanned with CamScanner

Since W_1 is a subspace of V, therefore $a, b \in F$

and $\alpha_1, \beta_1 \in W_1 \Rightarrow a\alpha_1 + b\beta_1 \in W_1$.

Similarly $\quad a\alpha_2 + b\beta_2 \in W_2$.

Consequently $a\alpha + b\beta = (a\alpha_1 + b\beta_1) + (a\alpha_2 + b\beta_2) \in W_1 + W_2$.

Thus $a, b \in F$ and $\alpha, \beta \in W_1 + W_2 \Rightarrow a\alpha + b\beta \in W_1 + W_2$.

Hence $W_1 + W_2$ is a subspace of $V(F)$.

(ii) Since W_2 contains the zero vector, therefore if $\alpha_1 \in W_1$, then we can write

$$\alpha_1 = \alpha_1 + 0 \in W_1 + W_2.$$

Thus $\quad W_1 \subseteq W_1 + W_2$.

Similarly $\quad W_2 \subseteq W_1 + W_2$.

Hence $\quad W_1 \cup W_2 \subseteq W_1 + W_2$.

Therefore $W_1 + W_2$ is a subspace of $V(F)$ containing $W_1 \cup W_2$.

Now to prove that $W_1 + W_2 = \{W_1 \cup W_2\}$ we should prove that $W_1 + W_2 \subseteq L(W_1 \cup W_2)$ and $L(W_1 \cup W_2) \subseteq W_1 + W_2$.

Let $\alpha = \alpha_1 + \beta_1$ be any element of $W_1 + W_2$. Then $\alpha_1 \in W_1$ and $\beta_1 \in W_2$. Therefore $\alpha_1, \beta_1 \in W_1 \cup W_2$. We can write

$$\alpha_1 + \beta_1 = 1\alpha_1 + 1\beta_1.$$

Thus $\alpha_1 + \beta_1$ is a linear combination of a finite number of elements $\alpha_1, \beta_1 \in W_1 \cup W_2$.

Therefore $\quad \alpha_1 + \beta_1 \in L(W_1 \cup W_2)$.

$\therefore \quad W_1 + W_2 \subseteq L(W_1 \cup W_2)$.

Also $L(W_1 \cup W_2)$ is the smallest subspace containing $W_1 \cup W_2$ and $W_1 + W_2$ is a subspace containing $W_1 \cup W_2$. Therefore $L(W_1 \cup W_2)$ must be contained in $W_1 + W_2$. Consequently

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

Hence $W_1 + W_2 = L(W_1 \cup W_2) = \{W_1 \cup W_2\}$.

Note. If $W_1, W_2, \ldots, W_k$ are subspaces of the vector space V, then their linear sum, denoted by $W_1 + W_2 + \ldots + W_k$, is the set of all sums $\alpha_1 + \alpha_2 + \ldots + \alpha_k$ such that $\alpha_i \in W_i$. It can be proved that $W_1 + W_2 + \ldots + W_k$ is a subspace of V which contains each of the subspaces W_i and which is spanned by the union of $W_1, W_2, \ldots, W_k$.

Solved Examples

Ex. 1. *If S, T are subsets of $V(F)$, then*

(*i*) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$.

(*ii*) $L(S \cup T) = L(S) + L(T)$.

(*iii*) *S is a subspace of V* $\Leftrightarrow L(S) = S$.

(*iv*) $L(L(S)) = L(S)$.

Scanned with CamScanner

Solution. (i) Let $\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \in L(S)$ where $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a finite subset of S. Since $S \subseteq T$, therefore $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is also a finite subset of T. So $\alpha \in L(T)$.

Thus $\qquad \alpha \in L(S) \Rightarrow \alpha \in L(T)$.

$$\therefore \quad L(S) \subseteq L(T).$$

(ii) Let α be any element of $L(S \cup T)$. Then

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m + b_1\beta_1 + b_2\beta_2 + \ldots + b_p\beta_p$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_p\}$ is a finite subset of $S \cup T$ such that $\{\alpha_1, \alpha_2, \ldots, \alpha_m\} \subseteq S$ and $\{\beta_1, \beta_2, \ldots, \beta_p\} \subseteq T$.

Now $\qquad a_1\alpha_1 + a_2\alpha_2 + \ldots + a_m\alpha_m \in L(S)$

and $\qquad b_1\beta_1 + b_2\beta_2 + \ldots + b_p\beta_p \in L(T)$.

Therefore $\qquad \alpha \in L(S) + L(T)$.

Consequently $L(S \cup T) \subseteq L(S) + L(T)$.

Now let γ be any element of $L(S) + L(T)$. Then $\gamma = \beta + \delta$ where $\beta \in L(S)$ and $\delta \in L(T)$. Now β will be a linear combination of a finite number of elements of S and δ will be a linear combination of a finite number of elements of T. Therefore $\beta + \delta$ will be a linear combination of a finite number of elements of $S \cup T$. Thus $\beta + \delta \in L(S \cup T)$. Consequently

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

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

(iii) Suppose S is a subspace of V. Then we are to prove that $L(S) = S$.

Let $\alpha \in L(S)$. Then $\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n$ where $a_1, \ldots, a_n \in F$ and $\alpha_1, \ldots, \alpha_n \in S$. But S is a subspace of V. Therefore it is closed with respect to scalar multiplication and vector addition. Hence $\alpha = a_1\alpha_1 + \ldots + a_n\alpha_n \in S$. Thus

$$\alpha \in L(S) \Rightarrow \alpha \in S.$$

Therefore $L(S) \subseteq S$. Also $S \subseteq L(S)$. Therefore $L(S) = S$.

Converse. Suppose $L(S) = S$. Then to prove that S is a subspace of V. We know that $L(S)$ is a subspace of V. Since $S = L(S)$, therefore S is also a subspace of V.

(iv) $L(L(S))$ is the smallest subspace of V containing $L(S)$. But $L(S)$ is a subspace of V. Therefore the smallest subspace of V containing $L(S)$ is $L(S)$ itself.

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

Ex. 2. *Show that the intersection of any collection of subspaces of a vector space is a subspace. Can you replace 'intersection' by 'Union' in this proposition ?* (Meerut 1973)

Ans. No.

Scanned with CamScanner

Ex. 3. *Let V be the vector space of all functions from $\mathbf{R}$ into $\mathbf{R}$; let V_e be the subset of even functions, $f(-x)=f(x)$; let V_o be the subset of odd functions, $f(-x)=-f(x)$.*

(i) *Prove that V_e and V_o are subspaces of V.*

(ii) *Prove that $V_e+V_o=V$.*

(iii) *Prove that $V_e \cap V_o=\{0\}$.*

(I.A.S. 1985)

Solution. (i) Suppose f_e and $g_e \in V_e$ and a is any scalar *i.e.* $a \in \mathbf{R}$. Then

$$(af_e+g_e)(-x)=a f_e(-x)+g_e(-x)$$
$$=af_e(x)+g_e(x)=(af_e+g_e)(x).$$

Therefore af_e+g_e is an even function.

Thus $a \in \mathbf{R}$ and $f_e, g_e \in V_e \Rightarrow af_e+g_e \in V_e$.

Hence V_e is a subspace of V.

Again suppose f_o and $g_o \in V_o$ and a is any scalar. Then

$$(af_o+g_o)(-x)=af_o(-x)+g_o(-x)=a[-f_o(x)]-g_o(x)$$
$$=-[af_o(x)+g_o(x)]=-(af_o+g_o)(x).$$

Therefore af_o+g_o is an odd function.

Thus $a \in \mathbf{R}$ and $f_o, g_o \in V_o \Rightarrow af_o+g_o \in V_o$.

Hence V_o is a subspace of V.

(ii) Since V_e and V_o are subspaces of V, therefore V_e+V_o is also a subspace of V and consequently $V_e+V_o \subseteq V$.

Now let $f \in V$. We shall show that f can be expressed as the sum of an even and an odd function

Let $fe_1(x)=\frac{1}{2}[f(x)+f(-x)]$ and

$fo_1(x)=\frac{1}{2}[f(x)-f(-x)]$.

Then obviously fe_1 is an even function and fo_1 is an odd function. We can easily see that $fe_1(-x)=fe_1(x)$

and $\quad fo_1(-x)=-fo_1(x)$.

Now $f(x)=\frac{1}{2}[f(x)+f(-x)]+\frac{1}{2}[f(x)-f(-x)]$

$$=fe_1(x)+fo_1(x)=(fe_1+fo_1)(x).$$

Therefore $f=fe_1+fo_1$ where $fe_1 \in V_e$, $o_1 \in V_o$.

Thus $f \in V \Rightarrow f \in V_e+V_o$. Therefore $V \subseteq V_e+V_o$.

Hence $\quad V=V_e+V_o$.

(iii) Let 0 denote the zero function *i.e.* $0(x)=0 \ \forall \ x \in \mathbf{R}$.

Then $0 \in V_e$ and also $0 \in V_o$.

Let $f \in V_e$ and also $f \in V_o$.

Then $f(-x)=f(x)=-f(x)$.

$\therefore \quad 2f(x)=0$

$$\Rightarrow f(x)=0=0(x).$$

Therefore $f=0$ (zero function).

Thus $f \in V_e \cap V_o \Leftrightarrow f=0$. Hence $V_e \cap V_o = \{0\}$.

§ 13. Linear dependence and linear independence of vectors.
(Kakatiya 1991; Osmania 90; Nagarjuna 78)

Linear dependence. Definition. *Let $V(F)$ be a vector space. A finite set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of vectors of V is said to be linearly dependent if there exist scalars $a_1, a_2, \ldots, a_n \in F$ not all of them 0 (some of them may be zero) such that*

$$a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 + \ldots + a_n\alpha_n = 0.$$

Linear independence. Definition. (Meerut 1980; Nagarjuna 78)

Let $V(F)$ be a vector space. A finite set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of vectors of V is said to be linearly independent if every relation of the form

$$a_1\alpha_1 + a_2\alpha_2 + a_3\alpha_3 + \ldots + a_n\alpha_n = 0,\ a_i \in F,\ 1 \leqslant i \leqslant n$$
$$\Rightarrow a_i = 0 \text{ for each } 1 \leqslant i \leqslant n.$$

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

Illustrative Examples

Example 1. *Prove that if two vectors are linearly dependent, one of them is a scalar multiple of the other.*

Solution. Let α, β be two linearly dependent vectors of the vector space V. Then $\exists$ scalars a, b not both zero, such that

$$a\alpha + b\beta = 0.$$

If $a \neq 0$, then we get

$$a\alpha = -b\beta$$

$$\Rightarrow \alpha = \left(-\frac{b}{a}\right)\beta \Rightarrow \alpha \text{ is a scalar multiple of } \beta.$$

If $b \neq 0$, then we get

$$b\beta = -a\alpha$$

$$\Rightarrow \beta = \left(-\frac{a}{b}\right)\alpha \Rightarrow \beta \text{ is a scalar multiple of } \alpha.$$

Thus one of the vectors α and β is a scalar multiple of the other.

Example 2. *In the vector space $V_n(F)$, the system of n vectors*

$$e_1 = (1, 0, 0, \ldots, 0),\ e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 0, 1)$$

is linearly independent where 1 denotes the unity of the field F.

Solution. If $a_1, a_2, a_3, \ldots, a_n$ be any scalars, then

$$a_1e_1 + a_2e_2 + \ldots + a_ne_n = 0$$

$\Rightarrow a_1 (1, 0, 0, \ldots, 0) + a_2 (0, 1, 0, \ldots, 0) + \ldots$
$+ a_n (0, 0, \ldots, 0, 1) = \mathbf{0}$

$\Rightarrow (a_1, a_2, \ldots, a_n) = (0, 0, \ldots, 0)$

$\Rightarrow a_1 = 0, a_2 = 0, \ldots, a_n = 0.$

Therefore the given set of n vectors is linearly independent.

In particular $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ *is a linearly indepen-ent subset of* $V_3 (F)$.

(Nagarjuna 1980)

Example 3. *If the set* $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *f vectors of* $V (F)$ *is linearly independent, then none of the vectors* $_1, \alpha_2, \ldots, \alpha_n$ *can be zero vector.*

Solution. Let α_r be equal to zero vector where $1 \leqslant r \leqslant n$.

Then $0\alpha_1 + 0\alpha_2 + \ldots + a\alpha_r + 0\alpha_{r+1} + \ldots + 0\alpha_n = \mathbf{0}$

r any $a \neq 0$ in F.

Since $a \neq 0$, therefore from this relation we conclude that S is inearly dependent. Thus we get a contradiction because it is iven that S is linearly independent. Hence none of the vectors $_1, \alpha_2, \ldots, \alpha_n$ can be zero vector. We also conclude that *a set of ectors which contains the zero vector is necessarily linearly depen-lent.*

(Kanpur 1981; Meerut 88)

Example 4. *Every superset of a linearly dependent set of vectors s linearly dependent.*

Solution. Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linearly dependent set of ectors. Then there exist scalars $a_1, a_2, \ldots, a_n$ not all zero such that

$$a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n = \mathbf{0}. \qquad \ldots(1)$$

Now let $S' = \{\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \ldots, \beta_m\}$ be a superset of S. Then we have from (1)

$$a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n + 0\beta_1 + 0\beta_2 + \ldots + 0\beta_m = \mathbf{0}. \qquad \ldots(2)$$

Since in the relation (2) the scalar coefficients are not all 0, therefore S' is linearly dependent.

From this we also conclude that *any subset of a linearly independent set of vectors is also linearly independent.*

(Nagarjuna 1990)

Example 5. *A system consisting of a single non-zero vector is always linearly independent.*

Solution. Let $S = \{\alpha\}$ be a subset of a vector space V and let α be not equal to zero vector. If a is any scalar, then

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

[Since α is not zero vector]

$\therefore$ the set S is linearly independent.

Example 6. *Show that*

$S=\{(1, 2, 4), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

is a linearly dependent subset of the vector space V_3 (**R**) *where* **R** *the field of real numbers.* **(Poona 1972)**

Solution. We have

$1(1, 2, 4)+(-1)(1, 0, 0)+(-2)(0, 1, 0)+(-4)(0, 0, 1)$

$=(1, 2, 4)+(-1, 0, 0)+(0, -2, 0)+(0, 0, -4)$

$=(0, 0, 0)$ *i.e.*, zero vector.

Since in this relation the scalar coefficients $1, -1, -2, -4$ are not all zero, therefore the given system S is linearly dependent.

Example 7. *In* V_3 (**R**), *where* **R** *is the field of real numbers, examine each of the following sets of vectors for linear dependence:*

(i) $\{(2, 1, 2), (8, 4, 8)\}$

(ii) $\{(1, 2, 0), (0, 3, 1), (-1, 0, 1)\}$ **(Meerut 1989)**

(iii) $\{(-1, 2, 1), (3, 0, -1), (-5, 4, 3)\}$

(iv) $\{(2, 3, 5), (4, 9, 25)\}$

(v) $\{(1, 3, 2), (1, -7, -8), (2, 1, -1)\}$

(vi) $(1, 2, 1), (3, 1, 5), (3, -4, 7)$. **(Meerut 1986)**

Solution. (i) We have

$$4(2, 1, 2)+(-1)(8, 4, 8)$$

$$=(8, 4, 8)+(-8, -4, -8)=(0, 0, 0) \text{ i.e., the zero vector.}$$

Since in this relation the scalar coefficients $4, -1$ are not both zero, therefore the given set is linearly dependent.

(ii) Let a, b, c be scalars *i.e.* real numbers such that

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

i.e., $$(a-c, 2a+3b, b+c)=(0, 0, 0)$$

i.e., $$a+0b-c=0,$$
$$2a+3b+0c=0,$$
$$0a+b+c=0.$$

These equations will have a non-zero solution *i.e.*, a solution in which a, b, c are not all zero if the rank of the coefficient matrix is less than three *i.e.*, the number of unknowns a, b, c. If the rank is 3, then zero solution $a=0, b=0, c=0$, will be the only solution.

Coefficient matrix $A=\begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

We have $\quad |A|=1(3-0)-2(0+1)=1\neq 0.$

$\therefore$ Rank $A=3$. Hence $a=0$, $b=0$, $c=0$ is the only solution. Therefore the given system is linearly independent.

(iii) Let a, b, c be scalars such that

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

i.e., $$(-a+3b-5c, 2a+0b+4c, a-b+3c)=(0, 0, 0)$$

i.e., $$-a+3b-5c=0, 2a+0b+4c=0, a-b+3c=0.$$

The coefficient matrix A of these equations is

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

We have $|A|=-1(0+4)-2(9-5)+1(12-0)=0.$

$\therefore$ Rank A is < 3 i.e., the number of unknowns a, b, c.

Therefore the given system of equations will possess a non-zero solution. For example $a=-2$, $b=1$, $c=1$ is a non-zero solution. Hence the given system of vectors is linearly dependent.

(iv) Let a, b be scalars i.e., real numbers such that

$$a(2, 3, 5)+b(4, 9, 25)=(0, 0, 0)$$

i.e., $$(2a+4b, 3a+9b, 5a+25b)=(0, 0, 0)$$

i.e., $$2a+4b=0, 3a+9b=0, 5a+25b=0.$$

The coefficient matrix A of these equations is

$$A=\begin{bmatrix} 2 & 4 \\ 3 & 9 \\ 5 & 25 \end{bmatrix}.$$

Obviously rank $A=2$ i.e., equal to the number of unknowns a and b. Therefore these equations have the only solution $a=0$, $b=0$. Hence the given set of vectors is linearly independent.

(v) Let a, b, c be scalars i.e., real numbers such that

$$a(1, 3, 2)+b(1, -7, -8)+c(2, 1, -1)=(0, 0, 0)$$

i.e., $$(a+b+2c, 3a-7b+c, 2a-8b-c)=(0, 0, 0)$$

i.e., $$a+b+2c=0, \quad \ldots(1)$$

$$3a-7b+c=0, \quad \ldots(2)$$

$$2a-8b-c=0. \quad \ldots(3)$$

Eliminating c between (1) and (2), we get

$$5a-15b=0 \text{ or } a-3b=0.$$

Eliminating c between (2) and (3), we get

$$5a-15b=0 \text{ or } a-3b=0,$$

which is the same equation as obtained on eliminating c between (1) and (2).

Scanned with CamScanner

If we choose $b=1$, then $a=3$ and putting in any one of the equations (1), (2), (3), we get $c=-2$. Hence the given set is linearly dependent.

(vi) Let a, b, c be scalars *i.e.*, real numbers such that

$$a(1, 2, 1)+b(3, 1, 5)+c(3, -4, 7)=(0, 0, 0)$$

i.e., $$(a+3b+3c, 2a+b-4c, a+5b+7c)=(0, 0, 0)$$

i.e.,

$$a+3b+3c=0, \quad \ldots(1)$$
$$2a+b-4c=0, \quad \ldots(2)$$
$$a+5b+7c=0. \quad \ldots(3)$$

Multiplying (1) by 2, we get

$$2a+6b+6c=0. \quad \ldots(4)$$

Subtracting (4) from (2), we get

$$-5b-10c=0$$

or $$b+2c=0. \quad \ldots(5)$$

Again subtracting (3) from (1), we get

$$-2b-4c=0 \text{ or } b+2c=0. \quad \ldots(6)$$

The equations (5) and (6) are the same and give $b=-2c$.

Putting $b=-2c$ in (1), we get $a=3c$. If we take $c=1$, we get $b=-2$ and $a=3$. Thus $a=3$, $b=-2$, $c=1$ is a non-zero solution of the equations (1), (2) and (3). Hence the given set of vectors is linearly dependent.

Example 8. *If F is the field of real numbers, prove that the vectors (a_1, a_2) and (b_1, b_2) in $V_2(F)$ are linearly dependent iff*

$$a_1b_2-a_2b_1=0.$$

(Kanpur 1981; Gorakhpur 79)

Solution. Let $x, y \in F$. Then

$$x(a_1, a_2)+y(b_1, b_2)=(0, 0)$$
$$\Rightarrow (xa_1+yb_1, xa_2+yb_2)=(0, 0).$$

Therefore $$\left.\begin{array}{l} a_1x+b_1y=0 \\ a_2x+b_2y=0 \end{array}\right\}.$$

and

The necessary and sufficient condition for these equations to possess a non-zero solution is that

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}=0$$

i.e., $$a_1b_2-a_2b_1=0.$$

Hence the given system is linearly dependent iff

$$a_1b_2-a_2b_1=0.$$

Example 9. *If α_1 and α_2 are vectors of $V(F)$, and $a, b \in F$, show that the set $\{\alpha_1, \alpha_2, a\alpha_1+b\alpha_2\}$ is linearly dependent.*

(Poona 1972)

Solution. We have

$$(-a)\,\alpha_1+(-b)\,\alpha_2+1\,(a\alpha_1+b\alpha_2)$$
$$=(-a+a)\,\alpha_1+(-b+b)\,\alpha_2$$
$$=0\alpha_1+0\alpha_2=\mathbf{0} \text{ i.e., zero vector.}$$

Whatever may be the scalars $-a$ and $-b$ since $1\neq 0$, therefore the given set of vectors is linearly dependent.

Example 10. *Let $\alpha_1, \alpha_2, \alpha_3$ be the vectors of $V(F)$, $a, b \in F$. Show that the set $\{\alpha_1, \alpha_2, \alpha_3\}$ is linearly dependent if the set $\{\alpha_1+a\alpha_2+b\alpha_3, \alpha_2, \alpha_3\}$ is linearly dependent.*

Solution. Since the set $\{\alpha_1+a\alpha_2+b\alpha_3, \alpha_2, \alpha_3\}$ is linearly dependent, therefore there exist scalars, x, y, z not all zero such that

$$x\,(\alpha_1+a\alpha_2+b\alpha_3)+y\alpha_2+z\alpha_3=\mathbf{0}$$

i.e.
$$x\alpha_1+(xa+y)\,\alpha_2+(xb+z)\,\alpha_3=\mathbf{0}. \qquad \ldots(1)$$

If in the relation (1), the coefficients x, $xa+y$, $xb+z$ are not all zero, then the set $\{\alpha_1, \alpha_2, \alpha_3\}$ will also be linearly dependent.

If $x\neq 0$, then the problem is at once solved whatever y and z may be. However if $x=0$, then at least one of y and z is not zero. Therefore at least one of $xa+y$ and $xb+z$ will not be zero since when $x=0$ then $xa+y$ and $xb+z$ reduce to y and z respectively.

Hence in the relation (1) the scalar coefficients of $\alpha_1, \alpha_2, \alpha_3$ are not all zero. Therefore the set $\{\alpha_1, \alpha_2, \alpha_3\}$ is also linearly dependent.

Example 11. *If α, β, γ are linearly dependent vectors of $V(F)$ where F is any subfield of the field of complex numbers then so also are $\alpha+\beta, \beta+\gamma, \gamma+\alpha$.* **(Meerut 1986)**

Solution. Let a, b, c be scalars such that

$$a\,(\alpha+\beta)+b\,(\beta+\gamma)+c\,(\gamma+\alpha)=\mathbf{0}$$

i.e.,
$$(a+c)\,\alpha+(a+b)\,\beta+(b+c)\,\gamma=\mathbf{0}. \qquad \ldots(1)$$

But α, β, γ are linearly independent. Therefore (1) implies

$$a+0b+c=0$$
$$a+b+0c=0$$
$$0a+b+c=0.$$

The coefficient matrix A of these equations is

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

We have rank $A=3$ i.e., the number of unknowns a, b, c.

Scanned with CamScanner

Therefore $a=0, b=0, c=0$ is the only solution of the given equations.

Hence $\alpha+\beta$, $\beta+\gamma$, $\gamma+\alpha$ are also linearly independent.

Example 12. *If α, β, γ are linearly independent vectors of $V(F)$ where F is the field of complex numbers, then so also are*

$$\alpha+\beta,\ \alpha-\beta,\ \alpha-2\beta+\gamma.$$

Solution. Let a, b, c be scalars such that

$$a(\alpha+\beta)+b(\alpha-\beta)+c(\alpha-2\beta+\gamma)=0 \quad \ldots(1)$$

i.e., $$(a+b+c)\,\alpha+(a-b-2c)\,\beta+c\gamma=0. \quad \ldots(2)$$

But α, β, γ are linearly independent. Therefore (2) implies

$$a+b+c=0,\ a-b-2c=0,\ c=0.$$

The only solution of these equations is $c=0$, $a=0$, $b=0$.

Thus (1) implies $a=0$, $b=0$, $c=0$. Therefore the vectors $\alpha+\beta$, $\alpha-\beta$, $\alpha-2\beta+\gamma$ are linearly independent.

Example 13. *Show that the set $\{1, x, 1+x+x^2\}$ is a linearly independent set of vectors in the vector space of all polynomials over the real number field.* **(Meerut 1976)**

Solution. Let a, b, c be scalars (real numbers) such that $a(1)+bx+c(1+x+x^2)=0$. We have

$$\begin{aligned} & a(1)+bx+c(1+x+x^2)=0 \\ \Rightarrow\ & (a+c)+(b+c)\,x+cx^2=0 \\ \Rightarrow\ & a+c=0,\ b+c=0,\ c=0 \\ \Rightarrow\ & c=0,\ b=0,\ a=0. \end{aligned}$$

$\therefore$ the vectors $1, x, 1+x+x^2$ are linearly independent over the field of real numbers.

Example 14. *In the vector space $F[x]$ of all polynomials over the field F the infinite set $S=\{1, x, x^2, x^3, \ldots\}$ is linearly independent.*

Solution. Let $S'=\{x^{m_1}, x^{m_2}, \ldots, x^{m_n}\}$ be any finite subset of S having n vectors. Here $m_1, m_2, \ldots, m_n$ are some non-negative integers. Let $a_1, a_2, \ldots, a_n$ be scalars such that

$$a_1x^{m_1}+a_2x^{m_2}+\ldots+a_nx^{m_n}=0. \quad \ldots(1)$$

(*i.e.*, zero polynomial)

By the definition of equality of two polynomials we have from (1)

$$a_1=0,\ a_2=0, \ldots, a_n=0.$$

Thus every finite subset of S is linearly independent.

Therefore S is linearly independent.

Scanned with CamScanner

Example 15. *Is the vector (2, −5, 3) in the subspace of* $\mathbf{R}^3$ *spanned by the vectors (1, −3, 2), (2, −4, −1), (1, −5, 7)?*

(Meerut 1990)

Solution. Let $\alpha=(2, -5, 3)$, $\alpha_1=(1, -3, 2)$, $\alpha_2=(2, -4, -1)$, $\alpha_3=(1, -5, 7)$. If α can be expressed as a linear combination of the vectors $\alpha_1, \alpha_2, \alpha_3$ then it will be in the subspace of $\mathbf{R}^3$ spanned by these vectors otherwise it will not be.

Let $\alpha=a_1\alpha_1+a_2\alpha_2+a_3\alpha_3$ where $a_1, a_2, a_3 \in \mathbf{R}$.

Then $(2, -5, 3)=a_1(1, -3, 2)+a_2(2, -4, -1)+a_3(1, -5, 7)$

$\therefore$ $(2, -5, 3)=(a_1+2a_2+a_3, -3a_1-4a_2-5a_3, 2a_1-a_2+7a_3)$

$$\therefore \quad \left.\begin{aligned} a_1+2a_2+a_3&=2 \\ -3a_1-4a_2-5a_3&=-5 \\ 2a_1-a_2+7a_3&=3 \end{aligned}\right\} \qquad \begin{aligned} &\ldots(1) \\ &\ldots(2) \\ &\ldots(3) \end{aligned}$$

Multiplying the equation (1) by 3 and adding to (2), we get

$$2a_2-2a_3=1 \text{ or } a_2-a_3=\tfrac{1}{2}. \qquad \ldots(4)$$

Again multiplying the equation (1) by 2 and subtracting from (3), we get

$$-5a_2+5a_3=-1 \text{ or } a_2-a_3=1/5. \qquad \ldots(5)$$

The relations (4) and (5) show that the above equations are inconsistent. Hence the vector α cannot be expressed as a linear combination of the vectors $\alpha_1, \alpha_2, \alpha_3$. Therefore α is not in the subspace of $\mathbf{R}^3$ generated by the vectors $\alpha_1, \alpha_2, \alpha_3$.

Example 16. *In the vector space* $\mathbf{R}^3$, *let* $\alpha=(1, 2, 1)$, $\beta=(3, 1, 5)$, $\gamma=(3, -4, 7)$. *Show that the subspaces spanned by* $S=\{\alpha, \beta\}$ *and* $T=\{\alpha, \beta, \gamma\}$ *are the same.*

(Meerut 1977)

Solution. First we shall show that the vector γ can be expressed as a linear combination of the vectors α and β. Let

$$(3, -4, 7)=a(1, 2, 1)+b(3, 1, 5).$$

Then $a+3b=3$, $2a+b=-4$, $a+5b=7$. Solving the first two equations we get $a=-3$, $b=2$ and these satisfy the third equation also. Therefore we can write $\gamma=-3\alpha+2\beta$.

Now
$$S \subseteq T \Rightarrow L(S) \subseteq L(T).$$

Further let $\delta \in L(T)$. Then δ can be expressed as a linear combination of the vectors α, β and γ. In this linear combination the vector γ can be replaced by $-3\alpha+2\beta$. Thus δ can be expressed as a linear combination of the vectors α and β. Therefore $\delta \in L(S)$. Thus $\delta \in L(T) \Rightarrow \delta \in L(S)$. Therefore $L(T) \subseteq L(S)$.

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

Scanned with CamScanner

Example 17. *Show that the three vectors* (1, 1, −1), (2, −3, 5) *and* (−2, 1, 4) *of* $\mathbf{R}^3$ *are linearly independent.* **(Meerut 1986)**

Solution. Let a, b, c be scalars *i.e.*, real numbers such that

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

i.e., $(a+2b-2c, a-3b+c, -a+5b+4c)=(0, 0, 0)$

i.e.,

$$a+2b-2c=0 \quad ...(1)$$

$$a-3b+c=0 \quad ...(2)$$

$$-a+5b+4c=0. \quad ...(3)$$

Now we shall solve the simultaneous equations (1), (2) and (3).

Multiplying (2) by 2 and adding to (1), we get

$$3a-4b=0. \quad ...(4)$$

Again multiplying (1) by 2 and adding to (3), we get

$$a+9b=0. \quad ...(5)$$

Multiplying (5) by 3 and subtracting from (4), we get

$$-31b=0 \quad \text{or} \quad b=0.$$

Putting $b=0$ in (5), we get $a=0$.

Now putting $a=0, b=0$ in (1), we get $c=0$.

Thus $a=0, b=0, c=0$ is the only solution of the equations (1), (2) and (3).

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

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

Hence the vectors (1, 1, −1), (2, −3, 5), (−2, 1, 4) of $\mathbf{R}^3$ are linearly independent.

Example 18. *Show that the vectors* (1, 1, 2, 4), (2, −1, −5, 2), (1, −1, −4, 0) *and* (2, 1, 1, 6) *are linearly independent in* $\mathbf{R}^4$.

(Nagarjuna 1990; Meerut 89)

Solution. Let $(1, 1, 2, 4)=a(2, -1, -5, 2)+b(1, -1, -4, 0)+c(2, 1, 1, 6)$.

Then

$$2a+b+2c=1 \quad ...(1)$$

$$-a-b+c=1 \quad ...(2)$$

$$-5a-4b+c=2 \quad ...(3)$$

$$2a+0b+6c=4 \quad ...(4)$$

Now we shall solve the simultaneous equations (1), (2), (3) and (4).

Adding (1) and (2), we get $a+3c=2$ which is the same equation as (4).

If we take $c=0$, we get $a=2$.

Scanned with CamScanner

Putting $a=2$ and $c=0$ in (1), we get $b=-3$.

We see that $a=2$, $b=-3$, $c=0$ satisfy all the four equations (1), (2), (3) and (4).

$\therefore \quad (1, 1, 2, 4)=2(2, -1, -5, 2)-3(1, -1, -4, 0) +0(2, 1, 1, 6)$

or $\quad 1(1, 1, 2, 4)-2(2, -1, -5, 2)+3(1, -1, -4, 0) -0(2, 1, 1, 6)=(0, 0, 0, 0). \quad ...(1)$

Since in the linear relation (1) among the four given vectors the scalar coefficients 1, -2, 3, 0 are not all zero, therefore the given vectors are linearly dependent in $\mathbf{R}^4$.

Example 19. *Show that the vectors* (1, 1, 0, 0), (0, 1, -1, 0), (0, 0, 0, 3) *in* $\mathbf{R}^4$ *are linearly independent*

Solution. Let a, b, c be scalars *i.e.*, real numbers such that

$a(1, 1, 0, 0)+b(0, 1, -1, 0)+c(0, 0, 0, 3)=(0, 0, 0, 0) \quad ...(1)$

Then
$$a+0b+0c=0,$$
$$a+b+0c=0,$$
$$0a-b+0c=0,$$
$$0a+0b+3c=0.$$

The only solution of the above equations is
$$a=0,\ b=0,\ c=0.$$

Thus the linear relation (1) among the three given vectors is possible only if $\quad a=0; b=0, c=0.$

Hence the three given vectors in $\mathbf{R}^4$ are linearly independent.

Example 20. *Is the vector* (3, -1, 0, -1) *in the subspace of* $\mathbf{R}^4$ *spanned by the vectors* (2, -1, 3, 2), (-1, 1, 1, -3) *and* (1, 1, 9, -5) ? **(Meerut 1983)**

Solution. Let $\alpha=(3, -1, 0, -1)$, $\alpha_1=(2, -1, 3, 2)$,

$\alpha_2=(-1, 1, 1, -3)$, $\alpha_3=(1, 1, 9, -5)$.

If α can be expressed as a linear combination of the vectors $\alpha_1, \alpha_2, \alpha_3$, then it will be in the subspace of $\mathbf{R}^4$ spanned by these vectors otherwise it will not be.

Let $\alpha=a\alpha_1+b\alpha_2+c\alpha_3$ where $a, b, c \in \mathbf{R}$.

Then $\quad (3, -1, 0, -1)=a(2, -1, 3, 2)+b(-1, 1, 1, -3) +c(1, 1, 9, -5).$

$\therefore$
$$2a-b+c=3, \quad ...(1)$$
$$-a+b+c=-1, \quad ...(2)$$
$$3a+b+9c=0, \quad ...(3)$$
and
$$2a-3b-5c=-1. \quad ...(4)$$

Scanned with CamScanner

Adding the equations (1) and (2), we get

$$a+2c=2. \quad \ldots(5)$$

Again adding the equations (1) and (3) we get

$$5a+10c=3. \quad \ldots(6)$$

Multiplying the equation (5) by 5, we get

$$5a+10c=10. \quad \ldots(7)$$

The relations (6) and (7) show that the equations (1), (2), (3) and (4) are inconsistent *i.e.*, do not possess a common solution. Hence the vector α cannot be expressed as a linear combination of the vectors $\alpha_1, \alpha_2, \alpha_3$. Therefore α is not in the subspace of R^4 generated by the vectors $\alpha_1, \alpha_2, \alpha_3$.

Example 21. *Show that the set* $\{1, x, x(1-x)\}$ *is a linearly independent set of vectors in the space of all polynomials over the real number field.* **(Meerut 1985)**

Solution. The zero vector of the vector space of all polynomials over the real number field is the zero polynomial.

Let a, b, c be scalars (*i.e.*, real numbers) such that

$$a(1)+bx+c[x(1-x)]=0 \text{ i.e., zero polynomial.}$$

We have $a(1)+bx+c(x-x^2)=0$

$$\Rightarrow a+(b+c)x-cx^2=0. \quad \ldots(1)$$

Now two polynomials in x are said to be equal if the coefficients of like powers of x on both sides are equal. So by the definition of the equality of two polynomials, we have from (1),

$$a=0, b+c=0, -c=0$$

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

Thus $a(1)+bx+c[x(1-x)]=0$

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

$\therefore$ the vectors $1, x, x(1-x)$ are linearly independent over the field of real numbers.

Imp

Example 22. *Find whether the vectors* $2x^3+x^2+x+1$, x^3+3x^2+x-2 *and* x^3+2x^2-x+3 *of* $\mathbf{R}[x]$, *the vector space of all polynomials over the real number field, are linearly independent or not.* **(Meerut 1975)**

Solution. The zero vector of the vector space $\mathbf{R}[x]$ is the zero polynomial.

Let a, b, c be scalars (*i.e.*, real numbers) such that

$$a(2x^3+x^2+x+1)+b(x^3+3x^2+x-2)+c(x^3+2x^2-x+3)=0$$

i.e., zero polynomial.

Scanned with CamScanner

Then

$(2a+b+c)\, x^3+(a+3b+2c)\, x^2+(a+b-c)\, x+a-2b+3c=0.$ ...(1)

Equating the coefficients of like powers of x on both sides of (1), we get

$$\left.\begin{aligned} 2a+b+c&=0,\\ a+3b+2c&=0,\\ a+b-c&=0,\\ \text{and}\quad a-2b+3c&=0. \end{aligned}\right\}. \quad ...(2)$$

The coefficient matrix A of the system of equations (2) is

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

$$\sim\begin{bmatrix} 1 & 3 & 2\\ 2 & 1 & 1\\ 1 & 1 & -1\\ 1 & -2 & 3 \end{bmatrix}, \text{ by } R_1 \leftrightarrow R_2$$

$$\sim\begin{bmatrix} 1 & 3 & 2\\ 0 & -5 & -3\\ 0 & -2 & -3\\ 0 & -5 & 1 \end{bmatrix}, \text{ by } \begin{aligned} &R_2\to R_2-2R_1\\ &R_3\to R_3-R_1\\ &R_4\to R_4-R_1 \end{aligned}$$

$$\sim\begin{bmatrix} 1 & 3 & 2\\ 0 & 1 & 3/5\\ 0 & -2 & -3\\ 0 & -5 & 1 \end{bmatrix}, \text{ by } R_2\to -\tfrac{1}{5}R_2$$

$$\sim\begin{bmatrix} 1 & 3 & 2\\ 0 & 1 & 3/5\\ 0 & 0 & -9/5\\ 0 & 0 & 4 \end{bmatrix}, \text{ by } \begin{aligned} &R_3\to R_3+2R_2,\\ &R_4\to R_4+5R_2 \end{aligned}$$

$$\sim\begin{bmatrix} 1 & 3 & 2\\ 0 & 1 & 3/5\\ 0 & 0 & 1\\ 0 & 0 & 4 \end{bmatrix}, \text{ by } R_3\to -\tfrac{5}{9}R_3$$

$$\sim\begin{bmatrix} 1 & 3 & 2\\ 0 & 1 & 3/5\\ 0 & 0 & 1\\ 0 & 0 & 0 \end{bmatrix}, \text{ by } R_4\to R_4-4R_3$$

which is in echelon form.

$\therefore$ rank A=number of non-zero rows in its echelon form
=3=number of unknowns a, b, c in the system of equations (2).

Hence the system of equations (2) has the only solution

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

$\therefore$ the given set of vectors is linearly independent.

Ex. 23. *Prove that a set of vectors which contains the zero vector is linearly dependent.* **(Meerut 1988)**

Sol. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a set of vectors of the vector space $V(F)$.

Let α_r be equal to zero vector where $1 \leqslant r \leqslant n$.

To show that the set S of vectors is linearly dependent.

Obviously

$$0\alpha_1+0\alpha_2+\ldots+a\alpha_r+0\alpha_{r+1}+\ldots+0\alpha_n=\mathbf{0} \text{ i.e., zero vector} \quad \ldots(1)$$

for any non-zero scalar a *i.e.*, for any non-zero element a in the field F.

Since in the linear relation (1) among the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$, the scalar coefficient a is not zero, therefore the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly dependent.

Ex. 24. *Show that the system of three vectors* $(1, 3, 2)$, $(1, -7, -8)$, $(2, 1, -1)$ *of* V_3 (**R**) *is linearly dependent.*

(Meerut 1989; Gorakhpur 80)

Sol. Let a, b, c be scalars *i.e.*, real numbers such that

$$a(1, 3, 2)+b(1, -7, -8)+c(2, 1, -1)=(0, 0, 0)$$

i.e., $$(a+b+2c, 3a-7b+c, 2a-8b-c)=(0, 0, 0).$$

Then $$a+b+2c=0, \quad \ldots(1)$$

$$3a-7b+c=0, \quad \ldots(2)$$

and $$2a-8b-c=0. \quad \ldots(3)$$

If the equations (1), (2), (3) possess a non-zero solution, then the given vectors are linearly dependent.

Adding (2) and (3), we get

$$5a-15b=0 \quad \text{or} \quad a-3b=0. \quad \ldots(4)$$

Multiplying (3) by 2 and adding to (1), we get

$$5a-15b=0 \quad a-3b=0. \quad \ldots(5)$$

The equations (4) and (5) are the same and give $a=3b$.

Putting $a=3b$ in (1), we get $\quad 2c+4b=0 \quad$ or $\quad c=-2b$.

If we take $b=1$, we get $a=3, c=-2$.

Thus $a=3, b=1, c=-2$ is a non-zero solution of the equations (1), (2) and (3). Hence the given set of vectors is linearly dependent.

Alternative method. The coefficient matrix A of the system of equations (1), (2) and (3) is

$$A=\begin{bmatrix} 1 & 1 & 2 \\ 3 & -7 & 1 \\ 2 & -8 & -1 \end{bmatrix}.$$

We have $\det A=\begin{vmatrix} 1 & 1 & 2 \\ 3 & -7 & 1 \\ 2 & -8 & -1 \end{vmatrix}$

$$=\begin{vmatrix} 1 & 0 & 0 \\ 3 & -10 & -5 \\ 2 & -10 & -5 \end{vmatrix}, \text{ by } C_2-C_1 \text{ and } C_3-2C_1$$

$$=50-50=0.$$

$\therefore$ rank $A < 3$ *i.e.*, rank $A <$ the number of unknowns a, b, c in the equations (1), (2) and (3).

Therefore the equations (1), (2) and (3) must possess a non-zero solution. Hence the given vectors are linearly dependent.

Ex. 25. *Determine whether the following set of vectors in $V_3(\mathbf{Q})$ is linearly dependent or independent, $\mathbf{Q}$ being the field of rational numbers :*

$$\{(-1, 2, 1), (3, 1, -2)\}.$$ **(Meerut 1974)**

Sol. Let a, b be scalars (*i.e.*, $a, b \in \mathbf{Q}$) such that

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

i.e., $$(-a+3b, 2a+b, a-2b)=(0, 0, 0).$$

Then $$\left.\begin{aligned} -a+3b&=0, \\ 2a+b&=0, \\ a-2b&=0. \end{aligned}\right\} \quad ...(1)$$

The coefficient matrix A of the system of equations (1) is

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

We have $\begin{vmatrix} -1 & 3 \\ 2 & 1 \end{vmatrix}=-1-6=-7\neq 0.$

Thus there exists a 2-rowed minor of the matrix A which is not zero. Also the matrix A can have no minor of order greater than 2.

$\therefore$ rank $A=2=$the number of unknowns a and b.

Therefore the equations (1) have the only solution $a=0$, $b=0$. Hence the given set of vectors is linearly independent.

Note. If we do not want to use the concept of the rank of a matrix to discuss the solutions of the system of equations (1), we can directly say that solving the system of equations (1) we find that the only solution of the system of equations (1) is $a=0$, $b=0$.

Scanned with CamScanner

Ex. 26. *Find a linearly independent subset T of the set*

$$S=\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$$

where $\alpha_1=(1, 2, -1)$, $\alpha_2=(-3, -6, 3)$,

$\alpha_3=(2, 1, 3)$, $\alpha_4=(8, 7, 7) \in \mathbf{R}^3$

which spans the same space as S.

Sol. First we observe that $\alpha_2=-3\alpha_1$ so that the vectors α_1 and α_2 are linearly dependent.

$\therefore$ if $S_1=\{\alpha_1, \alpha_3, \alpha_4\}$, then the subspace of $\mathbf{R}^3$ spanned by S_1 is the same as that spanned by S.

Now there exists no real number c such that $\alpha_3=c\alpha_1$. Therefore the vectors α_1 and α_3 are linearly independent.

Let us now see whether the vector α_4 lies in the subspace of $\mathbf{R}^3$ spanned by the vectors α_1 and α_3 or not

Let $\alpha_4=a\alpha_1+b\alpha_3$, where $a, b \in \mathbf{R}$.

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

$$\therefore \quad a+2b=8,$$
$$2a+b=7,$$
and $$-a+3b=7.$$

Solving the first two of these three equations, we get $a=2$, $b=3$. These values of a and b also satisfy the third equation.

$$\therefore \quad \alpha_4=2\alpha_1+3\alpha_3.$$

Thus the vector α_4 has been expressed as a linear combination of α_1 and α_3 so that the subspace of $\mathbf{R}^3$ spanned by the vectors α_1, α_3 and α_4 is the same as that spanned by the vectors α_1 and α_3.

Hence $T=\{\alpha_1, \alpha_3\}$ is a linearly independent subset of S which spans the same subspace of $\mathbf{R}^3$ as is spanned by S.

Ex. 27. *Determine a basis of the subspace spanned by the vectors*

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

$\alpha_3=(1, -1, -4)$, $\alpha_4=(4, 2, -2)$. **(I.A.S. 1988)**

Sol. Proceed as in Ex. 26. **Ans.** $\{\alpha_1, \alpha_2\}$.

Ex. 28. *Find a maximal linearly independent subsystem of the system of vectors*

$\alpha_1=(2, -2, -4)$, $\alpha_2=(1, 9, 3)$, $\alpha_3=(-2, -4, 1)$

and $\alpha_4=(3, 7, -1)$. **(I.A.S. 1986)**

Sol. Let A denote the matrix $\begin{bmatrix} 2 & -2 & -4 \\ 1 & 9 & 3 \\ -2 & -4 & 1 \\ 3 & 7 & -1 \end{bmatrix}$

whose rows consist of the vectors α_1, α_2, α_3 and α_4.

We shall reduce the matrix A to echelon form by applying the row transformations. We have

$$A \sim \begin{bmatrix} 1 & -1 & -2 \\ 1 & 9 & 3 \\ -2 & -4 & 1 \\ 3 & 7 & -1 \end{bmatrix}, \text{ applying } R_1 \to \tfrac{1}{2} R_1$$

$$\sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & 10 & 5 \\ 0 & -6 & -3 \\ 0 & 10 & 5 \end{bmatrix}, \text{ by } \begin{array}{l} R_2 \to R_2 - R_1 \\ R_3 \to R_3 + 2R_1 \\ R_4 \to R_4 - 3R_1 \end{array}$$

$$\sim \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & \tfrac{1}{2} \\ 0 & 1 & \tfrac{1}{2} \\ 0 & 1 & \tfrac{1}{2} \end{bmatrix}, \text{ by } \begin{array}{l} R_2 \to \tfrac{1}{10} R_2 \\ R_3 \to -\tfrac{1}{6} R_3 \\ R_4 \to \tfrac{1}{10} R_4 \end{array}$$

$$\sim \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & \tfrac{1}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \text{ by } \begin{array}{l} R_3 \to R_3 - R_2 \\ R_4 \to R_4 - R_2 \end{array}$$

which is in echelon form.

We have rank A = the number of non-zero rows in its echelon form = 2.

$\therefore$ the maximum number of linearly independent row vectors in the matrix A = the rank A = 2.

The vectors α_1 and α_2 are linearly independent and so $\{\alpha_1, \alpha_2\}$ is a maximal linearly independent subsystem of the given system of vectors. We observe that none of the given four vectors is a scalar multiple of any of the remaining three vectors. So any two of the given four vectors form a maximal linearly independent subsystem of the given system of vectors.

Exercises

1. Fill up the blanks in the following statements:

   (i) Any set of vectors containing the zero vector as a member is linearly...... **(Meerut 1976)**

   (ii) Intersection of two linearly independent subsets of a vector space will be linearly...... **(Meerut 1976)**

   (iii) The subset $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ of the vector space $\mathbf{R}^3$ is linearly......

Scanned with CamScanner

(iv) A system consisting of a single non-zero vector is always linearly......

(v) In the vector space $\mathbf{R}^3$ the vectors $(1, 0, 1)$, $(3, 7, 0)$ and $(-1, 0, -1)$ are linearly......

2. State whether the following statements are true or false :—

(i) If A and B are subsets of a vector space, then
$A \neq B \Rightarrow L(A) \neq L(B)$. **(Meerut 19 5)**

(ii) A set containing a linearly independent set of vectors is itself linearly independent. **(Meerut 1976)**

(iii) Union of two linearly independent subsets of a vector space is linearly independent. **(Meerut 1976)**

(iv) The union of two subspaces of a vector space V is also a subspace of V.

(v) The intersection of two subspaces of a vector space V is also a subspace of V.

Ans. (i) False; (ii) false; (iii) false; (iv) false; (v) true.

3. Determine if the vectors $(1, -2, 1)$, $(2, 1, -1)$, $(7, -4, 1)$ in $\mathbf{R}^3$ are linearly independent. **(I.A.S. 1985)**

Ans. Linearly dependent.

4. In the vector space $\mathbf{R}^3$ express the vector $(1, -2, 5)$ as a linear combination of the vectors $(1, 1, 1)$, $(1, 2, 3)$ and $(2, -1, 1)$. **(Meerut 1976)**

Ans. $(1, -2, 5) = -6\,(1, 1, 1) + 3\,(1, 2, 3) + 2\,(2, -1, 1)$.

5. In the vector space $\mathbf{R}^4$ determine whether or not the vector $(3, 9, -4, 2)$ is a linear combination of the vectors $(1, -2, 0, 3)$, $(2, 3, 0, -1)$ and $(2, -1, 2, 1)$.

Ans. No. **(Meerut 1976)**

6. Prove that the four vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(1, 1, 1)$ in $V_3(\mathbf{C})$ form a linearly dependent set but any three of them are linearly independent. **(Meerut 1969)**

7. If α, β and γ are vectors such that $\alpha+\beta+\gamma=\mathbf{0}$, then α and β span the same subspace as β and γ. **(Meerut 1976, 80)**

§ 14. Some Theorems on linear dependence and linear independence.

Theorem 1. *Let $V(F)$ be a vector space. If $\alpha_1, \alpha_2, \ldots\ldots, \alpha_n$ are non-zero vectors $\in V$ then either they are linearly independent or some α_k, $2 \leqslant k \leqslant n$, is a linear combination of the preceding ones $\alpha_1, \alpha_2, \ldots, \alpha_{k-1}$.* **(Nagarjuna 1980, 91)**

Scanned with CamScanner

Proof. If $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$ are linearly independent we are nothing to prove. So let $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly dependent. Then there exists a relation of the form

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=0, \qquad \ldots(1)$$

where, not all the scalar coefficients $a_1, a_2, \ldots, a_n$ are 0. Let k be the largest integer for which $a_k \neq 0$, *i.e.*, $a_{k+1}=0$, $a_{k+2}=0$, …, $a_n=0$ and $a_k \neq 0$. There is no harm in this assumption because at the most if $a_n \neq 0$ then $k=n$.

Also $2 \leqslant k$. Because if $a_2=0, a_3=0, \ldots, a_n=0$, then $a_1\alpha_1=0$ and $\alpha_1 \neq 0 \Rightarrow a_1=0$. This contradicts the fact that not all the a's are 0.

Now the relation (1) reduces to

$$a_1\alpha_1+a_2\alpha_2+\ldots+a_k\alpha_k=0, \text{ where } a_k \neq 0$$

$$a_k\alpha_k=-a_1\alpha_1-a_2\alpha_2-\ldots-a_{k-1}\alpha_{k-1}$$

or $$a_k^{-1}(a_k\alpha_k)=a_k^{-1}(-a_1\alpha_1-a_2\alpha_2-\ldots-a_{k-1}\alpha_{k-1})$$

or $$\alpha_k=(-a_k^{-1}a_1)\,\alpha_1+(-a_k^{-1}a_2)\,\alpha_2+\ldots+(-a_k^{-1}a_{k-1})\,\alpha_{k-1}.$$

or Thus α_k is a linear combination of its preceding vectors.

Theorem 2. *The set of non-zero vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ of $V(F)$ is linearly dependent if some α_k, $2 \leqslant k \leqslant n$, is a linear combination of the preceding ones.* **(Nagarjuna 1980)**

Proof. If some α_k, $2 \leqslant k \leqslant n$, is a linear combination of the preceding ones $\alpha_1, \alpha_2, \ldots, \alpha_{k-1}$ then $\exists$ scalars $a_1, a_2, \ldots, a_{k-1}$ such that

$$\alpha_k=a_1\alpha_1+\ldots+a_{k-1}\,\alpha_{k-1}$$

$$\Rightarrow 1\alpha_k-a_1\alpha_1-a_2\alpha_2-\ldots-a_{k-1}\alpha_{k-1}=0$$

$$\Rightarrow \text{the set } \{\alpha_1, \alpha_2, \ldots, \alpha_k\} \text{ is linearly dependent.}$$

Hence the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of which $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is a subset, must be linearly dependent.

Theorem 3. *If in a vector space $V(F)$, a vector β is a linear combination of the set of vectors $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n$, then the set of vectors $\beta, \alpha_1, \alpha_2, \ldots, \alpha_n$ is linearly dependent.*

Proof. Since β is a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_n$, therefore there exist scalars $a_1, a_2, \ldots, a_n$ such that

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

$$\Rightarrow 1\beta-a_1\alpha_1-a_2\alpha_2-\ldots-a_n\alpha_n=0. \qquad \ldots(1)$$

In the relation (1) the scalar coefficient of β is 1 which is $\neq 0$. Hence in the relation (1) not all the scalar coefficients are 0. Therefore the set $\beta, \alpha_1, \ldots, \alpha_n$ is linearly dependent.

Theorem 4. *Let S be a linearly independent subset of a vector space V. Suppose β is a vector in V which is not in the subspace* *spanned by S. Then the set obtained by adjoining β to S is linearly independent.* **(Meerut 1977)**

Proof. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_n$ are distinct vectors in S. Let

$$c_1\alpha_1+c_2\alpha_2+\ldots+c_m\alpha_m+b\beta=\mathbf{0}. \quad \ldots(1)$$

Then b must be zero, for otherwise

$$\beta=\left(-\frac{c_1}{b}\right)\alpha_1+\ldots+\left(-\frac{c_m}{b}\right)\alpha_m$$

and consequently β is in the subspace spanned by S which is a contradiction.

Putting $b=0$ in (1), we get

$$c_1\alpha_1+c_2\alpha_2+\ldots+c_m\alpha_m=\mathbf{0}$$

$$\Rightarrow c_1=0,\ c_2=0,\ldots,\ c_m=0 \text{ because}$$

the set $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ is linearly independent since it is a subset of a linearly independent set S.

Thus the relation (1) implies

$$c_1=0,\ c_2=0,\ldots,\ c_m=0,\ b=0.$$

Therefore the set $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta\}$ is linearly independent. If S' is the set obtained by adjoining β to S, then we have proved that every finite subset of S' is linearly independent.

Hence S' is linearly independent.

Theorem 5. *The set of non-zero vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ of $V(F)$ is linearly dependent iff one of these vectors is a linear combination of the remaining $(n-1)$ vectors.*

Proof. This theorem can be easily proved.

§ 15. Basis of a Vector Space. Definition.

(Meerut 1990; Nagarjuna 90, Allahabad 76; S.V.U. Tirupati 90)

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

(i) S consists of linearly independent vectors.

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

Example 1. *A system S consisting of n vectors $e_1=(1, 0, 0, \ldots, 0)$, $e_2=(0, 1, 0, \ldots, 0), \ldots, e_n=(0, 0, \ldots, 0, 1)$ is a basis of $V_n(F)$.* **(Meerut 1990)**

Solution. First we should show that S is a linearly independent set of vectors. We have proved it in one of the previous examples.

Now we should prove that $L(S)=V_n(F)$. We have always $L(S) \subseteq V_n(F)$. So we should prove that $V_n(F) \subseteq L(S)$ *i.e.*, each vector in $V_n(F)$ is a linear combination of elements of S.

Let $\alpha=(a_1, a_2,..., a_n)$ be any vector in $V_n(F)$. We can write

$$(a_1, a_2,..., a_n)=a_1 (1, 0,..., 0)+a_2 (0, 1, 0,..., 0)+...$$
$$...+a_n (0, 0,..., 0, 1)$$

i.e.,

$$\alpha=a_1e_1+a_2e_2+...+a_ne_n.$$

Hence S is a basis of $V_n(F)$. We shall call this particular basis the **standard basis** of $V_n(F)$ or F^n.

Note. The set $\{(1, 0), (0, 1)\}$ is a basis of $V_2(F)$. The set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ is a basis of $V_3(F)$. As a particular case a basis of $F(F)$ is the set consisting of only the unity element of F.

Example 2. *Show that the infinite set*

$$S=\{1, x, x^2,..., x^n,...\}$$

is a basis of the vector space $F[x]$ *of polynomials over the field* F.

Solution. First we should prove that S is a linearly independent set of vectors. For proof refer some previous example.

Now we should show that S spans $F[x]$ *i.e.*, each polynomial in $F[x]$ can be expressed as a linear combination of a finite number of elements of S.

Let $f(x)=a_0+a_1x+a_2x^2+...+a_tx^t$ be a polynomial of degree t.

Then $f(x)=(a_0)\,1+a_1x+a_2x^2+...+a_tx^t$.

Hence S is a basis of $F[x]$.

Note. The vector space $F[x]$ has no finite basis. If we take any finite set S of polynomials, we can find a polynomial of degree greater than that of each of them. Such a polynomial cannot at any cost be expressed as a linear combination of the elements of S.

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

The vector space $V_n(F)$ of n-tuples is a finite dimensional vector space.

The vector space $F[x]$ of all polynomials over a field F is not finite dimensional. There exists no finite subset S of $F[x]$ which spans $F[x]$. *A vector space which is not finitely generated may be referred to as an infinite dimensional space.* Thus the vector space $F[x]$ of all polynomials over a field F is infinite dimensional.

Scanned with CamScanner

Existence of basis of a finite dimensional vector space.

Theorem. *There exi 's a basis for each finite dimensional vector space.* [Meerut 1987, 92; Allahabad 76]

Proof. Let $V(F)$ be a finitely generated vector space. Let $S=\{\alpha_1, \alpha_2, ..., \alpha_m\}$ be a finite subset of V such that $L(S)=V$. We may suppose that no member of S is $\mathbf{0}$.

If S is linearly independent, then S itself is a basis of V.

If S is linearly dependent, then there exists a vector $\alpha_i \in S$ which can be expressed as a linear combination of the preceding vectors $\alpha_1, \alpha_2, ..., \alpha_{i-1}$.

If we omit this vector α_i from S, then the remaining set S' of $m-1$ vectors

$$\alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_m$$

also generates V *i.e.*, $V=L(S')$. For if α is any element of V, then $L(S)=V$ implies that α can be written as a linear combination of $\alpha_1, \alpha_2, ..., \alpha_m$. Let $\alpha=a_1\alpha_1+a_2\alpha_2+...+a_{i-1}\alpha_{i-1}+a_i\alpha_i+a_{i+1}\alpha_{i+1}+...+a_m\alpha_m$. But α_i can be expressed as a linear combination of $\alpha_1, ..., \alpha_{i-1}$. Let $\alpha_i=b_1\alpha_1+...+b_{i-1}\alpha_{i-1}$. Putting this value of α_i in the expression for α, we get $\alpha=a_1\alpha_1+...+a_{i-1}\alpha_{i-1}+a_i(b_1\alpha_1+...+b_{i-1}\alpha_{i-1})+a_{i+1}\alpha_{i+1}+...+a_m\alpha_m$. Thus α has been expressed as a linear combination of the vectors $\alpha_1, \alpha_2, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_m$. In this way $\alpha \in V \Rightarrow \alpha$ can be expressed as a linear combination of the vectors belonging to the set S'. Thus S' generates V *i.e.*, $L(S')=V$.

If S' is linearly independent, then S' will be a basis of V. If S' is linearly dependent, then proceeding as above we shall get a new set of $n-2$ vectors which generates V. Continuing this process, we shall, after finite number of steps, obtain a linearly independent subset of S which generates V and which is therefore a basis of V.

At the most it may happen that we shall be left with a subset of S which contains only one non-zero vector and which spans V. We know that a set containing a single non-zero vector is definitely linearly independent and so it will form a basis of V.

Note. The above theorem may also be stated as below:

If a finite set S of vectors spans a finite dimensional vector space $V(F)$, there exists a subset of S which forms a basis of V.

Invariance of number of elements in the basis of a finite dimensional vector space.

Dimension theorem for vector spaces. *If $V(F)$ is a finite* *dimensional vector space, then any two bases of V have the same number of elements.* (Nagarjuna 1991; Tirupati 90, 93; Meerut 81, 82, 87, 89, 93; Poona 72; Allahabad 79)

Proof. Suppose $V(F)$ is a finite dimensional vector space. Then V definitely possesses a basis. Let

$$S_1=\{\alpha_1, \alpha_2,\ldots,\alpha_m\}$$

and $$S_2=\{\beta_1, \beta_2,\ldots,\beta_n\}$$

be two bases of V. We shall prove that $m=n$.

Since $V=L(S_1)$ and $\beta_1 \in V$, therefore β_1 can be expressed as a linear combination of $\alpha_1, \alpha_2,\ldots, \alpha_n$. Consequently the set $S_3=\{\beta_1, \alpha_1, \alpha_2,\ldots, \alpha_n\}$ which also obviously generates $V(F)$ is linearly dependent. Therefore there exists a member $\alpha_i \neq \beta_1$ of this set S_3 such that α_i is a linear combination of the preceding vectors $\beta_1, \alpha_1, \alpha_2,\ldots,\alpha_{i-1}$. if we omit the vector α_i from S_3 then V is also generated by the remaining set

$$S_4=\{\beta_1, \alpha_1, \alpha_2,\ldots,\alpha_{i-1}, \alpha_{i+1},\ldots,\alpha_m\}.$$

Since $V=L(S_4)$ and $\beta_2 \in V$, therefore β_2 can be expressed as a linear combination of the vectors belonging to S_4. Consequently the set

$$S_5=\{\beta_2, \beta_1, \alpha_1, \alpha_2,\ldots,\alpha_{i-1}, \alpha_{i+1},\ldots,\alpha_m\}$$

is linearly dependent. Therefore there exists a member α_j of this set S_5 such that α_j is a linear combination of the preceding vectors. Obviously α_j will be different from β_1 and β_2 since $\{\beta_1, \beta_2\}$ is a linearly independent set. If we exclude the vector α_j from S_5, then the remaining set will generate $V(F)$.

We may continue to proceed in this manner. Here each step consists in the exclusion of an α and the inclusion of a β in the set S_1.

Obviously the set S_1 of α's cannot be exhausted before the set S_2 of β's otherwise $V(F)$ will be a linear span of a proper subset of S_2 and thus S_2 will become linearly dependent. Therefore we must have

$$m \nless n.$$

Interchanging the roles of S_1 and S_2, we shall get that

$$n \nless m.$$

Hence $$n = m.$$

Example. For the vector space V_3, the set

$$S_1=\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

and $$S_2=\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$$

Scanned with CamScanner

are bases as can easily be seen. Both these bases contain the same number of elements *i.e.*, 3.

Dimension of a finitely generated vector space Definition. *The number of elements in any basis of a finite dimensional vector space $V(F)$ is called the dimension of the vector space $V(F)$ and will be denoted by dim V.* **(Marathwada 1971; S.V.U. Tirupati 93; Nagarjuna 80)**

The vector space $V_n(F)$ is of dimension n. The vector space $V_3(F)$ is of dimension 3. If a field F is regarded as a vector space over F, then F will be of dimension 1 and the set $S=\{1\}$ consisting of unity element of F alone is a basis of F. In fact every non-zero element of F will form a basis of F.

§ 17. Some Properties of finite dimensional vector spaces.

Theorem 1. Extension theorem. *Every linearly independent subset of a finitely generated vector space $V(F)$ forms a part of a basis of V.*

Or

Every linearly independent subset of a finitely generated vector space $V(F)$ is either a basis of V or can be extended to form a basis of V. **(Meerut 1980, 84; Nagarjuna 90, 91; Andhra 90; Allahabad 76)**

Proof. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a linearly independent subset of a finite dimensional vector space $V(F)$. If dim $V=n$, then V has a finite basis say, $\{\beta_1, \beta_2, \ldots, \beta_n\}$. Consider the set

$$S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_m; \beta_1, \beta_2, \ldots, \beta_n\}.$$

Obviously $L(S_1)=V$. Since the α's can be expressed as linear combinations of the β's therefore the set S_1 is linearly dependent.

Therefore there is some vector of S_1 which is a linear combination of its preceding vectors This vector cannot be any of the α's, since the α's are linearly independent. Therefore this vector must be some β, say β_i. Now omit the vector β_i from S_1 and consider the set

$$S_2=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_n\}.$$

Obviously $L(S_2)=V$. If S_2 is linearly independent, then S_2 will be a basis of V and it is the required extended set which is a basis of V. If S_2 is not linearly independent, then repeating the above process a finite number of times, we shall get a linearly independent set containing $\alpha_1, \alpha_2, \ldots, \alpha_m$ and spanning V. This set will be a basis of V and it will contain S. Since each basis of V

contains the same number of elements, therefore exactly $n-m$ elements of set of β's will be adjoined to S so as to form a basis of V.

Theorem 2. *Each set of $(n+1)$ or more vectors of a finite dimensional vector space $V(F)$ of dimension n is linearly dependent.*

(Nagarjuna 1978; Andhra 92; Meerut 68)

Proof. Let $V(F)$ be a finite dimensional vector space of dimension n. Let S be a linearly independent subset of V containing $(n+1)$ or more vectors. Then S will form a part of basis of V. Thus we shall get a basis of V containing more than n vectors. But every basis of V will contain exactly n vectors. Hence our assumption is wrong. Therefore if S contains $(n+1)$ or more vectors, then S must be linearly dependent.

Theorem 3. *Let V be a vector space which is spanned by a finite set of vectors $\beta_1, \beta_2, \ldots, \beta_m$. Then any linearly independent set of vectors in V is finite and contains no more than m vectors.*

(Meerut 1972, 73, 85; Allahabad 79; Nagarjuna 78)

Proof. Let $S=\{\beta_1, \beta_2, \ldots, \beta_m\}$.

Since $L(S)=V$, therefore V has a finite basis and dim $V \leqslant m$. Hence every subset S' of V which contains more than m vectors is linearly dependent. This proves the theorem.

Theorem 4. *If $V(F)$ is a finite dimensional vector space of dimension n, then any set of n linearly independent vectors in V forms a basis of V.* **(Andhra 1992)**

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linearly independent subset of a finite dimensional vector space $V(F)$ of dimension n. If S is not a basis of V, then it can be extended to form a basis of V. Thus we shall get a basis of V containing more than n vectors.

But every basis of V must contain exactly n vectors. Therefore our assumption is wrong and S must be a basis of V.

Theorem 5. *If a set S of n vectors of a finite dimensional vector space $V(F)$ of dimension n generates $V(F)$, then S is a basis of V.* **(Andhra 1992)**

Proof. Let $V(F)$ be a finite dimensional vector space of dimension n. Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a subset of V such that $L(S)=V$. If S is linearly independent, then S will form a basis of V. If S is not linearly independent, then there will exist a proper subset of S which will form a basis of V. Thus we shall get a basis of V containing less than n elements. But every basis of V must

Scanned with CamScanner

contain exactly n elements. Hence S cannot be linearly dependent and so S must be a basis of V.

Note. *If V is a finite dimensional vector space of dimension n, then V cannot be generated by fewer than n vectors.*

Theorem 6. Dimension of a subspace.

Each subspace W of a finite dimensional vector space $V(F)$ of dimension n is a finite dimensional space with dim $m \leqslant n$.

Also $V=W$ iff dim $V=$dim W. **(Nagarjuna 1980; Andhra 92; Kakatiya 91; Poona 72)**

Proof. Let $V(F)$ be a finite dimensional vector space of dim n. Let W be a subspace of V. Any subset of W containing $(n+1)$ or more vectors is also a subset of V and any $(n+1)$ vectors in V are linearly dependent. Therefore any linearly independent set of vectors in W can contain, at the most n vectors. Let

$$S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$$

be a linearly independent subset of W with maximum number of elements. We claim that S is a basis of W. The proof is as follows :

(i) S is a linearly independent subset of W.

(ii) $L(S)=W$ Let α be any element of W.

Then the $(m+1)$ vectors $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_m$ belonging to W are linearly dependent because we have supposed that the largest independent subset of W contains m vectors.

Now $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha\}$ is a linearly dependent set. Therefore there exists a vector belonging to it which can be expressed as a linear combination of the preceding vectors. Since $\alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent, therefore this vector cannot be any of these m vectors. So it must be α itself. Thus α can be expressed as a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_m$. Hence $L(S)=W$.

$\therefore$ S is a basis of W.

$\therefore$ dim $W=m$ and $m \leqslant n$.

Now if $V=W$, then every basis of V is also a basis of W. Hence dim $V=$dim $W=n$.

Conversely let dim $W=$dim $V=n$. Then to prove that $W=V$.

Let S be a basis of W. Then $L(S)=W$ and S contains n vectors. Since S is also a subset of V and S contains n linearly independent vectors, therefore S will also be a basis of V. Therefore $L(S)=V$. Hence $W=V$. We thus conclude :

If W is a proper subspace of a finite-dimensional vector space V, then W is finite dimensional and $\dim W < \dim V$.

Theorem 7. *If W is a subspace of a finite-dimensional vector space V, every linearly independent subset of W is finite and is part of a (finite) basis for W.* **(Meerut 1988)**

Proof. Let $\dim V = n$. Let W be a subspace of V. Let S_0 be a linearly independent subset of W. Let S be a linearly independent subset of W which contains S_0 and which has maximum number of elements. Then S is also a linearly independent subset of V. So S will have at the most n elements. Therefore S is finite and consequently S_0 is finite.

Now our claim is that S is a basis of W. The proof is as follows :

(i) S is a linearly independent subset of W.

(ii) $L(S) = W$. Because if $\beta \in W$, then β must be in the linear span of S. If β is not in the linear span of S, then the subset of W obtained by adjoining β to S will be linearly independent. Thus S will not remain the maximum linearly independent subset of W containing S_0. Hence $\beta \in W \Rightarrow \beta \in L(S)$. Thus $W \subseteq L(S)$. Since S is a subset of the subspace W, therefore $L(S) \subseteq W$.

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

Thus S is a finite basis of W and $S_0 \subseteq S$.

Theorem 8. *Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of a finite dimensional vector space $V(F)$ of dimension n. Then every element α of V can be* **uniquely** *expressed as*

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n \text{ where } a_1, a_2, \ldots, a_n \in F.$$

(Meerut 1989)

Proof. Since S is a basis of V, therefore $L(S) = V$. Therefore any vector $\alpha \in V$ can be expressed as

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

To show uniqueness let us suppose that

$$\alpha = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n.$$

Then we must show that $a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$.

We have $a_1\alpha_1 + a_2\alpha_2 + \ldots + a_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n$

$\Rightarrow (a_1 - b_1)\alpha_1 + (a_2 - b_2)\alpha_2 + \ldots + (a_n - b_n)\alpha_n = 0$

$\Rightarrow a_1 - b_1 = 0, a_2 - b_2 = 0, \ldots, a_n - b_n = 0$ since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly independent

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

Hence the theorem.

Scanned with CamScanner

Theorem 9. *If W_1, W_2 are two subspaces of a finite dimensional vector space $V(F)$, then*

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

(Andhra 1992; Kakatiya 91; I.A.S 85, 86, 88; Marathwada 71; Poona 70; Nagarjuna 74; Tirupati 90; Allahabad 77; Meerut 80, 83, 85, 87, 89, 92)

Proof. Let $\dim (W_1 \cap W_2)=k$ and let the set

$$S=\{\gamma_1, \gamma_2, \gamma_3, \dots, \gamma_k\}$$

be a basis of $W_1 \cap W_2$. Then $S \subseteq W_1$ and $S \subseteq W_2$.

Since S is linearly independent and $S \subseteq W_1$, therefore S can be extended to form a basis of W_1. Let

$$\{\gamma_1, \gamma_2, \dots, \gamma_k, \alpha_1, \alpha_2, \dots, \alpha_m\}$$

be a basis of W_1. Then $\dim W_1=k+m$. Similarly let

$$\{\gamma_1, \gamma_2, \dots, \gamma_k, \beta_1, \beta_2, \dots, \beta_t\}$$

be a basis of W_2. Then $\dim W_2=k+t$.

$\therefore$ $\dim W_1+\dim W_2-\dim (W_1 \cap W_2)=(m+k)+(k+t)-k$

$$=k+m+t.$$

$\therefore$ to prove the theorem we must show that

$$\dim (W_1+W_2)=k+m+t.$$

We claim that the set

$$S_1=\{\gamma_1, \gamma_2, \dots, \gamma_k, \alpha_1, \alpha_2, \dots, \alpha_m, \beta_1, \beta_2, \dots, \beta_t\}$$

is a basis of W_1+W_2.

First we show that S_1 is linearly independent. Let

$$c_1\gamma_1+c_2\gamma_2+\dots+c_k\gamma_k+a_1\alpha_1+a_2\alpha_2+\dots+a_m\alpha_m+b_1\beta_1+b_2\beta_2 +\dots+b_t\beta_t=0 \qquad \dots(1)$$

$$\Rightarrow b_1\beta_1+b_2\beta_2+\dots+b_t\beta_t=-(c_1\gamma_1+\dots+c_k\gamma_k+a_1\alpha_1+\dots+a_m\alpha_m) \qquad \dots(2)$$

Now $-(c_1\gamma_1+\dots+c_k\gamma_k+a_1\alpha_1+\dots+a_m\alpha_m) \in W_1$ since it is a linear combination of a basis of W_1. Again

$$b_1\beta_1+b_2\beta_2+\dots+b_t\beta_t \in W_2$$

since it is a linear combination of elements belonging to a basis of W_2.

Also by virtue of the equality (2), $b_1\beta_1+\dots+b_t\beta_t \in W_1$. Therefore $b_1\beta_1+b_2\beta_2+\dots+b_t\beta_t \in W_1 \cap W_2$. Therefore it can be expressed as a linear combination of the basis of $W_1 \cap W_2$. Thus we have a relation of the form

$$b_1\beta_1+b_2\beta_2+\dots+b_t\beta_t=d_1\gamma_1+\dots+d_k\gamma_k$$

$$\Rightarrow b_1\beta_1+b_2\beta_2+\dots+b_t\beta_t-d_1\gamma_1-d_2\gamma_2-\dots-d_k\gamma_k=0.$$

But $\beta_1, \beta_2, \dots, \beta_t, \gamma_1, \dots, \gamma_k$ are linearly independent vectors.

Scanned with CamScanner

Therefore we must have $b_1=0, b_2=0,\ldots, b_t=0$.

Putting these values of b's in (1), it reduces to

$$c_1\gamma_1+c_2\gamma_2+\ldots+c_k\gamma_k+a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m=0$$

$$\Rightarrow c_1=0, c_2=0,\ldots, c_k=0, a_1=0, a_2=0,\ldots, a_m=0$$

since the vectors $\gamma_1, \gamma_2,\ldots, \gamma_k, \alpha_1, \alpha_2,\ldots, \alpha_m$ are linearly independent. Thus the relation (1) implies that

$$c_1=0, c_2=0,\ldots, c_k=0, a_1=0,\ldots, a_m=0, b_1=0,\ldots, b_t=0.$$

Therefore the set S_1 of vectors

$$\gamma_1,\ldots, \gamma_k, \alpha_1,\ldots, \alpha_m, \beta_1,\ldots, \beta_t$$

is linearly independent.

Now to show that $L(S_1)=W_1+W_2$.

Since W_1+W_2 is a subspace of V and each element of S_1 belongs to W_1+W_2, therefore $L(S_1)\subseteq W_1+W_2$.

Again let α be any element of W_1+W_2. Then

$\alpha=$some element of W_1+some element of W_2

$=$a linear combination of elements of basis of W_1+a linear combination of elements of basis of W_2

$=$a linear combination of elements of S_1.

$\therefore\quad \alpha\in L(S_1)$. Hence $W_1+W_2\subseteq L(S_1)$.

$\therefore\quad L(S_1)=W_1+W_2$.

$\therefore\quad S_1$ is a basis of W_1+W_2 and consequently

$$\dim(W_1+W_2)=k+m+t.$$

Hence the theorem.

Solved Examples

Ex. 1. *Let V be the vector space of all 2×2 matrices over the field F. Prove that V has dimension 4 by exhibiting a basis for V which has 4 elements.* **(Meerut 1979; I.A.S. 85; Nagarjuna 78)**

Sol. Let $\alpha=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, $\beta=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$, $\gamma=\begin{bmatrix}0 & 0\\ 1 & 0\end{bmatrix}$

and $\delta=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}$ be four elements of V.

The subset $S=\{\alpha, \beta, \gamma, \delta\}$ of V is linearly independent because

$$a\alpha+b\beta+c\gamma+d\delta=0$$

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

$$\Rightarrow \begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

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

Scanned with CamScanner

Also $L(S)=V$ because if $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is any vector in V, then we can write

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\alpha + b\beta + c\gamma + d\delta.$$

Therefore S is a basis of V. Since the number of elements in S is 4, therefore dim $V=4$.

Ex. 2. *Show that if $S=\{\alpha, \beta, \gamma\}$ is a basis of* C^3 (C), *then the set*

$$S'=\{\alpha+\beta, \beta+\gamma, \gamma+\alpha\}$$

is also a basis of C^3 (C).

Sol. The vector space C^3 (C) is of dimension 3. Any subset of C^3 having three linearly independent vectors will form a basis of C^3. We have shown in one of the previous examples that if S is a linearly independent subset of C^3, then S' is also a linearly independent subset of C^3. (Give the proof here).

Therefore S' is also a basis of C^3.

Ex. 3. *If W_1 and W_2 are finite-dimensional subspaces with the same dimension, and if $W_1 \subseteq W_2$, then $W_1 = W_2$.*

Sol. Since $W_1 \subseteq W_2$, therefore W_1 is also a subspace of W_2. Now dim W_1=dim W_2. Therefore we must have $W_1 = W_2$.

Ex. 4. *Let V be the vector space of ordered pairs of complex numbers over the real field* **R** *i.e., let V be the vector space* **C (R)**. *Show that the set $S=\{(1, 0), (i, 0), (0, 1), (0, i)\}$ is a basis for V.*

Sol. **S is linearly independent.** We have

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

where $a, b, c, d \in \mathbf{R}$

$$\Rightarrow (a+ib, c+id)=(0, 0)$$

$$\Rightarrow a+ib=0, c+id=0$$

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

Therefore S is linearly independent.

Now we shall show that $L(S)=V$. Let any ordered pair $(a+ib, c+id) \in V$ where $a, b, c, d \in \mathbf{R}$. Then as shown above we can write $(a+ib, c+id)=a\,(1, 0)+b\,(i, 0)+c\,(0, 1)+d\,(0, i)$. Thus any vector in V is expressible as a linear combination of elements of S. Therefore $L(S)=V$ and so S is a basis for V.

Ex. 5. *In the vector space* $\mathbf{R}^3$, *let* $\alpha=(1, 2, 1)$, $\beta=(3, 1, 5)$, $\gamma=(3, -4, 7)$. *Show that there exists more than* [illegible] *basis for the subspace spanned by the set* $S=\{\alpha, \beta, \gamma\}$. (M[illegible]wada 1971)

Sol. First show that vector γ can be expressed as a linear combination of the vectors α and β. Therefore if $T=\{\alpha, \beta\}$, then $L(T)=L(S)$. Now the set $\{\alpha, \beta\}$ is linearly independent as can be easily shown. Therefore the set $\{\alpha, \beta\}$ is a basis for $L(T)=L(S)$. Therefore dim $L(S)=2$. Now $\{\alpha, \gamma\}$ is a linearly independent subset of $L(S)$ containing two vectors. Therefore $\{\alpha, \gamma\}$ is also a basis for $L(S)$. Similarly $\{\beta, \gamma\}$ is also a basis for $L(S)$.

Ex. 6. *Show that the vectors* $(1, 2, 1), (2, 1, 0), (1, -1, 2)$ *form a basis of* $\mathbf{R}^3$.

(Nagarjuna 1980; Meerut 90)

Sol. We know that the set $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ forms a basis for $\mathbf{R}^3$. Therefore dim $\mathbf{R}^3=3$. If we show that the set $S=\{(1, 2, 1), (2, 1, 0), (1, -1, 2)\}$ is linearly independent, then this set will also form a basis for $\mathbf{R}^3$. [See theorem 4 of § 17]

We have

$$a_1 (1, 2, 1)+a_2 (2, 1, 0)+a_3 (1, -1, 2)=(0, 0, 0)$$

$$\Rightarrow (a_1+2a_2+a_3, 2a_1+a_2-a_3, a_1+2a_3)=(0, 0, 0).$$

$$\therefore \quad a_1+2a_2+a_3=0 \qquad \ldots(1)$$

$$2a_1+a_2-a_3=0 \qquad \ldots(2)$$

$$a_1+2a_3=0. \qquad \ldots(3)$$

Now we shall solve these equations to get the values of a_1, a_2, a_3. Multiplying the equation (2) by 2, we get

$$4a_1+2a_2-2a_3=0. \qquad \ldots(4)$$

Subtracting (4) from (1), we get

$$-3a_1+3a_3=0$$

or

$$-a_1+a_3=0. \qquad \ldots(5)$$

Adding (3) and (5), we get $3a_3=0$ or $a_3=0$. Putting $a_3=0$ in (3), we get $a_1=0$. Now putting $a_3=0$ and $a_1=0$ in (1), we get $a_2=0$.

Thus solving the equations (1), (2) and (3), we get $a_1=0$, $a_2=0$, $a_3=0$. Therefore the set S is linearly independent. Hence it forms a basis for $\mathbf{R}^3$.

Ex. 7. *Determine whether or not the following vectors form a basis of* $\mathbf{R}^3$:

$$(1, 1, 2), (1, 2, 5), (5, 3, 4).$$

(Meerut 1980)

Sol. We know that dim $\mathbf{R}^3=3$. If the given set of vectors is linearly independent, it will form a basis of $\mathbf{R}^3$ otherwise not. We have

$$a_1 (1, 1, 2)+a_2 (1, 2, 5)+a_3 (5, 3, 4)=(0, 0, 0)$$

Scanned with CamScanner

$\Rightarrow (a_1+a_2+5a_3, a_1+2a_2+3a_3, 2a_1+5a_2+4a_3)=(0, 0, 0)$.

$$\therefore \left.\begin{array}{r} a_1+a_2+5a_3=0 \\ a_1+2a_2+3a_3=0 \\ 2a_1+5a_2+4a_3=0 \end{array}\right\}. \quad \begin{array}{r} \ldots(1) \\ \ldots(2) \\ \ldots(3) \end{array}$$

Now we shall solve these equations to get the values of a_1, a_2, a_3. Subtracting (2) from (1), we get

$$-a_2+2a_3=0. \qquad \ldots(4)$$

Multiplying (1) by 2, we get

$$2a_1+2a_2+10a_3=0. \qquad \ldots(5)$$

Subtracting (5) from (3), we get

$$3a_2-6a_3=0$$

or

$$a_2-2a_3=0. \qquad \ldots(6)$$

We see that the equations (4) and (6) are the same and give $a_2=2a_3$. Putting $a_2=2a_3$ in (1), we get $a_1=-7a_3$. If we put $a_3=1$, we get $a_2=2$ and $a_1=-7$. Thus $a_1=-7$, $a_2=2$, $a_3=1$ is a non-zero solution of the equations (1), (2) and (3). Hence the given set is linearly dependent so it does not form a basis of $\mathbf{R}^3$.

Ex. 8. *For the 3-dimensional space* $\mathbf{R}^3$ *over the field of real numbers* $\mathbf{R}$, *determine if the set* $\{(2, -1, 0), (3, 5, 1), (1, 1, 2)\}$ *is a basis.*

Sol. We have dim $\mathbf{R}^3=3$. If the given set containing three vectors is linearly independent, it will form a basis of $\mathbf{R}^3$ otherwise not.

Let $a, b, c \in \mathbf{R}$ be such that

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

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

$$\therefore \quad 2a+3b+c=0, \qquad \ldots(1)$$

$$-a+5b+c=0, \qquad \ldots(2)$$

and

$$b+2c=0. \qquad \ldots(3)$$

Now we shall solve these equations to get the values of a, b, c.

Multiplying (2) by 2 and adding to (1), we get

$$13b+3c=0. \qquad \ldots(4)$$

Multiplying (3) by 13 and then subtracting (4) from it, we get

$$23c=0 \text{ or } c=0.$$

Putting $c=0$ in (3), we get $b=0$.

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

Thus the only solution of the equations (1), (2) and (3) is

$a=0,\ b=0,\ c=0$. Therefore the three given vectors are linearly independent and so they form a basis of $\mathbf{R}^3$.

Ex. 9. *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* $\alpha_1, \alpha_2, \alpha_3$.

(Meerut 1981, 84P, 93P)

Sol. Let $\quad S=\{\alpha_1, \alpha_2, \alpha_3\}$.

First we shall show that the set S is linearly independent. 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., $\quad a(1, 0, -1)+b(1, 2, 1)+c(0, -3, 2)=(0, 0, 0)$

i.e., $\quad (a+b+0c,\ 0a+2b-3c,\ -a+b+2c)=(0, 0, 0)$

i.e.,
$$a+b=0 \qquad \ldots(1)$$
$$2b-3c=0 \qquad \ldots(2)$$
$$-a+b+2c=0. \qquad \ldots(3)$$

Adding both sides of the equations (1) and (3), we get

$$2b+2c=0 \qquad \ldots(4)$$

Subtracting (2) from (4), we get $5c=0$ or $c=0$.

Putting $c=0$ in (2), we get $b=0$ and then putting $b=0$ in (1), we get $a=0$.

Thus $a=0,\ b=0,\ c=0$ is the only solution of the equations (1), (2) and (3) and so

$$a\alpha_1+b\alpha_2+c\alpha_3=\mathbf{0} \Rightarrow a=0, b=0, c=0.$$

$\therefore$ The vectors $\alpha_1, \alpha_2, \alpha_3$ are linearly independent.

Now we shall show that the vectors $\alpha_1, \alpha_2, \alpha_3$ also generate $\mathbf{R}^3$. Let $\gamma=(p, q, r)$ be any vector in $\mathbf{R}^3$ and let

$$\gamma=(p, q, r)=x\alpha_1+y\alpha_2+z\alpha_3,\ \ x, y, z \in \mathbf{R}$$
$$=x(1, 0, -1)+y(1, 2, 1)+z(0, -3, 2). \qquad \ldots(5)$$

Then
$$x+y=p \qquad \ldots(6)$$
$$2y-3z=q \qquad \ldots(7)$$
$$-x+y+2z=r \qquad \ldots(8)$$

Adding both sides of equations (6) and (8), we get

$$2y+2z=p+r. \qquad \ldots(9)$$

Subtracting (7) from (9), we get

$$5z=p+r-q \ \text{ or } \ z=\tfrac{1}{5}p-\tfrac{1}{5}q+\tfrac{1}{5}r.$$

Then from (9), we get

$$y=-z+\tfrac{1}{2}p+\tfrac{1}{2}r=\tfrac{3}{10}p+\tfrac{1}{5}q+\tfrac{3}{10}r$$

and from (6), we get

$$x=p-y=p-\tfrac{3}{10}p-\tfrac{1}{5}q-\tfrac{3}{10}r=\tfrac{7}{10}p-\tfrac{1}{5}q-\tfrac{3}{10}r.$$

Scanned with CamScanner

Thus every vector $\gamma=(p, q, r)$ in $\mathbf{R}^3$ can be expressed as $\gamma=x\alpha_1+y\alpha_2+z\alpha_3$, where $x, y, z \in \mathbf{R}$ are as found above.

$\therefore$ the set S generates $\mathbf{R}^3$.

Since S is linearly independent and it also generates $\mathbf{R}^3$, therefore it is a basis of $\mathbf{R}^3$.

The relation (5) expresses the vector $\gamma=(p, q, r)$ as a linear combination of α_1, α_2 and α_3.

The standard basis vectors are

$$e_1=(1, 0, 0),\ e_2=(0, 1, 0) \text{ and } e_3=(0, 0, 1).$$

If $\gamma=e_1$, then $p=1$, $q=0$, $r=0$ and so

$$x=\tfrac{7}{10},\ y=\tfrac{3}{10} \text{ and } z=\tfrac{1}{5}.$$

$$\therefore\ e_1=\tfrac{7}{10}\alpha_1+\tfrac{3}{10}\alpha_2+\tfrac{1}{5}\alpha_3.$$

If $\gamma=e_2$, then $p=0$, $q=1$, $r=0$ and so

$$x=-\tfrac{1}{5},\ y=\tfrac{1}{5} \text{ and } z=-\tfrac{1}{5}.$$

$$\therefore\ e_2=-\tfrac{1}{5}\alpha_1+\tfrac{1}{5}\alpha_2-\tfrac{1}{5}\alpha_3.$$

Finally if $\gamma=e_3$, then $p=0$, $q=0$, $r=1$ and so

$$x=-\tfrac{3}{10},\ y=\tfrac{3}{10} \text{ and } z=\tfrac{1}{5}.$$

$$\therefore\ e_3=-\tfrac{3}{10}\alpha_1+\tfrac{3}{10}\alpha_2+\tfrac{1}{5}\alpha_3.$$

Ex. 10. *Show that the set* $\{(1, i, 0), (2i, 1, 1), (0, 1+i, 1-i)\}$ *is a basis for* $V_3(\mathbf{C})$. **(Meerut 1981)**

Sol. We know that dim $V_3(\mathbf{C})$ or $\mathbf{C}^3=3$. If the given set containing three vectors is linearly independent it will form a basis of $V_3(\mathbf{C})$ otherwise not.

Let $a, b, c \in \mathbf{C}$ be such that

$$a(1, i, 0)+b(2i, 1, 1)+c(0, 1+i, 1-i)=(0, 0, 0)$$

$$\Rightarrow (a+2ib+0c,\ ai+b+c[1+i],\ 0a+b+c[1-i])=(0, 0, 0).$$

$$\therefore\quad a+2ib=0, \quad \text{...(1)}$$

$$ai+b+c(1+i)=0, \quad \text{...(2)}$$

and

$$b+c(1-i)=0. \quad \text{...(3)}$$

Now we shall solve these equations to get the values of a, b, c.

Multiplying (1) by $-i$ and adding to (2), we get

$$3b+c(1+i)0=. \quad \text{...(4)}$$

Multiplying (3) by 3 and subtracting from (4), we get

$$c(1+i)-3c(1-i)=0$$

or $c(1+i-3+3i)=0$ or $c(-2+4i)=0$ or $c=0$.

Putting $c=0$ in (3), we get $b=0$.

Putting $b=0$ in (1), we get $a=0$.

Thus the only solution of the equations (1), (2) and (3) is

Scanned with CamScanner

$a=0, b=0, c=0$. Therefore the three given vectors are linearly independent and so they form a basis of $V_3(\mathbf{C})$.

Ex. 11. *Show that a systèm X consisting of the vectors* $\alpha_1=(1, 0, 0, 0)$, $\alpha_2=(0, 1, 0, 0)$, $\alpha_3=(0, 0, 1, 0)$ *and* $\alpha_4=(0, 0, 0, 1)$ *is a basis set of* $\mathbf{R}^4(\mathbf{R})$.

Sol. First we show that the set X is a linearly independent set of vectors.

If a_1, a_2, a_3, a_4 be any scalars *i.e.*, elements of the field $\mathbf{R}$, then

$$a_1\alpha_1+a_2\alpha_2+a_3\alpha_3+a_4\alpha_4=\text{zero vector}$$

$$\Rightarrow a_1 (1, 0, 0, 0)+a_2 (0, 1, 0, 0)+a_3 (0, 0, 1, 0)+a_4 (0, 0, 0, 1) =(0, 0, 0, 0)$$

$$\Rightarrow (a_1, a_2, a_3, a_4)=(0, 0, 0, 0)$$

$$\Rightarrow a_1=0, a_2=0, a_3=0, a_4=0.$$

Therefore the given set X of four vectors is linearly independent.

Now we shall show that X generates $\mathbf{R}^4$ *i.e.*, each vector of $\mathbf{R}^4$ can be expressed as a linear combination of the vectors of X.

Let (a, b, c, d) be any vector in $\mathbf{R}^4$. We can write

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

$$=a\alpha_1+b\alpha_2+c\alpha_3+d\alpha_4.$$

Thus (a, b, c, d) has been expressed as a linear combination of the vectors of X and so X generates $\mathbf{R}^4$.

Since X is a linearly independent subset of $\mathbf{R}^4$ and it also generates $\mathbf{R}^4$, therefore it is a basis of $\mathbf{R}^4$.

Ex. 12. *Show that the set* $S=\{1, x, x^2,\ldots,x^n\}$ *of* $n+1$ *polynomials in* x *is a basis of the vector space* $P_n(\mathbf{R})$, *of all polynomials in* x (*of degree at most* n) *over the field of real numbers.*

(I.A.S. 1977; Meerut 74)

Sol. $P_n(\mathbf{R})$ is the vector space of all polynomials in x (of degree at most n) over the field $\mathbf{R}$ of real numbers.

$S=\{1, x, x^2,\ldots,x^n\}$ is a subset of P_n consisting of $n+1$ polynomials. To prove that S is a basis of the vector space $P_n(\mathbf{R})$.

First we show that the vectors in the set S are linearly independent over the field $\mathbf{R}$.

The zero vector of the vector space $P_n(\mathbf{R})$ is the zero polynomial. Let $a_0, a_1, a_2,\ldots,a_n \in \mathbf{R}$ be such that

$$a_0 (1)+a_1x+a_2x^2+\ldots+a_nx^n=0 \text{ i.e., zero polynomial.}$$

Scanned with CamScanner

Now by the definition of the equality of two polynomials, we have

$$a_0+a_1x+a_2x^2+\ldots+a_nx^n=0$$

$$\Rightarrow \quad a_0=0,\ a_1=0,\ a_2=0,\ldots,a_n=0.$$

$\therefore$ the vectors $1, x, x^2, \ldots, x^n$ of the vector space $P_n(\mathbf{R})$ are linearly independent.

Now we shall show that the set S generates the whole vector space $P_n(\mathbf{R})$.

Let $\alpha=a_0+a_1x+a_2x^2+\ldots+a_nx^n$ be any arbitrary member of P_n, where $a_0, a_1, \ldots, a_n \in \mathbf{R}$. Then α is a linear combination of the polynomials $1, x, x^2, \ldots, x^n$ over the field $\mathbf{R}$. Therefore S generates $P_n(\mathbf{R})$.

Since S is a linearly independent subset of $P_n(\mathbf{R})$ and it also generates $P_n(\mathbf{R})$, therefore it is a basis of $P_n(\mathbf{R})$.

Ex. 13. *Select a basis, if any, of* $\mathbf{R}^3(\mathbf{R})$ *from the set* $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, *where* $\alpha_1=(1, -3, 2)$, $\alpha_2=(2, 4, 1)$, $\alpha_3=(3, 1, 3)$, $\alpha_4=(1, 1, 1)$.

Sol. Let $S=\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$.

If any three vectors in S are linearly independent, then they will form a basis of the vector space $\mathbf{R}^3(\mathbf{R})$.

First consider the set $S_1=\{\alpha_1, \alpha_2, \alpha_3\}$. Let us see whether the vectors in the set S_1 are linearly independent or not.

The determinant of order 3 whose columns consist of the coordinates of the vectors $\alpha_1, \alpha_2, \alpha_3$ is

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

$$=\begin{vmatrix} 1 & 0 & 0 \\ -3 & 10 & 10 \\ 2 & -3 & -3 \end{vmatrix}, \text{ by } C_2-2C_1 \text{ and } C_3-3C_1$$

$$=-30+30=0.$$

$\therefore$ the vectors $\alpha_1, \alpha_2, \alpha_3$ are linearly dependent and so they do not form a basis of $\mathbf{R}^3(\mathbf{R})$.

Now consider the set $S_2=\{\alpha_1, \alpha_2, \alpha_4\}$.

The determinant of order 3 whose columns consist of the coordinates of the vectors $\alpha_1, \alpha_2, \alpha_4$ is

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

[illegible]

$\therefore$ [illegible] and [illegible]

[illegible]

[illegible] are linearly independent.

[illegible] linearly independent subset of $\mathbf{R}^3$ [illegible] therefore it is a basis of $\mathbf{R}^3$.

Ex. [illegible]. [illegible] finite subset W of a vector space $V(F)$ is linearly dependent if and only if some element of W can be expressed as a linear combination of the others.

[illegible] be a finite subset of a vector space $V(F)$.

[illegible] suppose that W is linearly dependent. Then to show [illegible] can be expressed as a linear combination of [illegible]

[illegible] are linearly dependent, therefore [illegible] not all zero such that

$$[illegible] = 0 \qquad \ldots(1)$$

[illegible]

Then from (1), we have

$$[illegible]$$

$$= [illegible]$$

$$= [illegible]$$

[illegible] is a linear combination of the other vectors of the set W.

Conversely suppose that some vector of W is a linear combination of the others. Then to prove that W is linearly dependent.

Without loss of generality suppose [illegible] is a linear combination of [illegible]

[illegible], where $b_2, b_3, \ldots, b_n \in F$.

Then [illegible] $= 0$. $\qquad \ldots(2)$

[illegible] (2) among the vectors [illegible] [illegible], therefore the vectors [illegible] are linearly dependent.

Ex. [illegible]. Prove that any finite set S of vectors, not all the zero vector, contains a linearly independent subset T which spans the same space as S.

Sol. Let S be a finite set of vectors belonging to a vector space $V(F)$. Assume that no member of S is zero vector for if any member of S is zero vector, we can omit it from S without affecting the subspace spanned by S.

Let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$.

If S is linearly independent, then S itself is the required linearly independent subset T of S which spans the same subspace of V as S.

If S is linearly dependent, then there exists, a vector $\alpha_i \in S$ which can be expressed as a linear combination of the preceding vectors $\alpha_1, \alpha_2, \ldots, \alpha_{i-1}$.

If we omit this vector α_i from S, then the remaining subset S' of S containing $m-1$ vectors spans the same subspace of V as S.

If S' is linearly independent, then S' will be the required linearly independent subset of S which spans the same subspace of V as S. If S' is linearly dependent, then proceeding as above we shall get a new subset of S containing $m-2$ vectors which spans the same space as S. Continuing this process we shall, after a finite number of steps, obtain a linearly independent subset of S which spans the same space as S.

At the most it may happen that we shall be left with a subset of S which contains only one non-zero vector and which spans the same space as S. We know that a set containing a single non-zero vector is definitely linearly independent.

Hence any finite set S of vectors, not all the zero vectors, definitely contains a linearly independentsubs et T which spans the same space as S.

Ex. 16. *Let V be a vector space. Let W be a subspace of V generated by the vectors $\alpha_1, \ldots, \alpha_s$. Prove that W is spanned by a linearly independent subset of $\alpha_1, \ldots, \alpha_s$.*

Sol. W is a subspace of V generated by a finite set

$$S=\{\alpha_1, \ldots, \alpha_s\}.$$

To show that there exists a linearly independent subset T of S which also spans W.

Now for proof proceed as in solved example 15 above.

Ex. 17. *If W is a subspace of a finite dimensional vector space V, prove that any basis of W can be extended to form a basis of V.*

Sol. Let $V(F)$ be a finite dimensional vector space of dimension n. Let $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be a basis of V.

Scanned with CamScanner

Let W be a subspace of V. Then W itself is finite dimensional and dim $W \leqslant n$. Let dim $W=m$ and let $S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basis of W.

Then S is a linearly independent subset of V. To show that S can be extended to form a basis of V. For proof proceed as in theorem 1 of § 17.

Ex. 18. *If n vectors span a vector space V containing r linearly independent vectors, then show that $n \geqslant r$.*

Sol. Suppose a subset S of V containing n vectors spans V. Then there exists a linearly independent subset T of S which also spans V. This subset T of S will form a basis of V. Suppose the number of vectors in T is m. Then dim $V=m \leqslant n$.

Since dim $V=m$, therefore any subset of V containing more than m vectors will be linearly dependent.

Hence if S_1 is a linearly independent subset of V and S_1 contains r vectors, we must have

$$r \leqslant m$$

$$\Rightarrow \quad r \leqslant n. \qquad [\because \quad m \leqslant n]$$

Exercises

1. Tell with reason whether or not the vectors (2, 1, 0), (1, 1, 0), and (4, 2, 0) form a basis of $\mathbf{R}^3$. **(Meerut 1976)**
2. State whether the following statements are true or false :—
   (i) If a subset of an n-dimensional vector space V consists of n non-zero vectors, then it will be a basis of V. **(Meerut 1976)**
   (ii) If A and B are subspaces of a vector space, then dim $A <$ dim $B \Rightarrow A \subset B$. **(Meerut 1976)**
   (iii) If A and B are subspaces of a vector space, then $A \neq B \Rightarrow$ dim $A \neq$ dim B. **(Meerut 1976)**
   (iv) If M and N are finite dimensional subspaces with the same dimension, and if $M \subseteq N$, then $M=N$.
   (v) In an n-dimensional space any subset consisting of n linearly independent vectors will form a basis.

   Ans. (i) false; (ii) false; (iii) false; (iv) true; (v) true.
3. (i) Show that the vectors (2, 1, 4), (1, −1, 2), (3, 1, −2) form a basis for $\mathbf{R}^3$.
   (ii) Show that the vectors (0, 1, 1), (1, 0, 1) and (1, 1, 0) form a basis of $\mathbf{R}^3$. **(S.V.U. Tirupati 1993)**

Scanned with CamScanner

(iii) Determine whether or not the following vectors form a basis of $\mathbf{R}^3$:— (Meerut 1980)

$(1, 1, 2), (1, 2, 5), (5, 3, 4)$.

Ans. Do not form a basis of $\mathbf{R}^3$.

4. Show that the vectors $\beta_1=(1, 1, 0)$ and $\beta_2=(1, i, 1+i)$ are in the subspace W of $\mathbf{C}^3$ spanned b (1, 0, i) and $(1+i, 1, -1)$, and that β_1 and β_2 form a basis of W. **(Meerut 1974)**

5. Find three vectors in $\mathbf{R}^3$ which are linearly dependent, and are such that any two of them are linearly independent.

6. Prove that the space of all $m \times n$ matrices over the field F has dimension mn, by exhibiting a basis for this space.

7. If a vector space V is spanned by a finite set of m vectors, then show that any linearly independent set of vectors in V has at most m elements. **(Meerut 1973)**

8. In a vector space V over the field F, let $B=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ span V. Prove that the following two statements are equivalent :—

   (i) B is linearly independent.

   (ii) If $\alpha \in V$, then the expression $\alpha=\sum_{i=1}^{n} a_i\alpha_i$ with $a_i \in F$ is unique. **(Meerut 1979)**

9. If $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis of $V_3(\mathbf{R})$, show that $\{\alpha_1+\alpha_2, \alpha_2+\alpha_3, \alpha_3+\alpha_1\}$ is also a basis of $V_3(\mathbf{R})$. **(Nagarjuna 1991)**

§ 18. Homomorphism of vector spaces or Linear transformation.

Definition. *Let $U(F)$ and $V(F)$ be two vector spaces. Then a mapping*

$$f : U \to V$$

is called a homomorphism or a linear transformation of U **into** *V if*

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

and (*ii*) $f(a\alpha)=af(\alpha) \ \forall \ a \in F, \ \forall \ \alpha \in U.$

(Nagarjuna 1991; Poona 72)

The conditions (i) and (ii) can be combined into a single condition

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

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

Theorem. *If f is a homomorphism of $U(F)$ into $V(F)$, then*

(*i*) $f(\mathbf{0})=\mathbf{0}'$ *where $\mathbf{0}$ and $\mathbf{0}'$ are the zero vectors of U and V respectively.*

Scanned with CamScanner

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

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

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

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

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

$$\Rightarrow\quad 0'=f(0) \text{ by left cancellation law.}$$

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

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

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

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

§ 10. Isomorphism of Vector Spaces.

Definition. *Let $U(F)$ and $V(F)$ be two Vector spaces. Then a mapping*

$$f: U\to V$$

is called an isomorphism of U onto V if

(*i*) *f is one-one,*

(*ii*) *f is onto,*

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

Also then the two vector spaces U and V are said to be isomorphic and symbolically we write $U(F)\cong V(F)$. **(Kanpur 1980)**

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.

Isomorphism of finite dimensional vector spaces.

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

(Meerut 1988, 91, 92, 93P; S.V.U. Tirupati 90)

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

$$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$$

and

$$\{\beta_1, \beta_2, \ldots, \beta_n\}$$

be the bases of U and V respectively.

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

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

Scanned with CamScanner

Let $f: U \rightarrow V$ 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

$$\begin{aligned}
&f[a(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)+b(b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n)]\\
&=f[(aa_1+bb_1)\alpha_1+(aa_2+bb_2)\alpha_2+\ldots+(aa_n+bb_n)\alpha_n]\\
&=(aa_1+bb_1)\beta_1+(aa_2+bb_2)\beta_2+\ldots+(aa_n+bb_n)\beta_n\\
&=a(a_1\beta_1+a_2\beta_2+\ldots+a_n\beta_n)+b(b_1\beta_1+b_2\beta_2+\ldots+b_n\beta_n)\\
&=af(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)+bf(b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n).
\end{aligned}$$

$\therefore$ f is a linear transformation.

Hence f is an isomorphism of U onto V.

$$\therefore \quad U \cong V.$$

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 $a_1 f(\alpha_1)+a_2 f(\alpha_2)+\ldots+a_n f(\alpha_n)=\mathbf{0}'$ (zero vector of V)

$\Rightarrow f(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=\mathbf{0}'$

[$\because$ f is a linear transformation]

Scanned with CamScanner

$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=0$ [$\because$ f is one-one and $f(0)=0'$ where 0 is zero vector of U]

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

$\therefore$ S' is linearly independent.

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

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

Thus β is a linear combination of the vectors of S'.

Hence $V=L(S')$.

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

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

Theorem 2. *Every n-dimensional vector space $V(F)$ is isomorphic to $V_n(F)$* **(Nagarjuna 1980; Meerut 89, 93; Poona 72; Kanpur 80)**

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 $f: V(F) \to V_n(F)$ 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
$$\beta=b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n$$
be any two elements of V. We have

$$f(\alpha)=f(\beta)$$
$$\Rightarrow f(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=f(b_1\alpha_1+b_2\alpha_2+\ldots+b_n\alpha_n)$$
$$\Rightarrow (a_1, a_2, \ldots, a_n)=(b_1, b_2, \ldots, b_n)$$
$$\Rightarrow a_1=b_1, a_2=b_2, \ldots, a_n=b_n$$
$$\Rightarrow \alpha=\beta.$$
$\therefore$ f is one-one.

Scanned with CamScanner

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

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

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

$$f(a\alpha+b\beta)$$
$$=f[a(a_1\alpha_1+a_2\alpha_2+...+a_n\alpha_n)+b(b_1\alpha_1+b_2\alpha_2+...+b_n\alpha_n)]$$
$$=f[(aa_1+bb_1)\alpha_1+(aa_2+bb_2)\alpha_2+...+(aa_n+bb_n)\alpha_n]$$
$$=(aa_1+bb_1, aa_2+bb_2, ..., aa_n+bb_n)$$
$$=(aa_1, aa_2, ..., aa_n)+(bb_1, bb_2, ..., bb_n)$$
$$=a(a_1, a_2, ..., a_n)+b(b_1, b_2, ..., b_n)$$
$$=af(a_1\alpha_1+a_2\alpha_2+...+a_n\alpha_n)+bf(b_1\alpha_1+b_2\alpha_2+...+b_n\alpha_n)$$
$$=af(\alpha)+bf(\beta).$$

$\therefore$ f is a linear transformation.

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

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

Solved Examples

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

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

is a homomorphism of $V_3(F)$ onto $V_2(F)$. **[Kanpur 81]**

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

$$f(a\alpha+b\beta)=f[a(a_1, a_2, a_3)+b(b_1, b_2, b_3)]$$
$$=f[(aa_1+bb_1, aa_2+bb_2, aa_3+bb_3)]=(aa_1+bb_1, aa_2+bb_2)$$
$$=a(a_1, a_2)+b(b_1, b_2)=af(a_1, a_2, a_3)+bf(b_1, b_2, b_3)$$
$$=af(\alpha)+bf(\beta).$$

$\therefore$ f is a linear transformation.

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

Therefore f is a homomorphism of $V_3(F)$ onto $V_2(F)$.

Example 2. *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.*

Solution. T **is one-one.** Let

$$\alpha=a+ib, \beta=c+id$$

be any two members of $V(\mathbf{R})$. Then $a, b, c, d \in \mathbf{R}$.

Scanned with CamScanner

We have

$$T(\alpha)=T(\beta) \Rightarrow (a, b)=(c, d) \qquad [\because \; T(\alpha)=(a, b)]$$
$$\Rightarrow a=c,\, b=d \Rightarrow a+ib=c+id$$
$$\Rightarrow \alpha=\beta.$$

$\therefore$ T is one-one.

T is onto. Let (a, b) be an arbitrary member of V_2 (**R**). Then $\exists$ a vector $a+ib \in V$ (**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 (**R**) and k_1, k_2 be any two elements of the field **R**. Then

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

We have

$$T(k_1\alpha+k_2\beta)=(k_1a+k_2c,\, k_1b+k_2d), \text{ by definition of } T$$
$$=(k_1a,\, k_1b)+(k_2c,\, k_2d)=k_1(a, b)+k_2(c, d)$$
$$=k_1T(a+ib)+k_2T(c+id) \qquad [\text{by definition of } T]$$
$$=k_1T(\alpha)+k_2T(\beta).$$

Hence T is a linear transformation.

Hence T is an isomorphism.

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

(Poona 1970; Meerut 85)

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 $a_1 f(\alpha_1)+a_2 f(\alpha_2)+\ldots+a_n f(\alpha_n)=0$ (zero vector of V)

$$\Rightarrow f(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)=0 \qquad [\because \; f \text{ is linear transformation}]$$
$$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n=0 \qquad [\because \; f \text{ is one-one and } f(0)=0]$$
$$\Rightarrow a_1=0,\, a_2=0,\, \ldots,\, a_n=0 \text{ since } \alpha_1, \alpha_2, \ldots, \alpha_n \text{ are linearly independent.}$$

$\therefore$ 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'.

Scanned with CamScanner

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 4. *$V(F)$ 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 onto W.*

Solution. Proceed as in example 3.

Example 5. *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.*

(Meerut 1985)

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

Scanned with CamScanner

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

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

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

$\Rightarrow c_1-d_1=0,\ c_2-d_2=0, \ldots,\ c_n-d_n=0$ 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 6. *If V is finite dimensional and f is a homomorphism of V into itself which is not onto prove that there is some $\alpha \neq \mathbf{0}$ in V such that $f(\alpha)=\mathbf{0}$.* **(Meerut 1969)**

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

$$f(\beta)=f(\gamma)$$
$$\Rightarrow f(\beta)-f(\gamma)=\mathbf{0}$$
$$\Rightarrow f(\beta-\gamma)=\mathbf{0} \qquad [\because f \text{ is a linear transformation}]$$
$$\Rightarrow \beta-\gamma=\mathbf{0} \Rightarrow \beta=\gamma.$$

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

Example 7. *Define linear transformation of a vector space $V(F)$ into a vector space $W(F)$. Show that the mapping*

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

of $V_2(\mathbf{R})$ into itself is not a linear transformation.

Solution. Linear Transformation. Definition. Let $V(F)$ and $W(F)$ be two vector spaces over the same field F. A mapping

$$T : V \to W$$

is called a linear transformation of V into W if

$T(a\alpha+b\beta)=a\,T(\alpha)+b\,T(\beta)\ \forall\ a, b \in F$ and $\forall\ \alpha, \beta \in V$.

Now to show that the mapping

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

of $V_2(\mathbf{R})$ into itself is not a linear transformation.

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

Scanned with CamScanner

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

By the definition of the mapping T, we have

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

Also $T(\alpha)=T\,(1,2)=(1+2,\ 2+3)=(3,5)$
and $T(\beta)=T\,(1,3)=(1+2,\ 3+3)=(3,6)$.

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

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

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

Hence T is not a linear transformation of $V_2(\mathbf{R})$ into itself.

Example 8. *Let f be a linear transformation from a vector space U into a vector space V. If S is a subspace of U, prove that $f(S)$ will be a subspace of V.* **(Meerut 1974)**

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

$f: U\to V$ such that

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

Let S be a subspace of U. Then to prove that $f(S)$ is a subspace of V.

Let $a, b\in F$ and $f(\alpha), f(\beta)\in f(S)$ where $\alpha, \beta\in S$.

Since S is a subspace of U, therefore

$a, b\in F$ and $\alpha, \beta\in S \Rightarrow a\alpha+b\beta\in S$

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

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

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

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

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

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

Example 9. *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.* **(Nagarjuna 1990)**

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

$f: U\to V$ such that

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

Scanned with CamScanner

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

$$a_1 f(\alpha_1)+a_2 f(\alpha_2)+\ldots+a_r f(\alpha_r)=0,$$

where $a_1, a_2, \ldots, a_r \in F$

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

[$\because$ f is a linear transformation]

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

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

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

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

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

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

[$\because$ f is a linear transformation]

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

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.

Exercises

1. 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.
2. 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.
3. 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.
4. Prove that a finite dimensional vector space $V(\mathbf{R})$ with dimension $V=n$ is isomorphic to $\mathbf{R}^n$. (Meerut 1987)

Scanned with CamScanner

5. Let V be a finite dimensional vector space. If $f : V \to V$ is a one-one linear transformation, show that f is an isomorphism of V onto itself. **(Poona 1971)**
6. Give an example of a one-one linear transformation of an infinite dimensional vector space which is not an isomorphism.
7. If T be a linear operator on a finite dimensional vector space V, show that T is one-one if and only if T is onto. **(Meerut 1985)**

§ 20. Quotient Space. Let W be any subspace of a vector space $V(F)$. Let α be any element of V. Then the set

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

is called a **right coset** of W in V generated by α. Similarly the set

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

is called a **left coset** of W in V generated by α.

Obviously $W+\alpha$ and $\alpha+W$ are both subsets of V. Since addition in V is commutative, therefore we have $W+\alpha=\alpha+W$. Hence we shall call $W+\alpha$ as simply a coset of W in V generated by α.

The following results about cosets are both to be remembered:

(i) We have $0 \in V$ and $W+0=W$. *Therefore W itself is a coset of W in V.*

(ii) $\alpha \in W \Rightarrow W+\alpha=W$.

Proof. First we shall prove that $W+\alpha \subseteq W$.

Let $\gamma+\alpha$ be any arbitrary element of $W+\alpha$. Then $\gamma \in W$.

Now W is a subspace of V. Therefore

$$\gamma \in W, \alpha \in W \Rightarrow \gamma+\alpha \in W.$$

Thus every element of $W+\alpha$ is also an element of W. Hence

$$W+\alpha \subseteq W.$$

Now we shall prove that $W \subseteq W+\alpha$.

Let $\beta \in W$. Since W is a subspace, therefore

$$\alpha \in W \Rightarrow -\alpha \in W.$$

Consequently $\beta \in W, -\alpha \in W \Rightarrow \beta-\alpha \in W$. Now we can write

$$\beta=(\beta-\alpha)+\alpha \in W+\alpha \text{ since } \beta-\alpha \in W.$$

Thus $\beta \in W \Rightarrow \beta \in W+\alpha$. Therefore $W \subseteq W+\alpha$.

Hence $W=W+\alpha$.

(iii) *If $W+\alpha$ and $W+\beta$ are two cosets of W in V, then*

$$W+\alpha=W+\beta \Leftrightarrow \alpha-\beta \in W.$$

Proof. Since $0 \in W$, therefore $0+\alpha \in W+\alpha$. Thus

$$\alpha \in W+\alpha.$$

Now $\qquad W+\alpha=W+\beta \Rightarrow \alpha \in W+\beta$

$$\Rightarrow \alpha-\beta \in W+(\beta-\beta)$$
$$\Rightarrow \alpha-\beta \in W+\mathbf{0} \Rightarrow \alpha-\beta \in W.$$

Conversely,
$$\alpha-\beta \in W \Rightarrow W+(\alpha-\beta)=W$$
$$\Rightarrow W+[(\alpha-\beta)+\beta]=W+\beta$$
$$\Rightarrow W+\alpha=W+\beta.$$

Let V/W denote the set of all cosets of W in V *i.e.*, let
$$V/W=\{W+\alpha : \alpha \in V\}.$$

We have just seen that if $\alpha-\beta \in W$, then $W+\alpha=W+\beta$. Thus a coset of W in V can have more than one representation.

Now if $V(F)$ is a vector space, then we shall give a vector space structure to the set V/W over the same field F. For this we shall have to define addition in V/W *i.e.*, addition of cosets of W in V and multiplication of a coset by an element of F *i.e.*, scalar multiplication.

Theorem. *If W is any subspace of a vector space $V(F)$, then the set V/W of all cosets $W+\alpha$ where α is any arbitrary element of V, is a vector space over F for the addition and scalar multiplication compositions defined as follows :*
$$(W+\alpha)+(W+\beta)=W+(\alpha+\beta) \ \forall \ \alpha, \beta \in V$$
and
$$a\,(W+\alpha)=W+a\alpha \ ; \ a \in F, \alpha \in V.$$
(Meerut 1993; Kanpur 80; Nagarjuna 90)

Proof. We have $\alpha, \beta \in V \Rightarrow \alpha+\beta \in V.$

Also $\quad a \in F, \alpha \in V \Rightarrow a\alpha \in V.$

Therefore $W+(\alpha+\beta) \in V/W$ and also $W+a\alpha \in V/W$. Thus V/W is closed with respect to addition of cosets and scalar multiplication as defined above. Now first of all we shall show that these two compositions are well defined *i.e.*, are independent of the particular representative chosen to denote a coset.

Let $\quad W+\alpha=W+\alpha', \ \alpha, \alpha' \in V$

and $\quad W+\beta=W+\beta', \ \beta, \beta' \in V.$

We have $\quad W+\alpha=W+\alpha' \Rightarrow \alpha-\alpha' \in W$

and $\quad W+\beta=W+\beta' \Rightarrow \beta-\beta' \in W.$

Now W is a subspace, therefore
$$\alpha-\alpha' \in W, \beta-\beta' \in W \Rightarrow (\alpha-\alpha')+(\beta-\beta') \in W$$
$$\Rightarrow (\alpha+\beta)-(\alpha'+\beta') \in W$$
$$\Rightarrow W+(\alpha+\beta)=W+(\alpha'+\beta')$$
$$\Rightarrow (W+\alpha)+(W+\beta)=(W+\alpha')+(W+\beta').$$

Therefore addition in V/W is well defined.

Scanned with CamScanner

Again $a' \in F$, $\alpha-\alpha' \in W \Rightarrow a\,(\alpha-\alpha') \in W$

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

$\therefore$ scalar multiplication in V/W is also well defined.

Commutativity of addition. Let $W+\alpha$, $W+\beta$ be any two elements of V/W. Then

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

Associativity of addition. Let $W+\alpha$, $W+\beta$, $W+\gamma$ be any three elements of V/W. Then

$$(W+\alpha)+[(W+\beta)+(W+\gamma)]=(W+\alpha)+[W+(\beta+\gamma)]$$
$$=W+[\alpha+(\beta+\gamma)]$$
$$=W+[(\alpha+\beta)+\gamma]$$
$$=[W+(\alpha+\beta)]+(W+\gamma)$$
$$=[(W+\alpha)+(W+\beta)]+(W+\gamma).$$

Existence of additive identity. If **0** is the zero vector of V, then $W+0=W \in V/W$. If $W+\alpha$ is any element of V/W, then

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

$\therefore$ $W+0=W$ is the additive identity.

Existence of additive inverse. If $W+\alpha$ is any element of V/W, then $W+(-\alpha)=W-\alpha \in V/W$. Also we have

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

$\therefore$ $W-\alpha$ is the additive inverse of $W+\alpha$.

Thus V/W is an abelian group with respect to addition composition. Further we observe that if

$a, b \in F$ and $W+\alpha$, $W+\beta \in V/W$, then

1. $a\,[(W+\alpha)+(W+\beta)]=a\,[W+(\alpha+\beta)]$
$$=W+a(\alpha+\beta)=W+(a\alpha+a\beta)$$
$$=(W+a\alpha)+(W+a\beta)$$
$$=a\,(W+\alpha)+a\,(W+\beta).$$

2. $(a+b)\,(W+\alpha)=W+(a+b)\,\alpha$
$$=W+(a\alpha+b\alpha)$$
$$=(W+a\alpha)+(W+b\alpha)$$
$$=a\,(W+\alpha)+b\,(W+\alpha).$$

3. $(ab)\,(W+\alpha)=W+(ab)\,\alpha=W+a\,(b\alpha)$
$$=a\,(W+b\alpha)=a\,[b\,(W+\alpha)].$$

4. $1\,(W+\alpha)=W+1\alpha=W+\alpha.$

$\therefore$ V/W is a vector space over F for these two compositions. The vector space V/W is called the **Quotient Space** of V relative to W. The coset W is the zero vector of this vector space.

Scanned with CamScanner

Dimension of a Quotient Space. Theorem.

If W be a subspace of a finite dimensional vector space $V(F)$, then dim $V/W=$dim $V-$dim W.

(**Meerut 1990, 93; I.A.S 89; Allahabad 78, Nagarjuna 91; Andhra 92; Poona 72**)

Proof. Let m be the dimension of the subspace W of the vector space $V(F)$. Let

$$S=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$$

be a basis of W. Since S is a linearly independent subset of V, therefore it can be extended to form a basis of V. Let

$$S'=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l\}$$

be a basis of V. Then dim $V=m+l$.

$\therefore$ dim $V-$dim $W=(m+l)-m=l$.

So we should prove that dim $V/W=l$.

We claim that the set of l cosets

$$S_1=\{W+\beta_1, W+\beta_2, \ldots, W+\beta_l\}$$

is a basis of V/W.

First we show that S_1 is linearly independent. The zero vector of V/W is W.

Let $a_1(W+\beta_1)+a_2(W+\beta_2)+\ldots+a_l(W+\beta_l)=W$

$\Rightarrow (W+a_1\beta_1)+(W+a_2\beta_2)+\ldots+(W+a_l\beta_l)=W$

$\Rightarrow W+(a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l)=W+0$

$\Rightarrow a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l \in W$

$\Rightarrow a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l=b_1\alpha_1+b_2\alpha_2+\ldots+b_m\alpha_m$

[$\because$ any vector in W can be expressed as a linear combination of its basis vectors]

$\Rightarrow a_1\beta_1+a_2\beta_2+\ldots+a_l\beta_l-b_1\alpha_1-b_2\alpha_2-\ldots-b_m\alpha_m=0$

$\Rightarrow a_1=0, a_2=0, \ldots, a_l=0$ since the vectors $\beta_1, \beta_2, \ldots, \beta_l, \alpha_1, \alpha_2, \ldots, \alpha_m$ are linearly independent.

$\therefore$ The set S_1 is linearly independent.

Now to show that $L(S_1)=V/W$. Let $W+\alpha$ be any element of V/W. The vector $\alpha \in V$ can be expressed as

$$\alpha=c_1\alpha_1+c_2\alpha_2+\ldots+c_m\alpha_m+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l$$
$$=\gamma+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l \text{ where}$$
$$\gamma=c_1\alpha_1+c_2\alpha_2+\ldots+c_m\alpha_m \in W.$$

So $W+\alpha=W+(\gamma+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l)$

$=(W+\gamma)+d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l$

$=W+(d_1\beta_1+d_2\beta_2+\ldots+d_l\beta_l)$

[$\because \gamma \in W \Rightarrow W+\gamma=W$]

$=(W+d_1\beta_1)+(W+d_2\beta_2)+\ldots+(W+d_l\beta_l)$

$=d_1(W+\beta_1)+d_2(W+\beta_2)+\ldots+d_l(W+\beta_l)$

Scanned with CamScanner

$\therefore$ Thus any element $W+\alpha$ of V/W can be expressed as a linear combination of S_1.

$\therefore \quad V/W=L(S_1)$.

$\therefore \quad S_1$ is a basis of V/W.

$\therefore \quad \dim V/W=l$.

Hence the theorem.

§ 21. Direct sum of spaces.

Vector space as a direct sum of subspaces.

Definition. *Let $V(F)$ be a vector space and let $W_1, W_2, \ldots, W_m$ be subspaces of V. Then V is said to be the direct sum of $W_1, W_2, \ldots, W_m$ if every element $\alpha \in V$ can be written in one and only one way as $\alpha=\alpha_1+\alpha_2+\ldots+\alpha_m$ where*

$$\alpha_1 \in W_1, \alpha_2 \in W_2, \ldots, \alpha_m \in W_m. \qquad \textbf{(Nagarjuna 1978)}$$

If a vector space $V(F)$ is a direct sum of its two subspaces W_1 and W_2 then we should have not only $V=W_1+W_2$ but also that each vector of V can be uniquely expressed as sum of an element of W_1 and an element of W_2. Symbolically the direct sum is represented by the notation $V=W_1 \oplus W_2$.

Example. Let $V_2(F)$ be the vector space of all ordered pairs of F. Then $W_1=\{(a, 0) : a \in F\}$ and $W_2=\{(0, b) : b \in F\}$ are two subspaces of $V_2(F)$. Obviously any element $(x, y) \in V_2(F)$ can be uniquely expressed as sum of an element of W_1 and an element of W_2. The unique expression is $(x, y)=(x, 0)+(0, y)$. Thus $V_2(F)$ is the direct sum of W_1 and W_2. Also we observe that the only element common to both W_1 and W_2 is the zero vector $(0, 0)$.

Disjoint subspaces. Definition. *Two subspaces W_1 and W_2 of the vector space $V(F)$ are said to be disjoint if their intersection is the zero subspace i.e. if $W_1 \cap W_2=\{0\}$.*

Theorem. *The necessary and sufficient conditions for a vector space $V(F)$ to be a direct sum of its two subspaces W_1 and W_2 are that*

(i) $V=W_1+W_2$

and (ii) $W_1 \cap W_2=\{0\}$ *i.e., W_1 and W_2 are disjoint.*

(Meerut 1987, 90, 91, 92; Kanpur 81; Andhra 92)

Proof. The conditions are necessary.

Let V be direct sum of its two subspaces W_1 and W_2. Then each element of V is expressible uniquely as sum of an element of W_1 and an element of W_2. Therefore we have

$$V=W_1+W_2.$$

Let, if possible $0 \neq \alpha \in W_1 \cap W_2$. Then $\alpha \in W_1, \alpha \in W_2$. Also $\alpha \in V$ and we can write

$$\alpha=0+\alpha \text{ where } 0 \in W_1, \alpha \in W_2$$

and $\quad \alpha=\alpha+0$ where $\alpha \in W_1, 0 \in W_2$.

Thus $\alpha \in V$ can be expressed in at least two different ways as sum of an element of W_1 and an element of W_2. This contradicts the fact that V is direct sum of W_1 and W_2. Hence 0 is the only vector common to both W_1 and W_2 i.e. $W_1 \cap W_2=\{0\}$. Thus the conditions are necessary.

The conditions are sufficient.

Let $V=W_1+W_2$ and $W_1 \cap W_2=\{0\}$. Then to show that V is direct sum of W_1 and W_2.

$V=W_1+W_2 \Rightarrow$ that each element of V can be expressed as sum of an element of W_1 and an element of W_2. Now to show that this expression is unique.

Let, if possible,

$$\alpha=\alpha_1+\alpha_2, \ \alpha \in V, \alpha_1 \in W_1, \alpha_2 \in W_2,$$

and $\quad \alpha=\beta_1+\beta_2, \ \beta_1 \in W_1, \beta_2 \in W_2$.

Then to show that $\alpha_1=\beta_1$ and $\alpha_2=\beta_2$.

We have $\quad \alpha_1+\alpha_2=\beta_1+\beta_2$

$$\Rightarrow \alpha_1-\beta_1=\beta_2-\alpha_2.$$

Since W_1 is a subspace, therefore

$$\alpha_1 \in W_1, \beta_1 \in W_1 \Rightarrow \alpha_1-\beta_1 \in W_1.$$

Similarly $\beta_2-\alpha_2 \in W_2$.

$\therefore \quad \alpha_1-\beta_1=\beta_2-\alpha_2 \in W_1 \cap W_2$.

But 0 is the only vector which belongs to $W_1 \cap W_2$. Therefore $\alpha_1-\beta_1=0 \Rightarrow \alpha_1=\beta_1$. Also $\beta_2-\alpha_2=0 \Rightarrow \alpha_2=\beta_2$.

Thus each vector $\alpha \in V$ is uniquely expressible as sum of an element of W_1 and an element of W_2. Hence $V=W_1 \oplus W_2$.

Dimension of a Direct Sum. *If a finite dimensional vector space $V(F)$ is a direct sum of two subspaces W_1 and W_2, then*

$$\dim V=\dim W_1+\dim W_2. \qquad \textbf{(Nagarjuna 1980)}$$

Proof. Let $\dim W_1=m$ and $\dim W_2=l$. Also let the sets of vectors

$$S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$$

and $\quad S_2=\{\beta_1, \beta_2, \ldots, \beta_l\}$

be the bases of W_1 and W_2 respectively.

We have $\dim W_1+\dim W_2=m+l$.

In order to prove that dim $V=\dim W_1+\dim W_2$, we should therefore prove that dim $V=m+l$. We claim that the set

$$S=S_1 \cup S_2=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l\}$$

is a basis of V.

First we show that the set S is linearly independent. Let

$a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m+b_1\beta_1+b_2\beta_2+\ldots+b_l\beta_l=0$

$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m=-(b_1\beta_1+b_2\beta_2+\ldots+b_l\beta_l).$

Now $a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m \in W_1$

and $-(b_1\beta_1+b_2\beta_2+\ldots+b_l\beta_l) \in W_2$. Therefore

$a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m \in W_1 \cap W_2$

and $-(b_1\beta_1+b_2\beta_2+\ldots+b_l\beta_l) \in W_1 \cap W_2$.

But V is the direct sum of W_1 and W_2. Therefore 0 is the only vector belonging to $W_1 \cap W_2$. Then we have

$a_1\alpha_1+a_2\alpha_2+\ldots+a_m\alpha_m=0,\ b_1\beta_1+b_2\beta_2+\ldots+b_l\beta_l=0.$

Since both the sets $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ and $\{\beta_1, \beta_2, \ldots, \beta_l\}$ are linearly independent, therefore we have

$a_1=0,\ a_2=0, \ldots,\ a_m=0,\ b_1=0,\ b_2=0, \ldots,\ b_l=0.$

Therefore S is linearly independent.

Now we shall show that $L(S)=V$. Let α be any element of V. Then

α=an element of W_1+an element of W_2.

=a linear combination of S_1+a linear combination of S_2

=a linear combination of elements of S.

$\therefore\ L(S)=V$.

$\therefore\ S$ is a basis of V. Therefore dim $V=m+l$.

Hence the theorem.

A sort of converse of this theorem is true. It has been proved in the following theorem.

Theorem. *Let V be a finite dimensional vector space and let W_1, W_2 be subspaces of V such that $V=W_1+W_2$ and*

dim V=dim W_1+dim W_2. Then $V=W_1 \oplus W_2$.

Proof. Let dim $W_1=l$ and dim $W_2=m$. Then

dim $V=l+m$.

Let $S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_l\}$

be a basis of W_1 and

$S_2=\{\beta_1, \beta_2, \ldots, \beta_m\}$

be a basis of W_2. We shall show that $S_1 \cup S_2$ is a basis of V.

Let $\alpha \in V$. Since $V=W_1+W_2$, therefore we can write

$\alpha=\gamma+\delta$ where $\gamma \in W_1$, $\delta \in W_2$.

Now $\gamma \in W_1$ can be expressed as a linear combination of the elements of S_1 and $\delta \in W_2$ can be expressed as a linear combination of the elements of S_2. Therefore $\alpha \in V$ can be expressed as a linear combination of the elements of $S_1 \cup S_2$. Therefore $V=L(S_1 \cup S_2)$.

Since dim $V=l+m$ and $L(S_1 \cup S_2)=V$, therefore the number of distinct elements in $S_1 \cup S_2$ cannot be less than $l+m$. Thus $S_1 \cup S_2$ has $l+m$ distinct elements and therefore $S_1 \cup S_2$ is a basis of V. Therefore the set

$\{\alpha_1, \alpha_2, \ldots, \alpha_l, \beta_1, \beta_2, \ldots, \beta_m\}$

is linearly independent.

Now we shall show that $W_1 \cap W_2=\{0\}$.

Let $\alpha \in W_1 \cap W_2$. Then $\alpha \in W_1$, $\alpha \in W_2$.

Therefore $\alpha=a_1\alpha_1+a_2\alpha_2+\ldots+a_l\alpha_l$

and $\alpha=b_1\beta_1+b_2\beta_2+\ldots+b_m\beta_m$

for some a's and b's $\in F$.

$\therefore\ a_1\alpha_1+a_2\alpha_2+\ldots+a_l\alpha_l=b_1\beta_1+b_2\beta_2+\ldots+b_m\beta_m$

$\Rightarrow a_1\alpha_1+a_2\alpha_2+\ldots+a_l\alpha_l-b_1\beta_1-b_2\beta_2-\ldots-b_m\beta_m=0$

$\Rightarrow a_1=0,\ a_2=0, \ldots,\ a_l=0,\ b_1=0,\ b_2=0, \ldots,\ b_m=0$

$\Rightarrow \alpha=0$.

$\therefore\ W_1 \cap W_2=\{0\}$.

Complementary subspaces. Definition. *Let $V(F)$ be a vector space and W_1, W_2 be two subspaces of V. Then the subspace W_2 is called the complement of W_1 in V if V is the direct sum of W_1 and W_2.*

Existence of complementary subspaces. Theorem. *Corresponding to each subspace W_1 of a finite dimensional vector space $V(F)$, there exists a subspace W_2 such that V is the direct sum of W_1 and W_2.*

(I.A.S. 1972; Meerut 68)

Proof. Let dim $W_1=m$. Let the set

$S_1=\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$

be a basis of W_1.

Since S_1 is a linearly independent subset of V, therefore S_1 can be extended to form a basis of V. Let the set

$S=\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_l\}$

be a basis of V.

Let W_2 be the subspace of V generated by the set

$S_2=\{\beta_1, \beta_2, \ldots, \beta_l\}$.

We shall prove that V is the direct sum of W_1 and W_2. For this we shall prove that $V=W_1+W_2$ and $W_1 \cap W_2=\{0\}$.

Let α be any element of V. Then we can express

α=a linear combination of elements of S

=a linear combination of S_1+a linear combination of S_2

=an element of W_1+an element of W_2.

$\therefore\ V=W_1+W_2$.

Scanned with CamScanner

Suppose that also

$$\alpha=\beta_1+\ldots+\beta_k \text{ with } \beta_i \text{ in } W_i.$$

Then $\qquad \alpha_1+\ldots+\alpha_k=\beta_1+\ldots+\beta_k$

$\Rightarrow (\alpha_1-\beta_1)+\ldots+(\alpha_k-\beta_k)=0$ with $\alpha_i-\beta_i$ in W_i

as W_i is a subspace

$\Rightarrow \alpha_i-\beta_i=0 \qquad$ [$\because$ $W_1, \ldots, W_k$ are independent]

$\Rightarrow \alpha_i=\beta_i, \ i=1, \ldots, k.$

Therefore the α_i are uniquely determined by α.

(ii) $\Rightarrow$ (iii). Let $\alpha \in W_j \cap (W_1+\ldots+W_{j-1})$.

Then $\alpha \in W_j$ and $\alpha \in W_1+\ldots+W_{j-1}$.

Now $\alpha \in W_1+\ldots+W_{j-1}$ implies that there exist vectors $\alpha_1, \ldots, \alpha_{j-1}$ with α_i in W_i such that

$$\alpha=\alpha_1+\ldots+\alpha_{j-1}.$$

Also $\qquad \alpha \in W_j.$

Therefore we get two expressions for α as a sum of vectors, one in each W_i. These are

$$=\alpha_1+\ldots+\alpha_{j-1}+0+\ldots+0$$

in which the vector belonging to W_j is 0

and $\qquad \alpha$ [illegible] $\ldots+0+\alpha+\ldots+0$

in which the [illegible] belonging to W_j is α.

Since the expression for α is given to be unique, therefore we must have

$$\alpha_1=\ldots=\alpha_{j-1}=0=\alpha.$$

Thus $W_j \cap (W_1+\ldots+W_{j-1})=\{0\}$.

(iii) $\Rightarrow$ (i). Let $\alpha_1+\ldots+\alpha_k=0$

where $\alpha_i \in W_i, \ i=1, \ldots, k.$ ...(1)

Then we are to prove that each $\alpha_i=0$.

Suppose that for some i we have $\alpha_i \neq 0$.

Let j be the largest integer i between 1 and k such that $\alpha_i \neq 0$. Obviously j must be $\geqslant 2$ and at the most j can be equal to k. Then (1) reduces to

$$\alpha_1+\ldots+\alpha_j=0, \ \alpha_j \neq 0$$

$\Rightarrow \alpha_j=-\alpha_1-\ldots-\alpha_{j-1}$

$\Rightarrow \alpha_j \in W_1+\ldots+W_{j-1}$

[$\because$ $-\alpha_1-\ldots-\alpha_{j-1} \in W_1+\ldots+W_{j-1}$]

$\Rightarrow \alpha_j \in W_j \cap (W_1+\ldots+W_{j-1}) \qquad$ [$\because$ $\alpha_j \in W_j$]

$\Rightarrow \alpha_j=0.$

Thus we get a contradiction. Hence each $\alpha_i=0$.

Note. If any (and hence all) of three cond[illegible] theorem 1 hold for $W_1, \ldots, W_k$, then we shall say that W [illegible] ct sum of

$W_1, \ldots, W_k$ and we write

$$W=W_1 \oplus \ldots \oplus W_k.$$

Theorem 2. *Let $V(F)$ be a vector space. Let $W_1, \ldots, W_n$ be subspaces of V. Suppose that $V=W_1+\ldots+W_n$ and that*

$$W_i \cap (W_1+\ldots+W_{i-1}+W_{i+1}+\ldots+W_n)=\{0\}$$

for every $i=1, 2, \ldots, n$. Prove that V is the direct sum of $W_1, \ldots, W_n$.

Proof. In order to prove that V is the direct sum of $W_1, \ldots, W_n$, we should prove that each vector $\alpha \in V$ can be uniquely expressed as

$$\alpha=\alpha_1+\ldots+\alpha_n \text{ where } \alpha_i \in W_i, \ i=1, \ldots, n.$$

Since $V=W_1+\ldots+W_n$, therefore any vector α in V can be written as

$$\alpha=\alpha_1+\ldots+\alpha_n \text{ where } \alpha_i \in W_i. \qquad \ldots(1)$$

To show that $\alpha_1, \ldots, \alpha_n$ are unique.

Let $\alpha=\beta_1+\ldots+\beta_n$ where $\beta_i \in W_i$. ...(2)

From (1) and (2), we get

$$\alpha_1+\ldots+\alpha_n=\beta_1+\ldots+\beta_n$$

$$\Rightarrow (\alpha_1-\beta_1)+\ldots+(\alpha_{i-1}-\beta_{i-1})+(\alpha_i-\beta_i)$$
$$+(\alpha_{i+1}-\beta_{i+1})+\ldots+(\alpha_n-\beta_n)=0. \qquad \ldots(3)$$

Now each W_i is a subspace of V. Therefore $\alpha_i-\beta_i$ and also its additive inverse $\beta_i-\alpha_i \in W_i, \ i=1, \ldots, n.$

From (3), we get

$$(\alpha_i-\beta_i)=(\beta_1-\alpha_1)+\ldots+(\beta_{i-1}-\alpha_{i-1})$$
$$+(\beta_{i+1}-\alpha_{i+1})+\ldots+(\beta_n-\alpha_n). \qquad \ldots(4)$$

Now the vector on the right hand side of (4) and consequently the vector $\alpha_i-\beta_i$ is in $W_1+\ldots+W_{i-1}+W_{i+1}+\ldots+W_n$.

Also $\alpha_i-\beta_i \in W_i$.

$\therefore \quad \alpha_i-\beta_i \in W_i \cap (W_1+\ldots+W_{i-1}+W_{i+1}+\ldots+W_n)$.

But for every $i=1, \ldots, n$, it is given that

$$W_i \cap (W_1+\ldots+W_{i-1}+W_{i+1}+\ldots+W_n)=\{0\}.$$

Therefore $\alpha_i-\beta_i=0, \ i=1, \ldots, n$

$\Rightarrow \alpha_i=\beta_i, \ i=1, \ldots, n$

$\Rightarrow$ the expression (1) for α is unique.

Hence V is the direct sum of $W_1, \ldots, W_n$.

Theorem 3. *Let $V(F)$ be a finite dimensional vector space and let $W_1, \ldots, W_k$ be subspaces of V. Then the following two statements are equivalent.*

(*i*) *V is the direct sum of $W_1, \ldots, W_k$.*

(*ii*) *If B_i is a basis of $W_i, i=1, \ldots, k$, then the union*

Scanned with CamScanner

$B = \bigcup_{i=1}^{k} B_i$ *is also a basis for* V. (Meerut 1973, 75, 80, 83)

Proof. Let $B_i = \{\alpha_1^i, \alpha_2^i, ..., \alpha_{n_i}^i\}$ be a basis for W_i. Here $n_i = \dim W_i$ = number of vectors in B_i. Also let B be the union of the bases B_i.

(i) $\Rightarrow$ (ii). It is given that V is the direct sum of $W_1, ..., W_k$, therefore for any $\alpha \in V$, we can write

$$\alpha = \alpha_1 + ... + \alpha_k \text{ for } \alpha_i \in W_i,\ i = 1, ..., k.$$

Now α_i can be expressed as a linear combination of the vectors in B_i which is a basis of W_i. Therefore α can be expressed as a linear combination of the elements of $B = \bigcup_{i=1}^{k} B_i$. Therefore $L(B) = V$ i.e., B spans V.

Now to show that B is linearly independent. Let

$$\sum_{i=1}^{k} (a_1^i \alpha_1^i + a_2^i \alpha_2^i + ... + a_{n_i}^i \alpha_{n_i}^i) = 0. \quad ...(1)$$

Since V is the direct sum of $W_1, ..., W_k$, therefore $0 \in V$ can be uniquely expressed as a sum of vectors one in each W_i. This unique expression is

$$0 = 0 + ... + 0 \text{ where } 0 \in W_i,\ i = 1, ..., k.$$

Now $a_1^i \alpha_1^i + ... + a_{n_i}^i \alpha_{n_i}^i \in W_i$. Therefore from (1) which is an expression for $0 \in V$ as a sum of vectors one in each W_i, we get

$$a_1^i \alpha_1^i + ... + a_{n_i}^i \alpha_{n_i}^i = 0,\ i = 1, ..., k$$

$$\Rightarrow a_1^i = ... = a_{n_i}^i = 0 \text{ since } \{\alpha_1^i, ..., \alpha_{n_i}^i\}$$

is linearly independent being a basis for W_i.

Therefore $B = \bigcup_{i=1}^{k} B_i$ is linearly independent. Therefore B is a basis of V.

(ii) $\Rightarrow$ (i). It is given that $B = \bigcup_{i=1}^{k} B_i$ is a basis of V.

Therefore for any $\alpha \in V$, we can write

$$\alpha = \sum_{i=1}^{k} (a_1^i \alpha_1^i + a_2^i \alpha_2^i + a_{n_i}^i \alpha_{n_i}^i)$$

$$= \alpha_1 + \alpha_2 + ... + \alpha_k \quad ...(2)$$

where $\alpha_i = a_1^i \alpha_1^i + ... + a_{n_i}^i \alpha_{n_i}^i \in W_i$.

Thus each vector in V can be expressed as a sum of vectors one in each W_i.

Now V will be the direct sum of $W_1, ..., W_k$ if the expression (2) for α is unique. Let

$$\alpha = \beta_1 + \beta_2 + ... + \beta_k$$

where $\beta_i = b_1^i \alpha_1^i + ... + b_{n_i}^i \alpha_{n_i}^i \in W_i$. ...(3)

From (2) and (3), we get

$$\alpha_1 + ... + \alpha_k = \beta_1 + ... + \beta_k$$

$$\Rightarrow (\alpha_1 - \beta_1) + ... + (\alpha_i - \beta_i) + ... + (\alpha_k - \beta_k) = 0$$

$$\Rightarrow \sum_{i=1}^{k} [(a_1^i - b_1^i) \alpha_1^i + ... + (a_{n_i}^i - b_{n_i}^i) \alpha_{n_i}^i] = 0$$

$$\Rightarrow a_1^i - b_1^i = ... = a_{n_i}^i - b_{n_i}^i = 0,\ i = 1, ..., k$$

[$\because \bigcup_{i=1}^{k} B_i$ is linearly independent being a basis of V]

$$\Rightarrow a_1^i = b_1^i, ..., a_{n_i}^i = b_{n_i}^i,\ i = 1, ..., k$$

$$\Rightarrow \alpha_i = \beta_i,\ i = 1, ..., k$$

$\Rightarrow$ the expression (2) for α is unique.

Hence V is the direct sum of $W_1, ..., W_k$.

Note. While proving this theorem we have proved that *if a finite dimensional vector space* $V(F)$ *is the direct sum of its subspaces* $W_1, ..., W_k$, *then* $\dim V = \dim W_1 + ... + \dim W_k$.

§ 22. Co-ordinates. (Meerut 1983 P). Let $V(F)$ be a finite dimensional vector space. Let

$$B = \{\alpha_1, \alpha_2, ..., \alpha_n\}$$

be an ordered basis for V. By an ordered basis we mean that the vectors of B have been enumerated in some well-defined way *i.e.*, the vectors occupying the first, second, ..., n^{th} places in the set B are fixed.

Let $\alpha \in V$. Then there exists a unique n-tuple $(x_1, x_2, ..., x_n)$ of scalars such that

$$\alpha = x_1\alpha_1 + x_2\alpha_2 + ... + x_n\alpha_n = \sum_{i=1}^{n} x_i\alpha_i.$$

The n-tuple $(x_1, x_2, ..., x_n)$ is called the **n-tuple of co-ordinates of α relative to the ordered basis B.** The scalar x_i is called i^{th}

Scanned with CamScanner

coordinate of α relative to the ordered basis B. The $n \times 1$ matrix

$$X=\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

is called the **coordinate matrix** of α **relative to the ordered basis** B. We shall use the symbol

$$[\alpha]_B$$

for the coordinate matrix of the vector α relative to the ordered basis B.

It should be noted that for the same basis set B, the coordinates of the vector α are unique only with respect to a particular ordering of B. The basis set B can be ordered in several ways. The coordinates of α may change with a change in the ordering of B.

Solved Examples

Example 1. *Show that the set*

$$S=\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$$

is a basis of $\mathbf{R}^3$ (R) *where* **R** *is the field of real numbers. Hence find the coordinates of the vector* (a, b, c) *with respect to the above basis.* **(Meerut 1989)**

Solution. The dimension of the vector space $\mathbf{R}^3$ (R) is 3. If the set S is linearly independent, then S will form a basis of $\mathbf{R}^3$ (R) because S contains 3 vectors. Let x, y, z be scalars in **R** such that

$x(1, 0, 0)+y(1, 1, 0)+z(1, 1, 1)=\mathbf{0}=(0, 0, 0)$

$\Rightarrow (x+y+z, y+z, z)=(0, 0, 0)$

$\Rightarrow x+y+z=0, y+z=0, z=0$

$\Rightarrow x=0, y=0, z=0$

$\Rightarrow$ the set S is linearly independent.

$\therefore$ S is a basis of $\mathbf{R}^3$ (R).

Now to find the coordinates of (a, b, c) with respect to the ordered basis S. Let p, q, r be scalars in **R** such that

$(a, b, c)=p(1, 0, 0)+q(1, 1, 0)+r(1, 1, 1)$

$\Rightarrow (a, b, c)=(p+q+r, q+r, r)$

$\Rightarrow p+q+r=a, q+r=b, r=c$

$\Rightarrow r=c, q=b-c, p=a-b.$

Hence the coordinates of the vector (a, b, c) are (p, q, r) *i.e.*, $(a-b, b-c, c)$.

Ex. 2. *Find the coordinates of the vector* $(2, 1, -6)$ *of* $\mathbf{R}^3$ *relative to the basis* $\alpha_1=(1, 1, 2)$, $\alpha_2=(3, -1, 0)$, $\alpha_3=(2, 0, -1)$.

Sol. To find the coordinates of the vector $(2, 1, -6)$ relative to the ordered basis $\{\alpha_1, \alpha_2, \alpha_3\}$, we shall express the vector $(2, 1, -6)$

as a linear combination of the vectors $\alpha_1, \alpha_2, \alpha_3$. Let p, q, r be scalars in **R** such that

$(2, 1, -6)=p\alpha_1+q\alpha_2+r\alpha_3$

$\Rightarrow (2, 1, -6)=p(1, 1, 2)+q(3, -1, 0)+r(2, 0, -1)$

$\Rightarrow (2, 1, -6)=(p+3q+2r, p-q, 2p-r).$

$$\therefore \quad \left.\begin{aligned} p+3q+2r&=2, \\ p-q&=1, \\ \text{and} \quad 2p-r&=-6. \end{aligned}\right\} \qquad \ldots(1)$$

Solving the equations (1), we get

$p=-7/8$, $q=-15/8$ and $r=17/4$.

Hence the required coordinates of the vector $(2, 1, -6)$ relative to the ordered basis $\{\alpha_1, \alpha_2, \alpha_3\}$ are (p, q, r) *i.e.*, $(-7/8, -15/8, 17/4)$.

Example 3. *Construct three subspaces* W_1, W_2, W_3 *of a vector space* V *so that* $V=W_1 \oplus W_2=W_1 \oplus W_3$ *but* $W_2 \neq W_3$. **(Kanpur 1980)**

Solution. Take the vector space $V=\mathbf{R}^2$.

Obviously $W_1=\{(a, 0) : a \in \mathbf{R}\}$,

$W_2=\{(0, a) : a \in \mathbf{R}\}$,

and $W_3=\{(a, a) : a \in \mathbf{R}\}$

are three subspaces of $\mathbf{R}^2$.

We have $V=W_1+W_2$ and $W_1 \cap W_2=\{(0, 0)\}$.

$\therefore \quad V=W_1 \oplus W_2$.

Also it can be easily shown that

$V=W_1+W_3$ and $W_1 \cap W_3=\{(0, 0)\}$.

$\therefore \quad V=W_1 \oplus W_3$.

Thus $V=W_1 \oplus W_2=W_1 \oplus W_3$ but $W_2 \neq W_3$.

Exercises

1. Let W_1 and W_2 be two subspaces of a finite dimensional vector space V. If
   dim V=dim W_1+dim W_2 and $W_1 \cap W_2=\{\mathbf{0}\}$, prove that
   $$V=W_1 \oplus W_2.$$
2. Show that the set $S=\{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ is a basis of $\mathbf{C}^3$ (C) where C is the field of complex numbers. Hence find the coordinates of the vector $(3+4i, 6i, 3+7i)$ in $\mathbf{C}^3$ with respect to the above basis.
   Ans. $(3-2i, -3-i, 3+7i)$.
3. Let $B=\{\alpha_1, \alpha_2, \alpha_3\}$ be an ordered basis for $\mathbf{R}^3$, where $\alpha_1=(1, 0, -1)$, $\alpha_2=(1, 1, 1)$, $\alpha_3=(1, 0, 0)$.
   Obtain the coordinates of the vector (a, b, c) in the ordered

Scanned with CamScanner

basis B. (Meerut 1973, 77, 85)

Ans. $(b-c, b, a-2b+c)$.

4. Let V be the vector space of all polynomial functions of degree less than or equal to two from the field of real numbers **R** into itself. For a fixed $t \in \mathbf{R}$, let

$$g_1(x)=1,\ g_2(x)=x+t,\ g_3(x)=(x+t)^2.$$

Prove that $\{g_1, g_2, g_3\}$ is a basis for V and obtain the coordinates of $c_0+c_1x+c_2x^2$ in this ordered basis. (Meerut 1973)

Ans. $(c_0-c_1t+c_2t^2,\ c_1-2c_2t,\ c_2)$.

5. Let V be a finite-dimensional vector space and let $W_1, \ldots, W_k$ be subspaces of V such that

$V=W_1+\ldots+W_k$ and $\dim V=\dim W_1+\ldots+\dim W_k$.

Prove that $V=W_1 \oplus \ldots \oplus W_k$. (Meerut 1976)

2
Linear Transformations

§ 1. Linear transformations or Vector space homomorphism.

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.

(Meerut 1978, 79; Allahabad 77; Nagarjuna 80, 91)

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 for 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. (Meerut 1983)

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

Example 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

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

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

Scanned with CamScanner

2

Linear Transformations

§ 1. Linear transformations *or* Vector space homomorphism.

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.

(Meerut 1978, 79; Allahabad 77; Nagarjuna 80, 91)

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 for 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. **(Meerut 1983)**

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

Example 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

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

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

Example 2. *Let $V(F)$ be the vector space of all $m \times n$ matrices over the field F. Let P be a fixed $m \times m$ matrix over F, and let Q be a fixed $n \times n$ matrix over F. The correspondence T from V into V defined by* $T(A)=PAQ \; \forall \; A \in V$
is a linear operator on V.

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

$$T(aA+bB)=P(aA+bB)Q \quad \text{[by def. of } T]$$
$$=(aPA+bPB)Q=a\,PAQ+b\,PBQ=aT(A)+bT(B)$$

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

Example 3. *Let $V(F)$ be the vector space of all polynomials over the field F. Let $f(x)=a_0+a_1+a_2x^2+\ldots+a_nx^n \in V$ be a polynomial of degree n in the indeterminate x. Let us define*

$$Df(x)=a_1+2a_2x+\ldots+na_nx^{n-1} \text{ if } n>1$$
and $$Df(x)=0 \text{ if } f(x) \text{ is a constant polynomial.}$$

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

$$D[a\,f(x)+bg(x)]=a\,Df(x)+bDg(x).$$

$\therefore$ 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.

Example 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

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

$\therefore$ 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.

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

$$(-T)(a\alpha+b\beta)=-[T(a\alpha+b\beta)] \quad \text{[by def. of } -T]$$
$$=-[a\,T(\alpha)+bT(\beta)] \quad [\because T \text{ is a linear transformation}]$$
$$=a[-T(\alpha)]+b[-T(\beta)]=a[(-T)\alpha]+b[(-T)\beta].$$

$\therefore$ $-T$ 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.* **(Meerut 1974, 76, 79)**

(ii) $T(-\alpha)=-T(\alpha) \;\forall\; \alpha \in U.$ **(Meerut 1979)**

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

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

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

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

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

$=T(\alpha)+T(0)$ [$\because$ T is a linear transformation]

Now in the vector space V, we have

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

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

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

(ii) We have $T[\alpha+(-\alpha)]=T(\alpha)+T(-\alpha)$

[$\because$ T is a linear transformation]

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

Thus in V, we have

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

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

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

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

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

Then $T(a_1\alpha_1+a_2\alpha_2+\ldots+a_n\alpha_n)$

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

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

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

(Marathwada 1971)

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

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

Theorem 1. *If $U(F)$ and $V(F)$ are two vector spaces and T is a linear transformation from U into V, then range of T is a subspace of V.* **(Meerut 1980)**

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

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

$=T\,(a\alpha_1+b\alpha_2)$ [$\because$ T 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 a linear transformation. Definition.

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

$$N(T)=\{\alpha \in U : T(\alpha)=0 \in V\}.$$ **(Marathwada 1971)**

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. **(Nagarjuna 1991; Allahabad 75)**

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.*

(Meerut 1980, 90; Allahabad 75; Nagarjuna 91; Andhra 92)

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

Since $T(0)=0 \in V$, therefore at least $0 \in N(T)$

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

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

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

$T\,(a\alpha_1+b\alpha_2)=a\ T(\alpha_1)+b\ T(\alpha_2)$

[$\because$ T is a linear transformation]

$=a0+b0=0+0=0 \in V$.

Scanned by TapScanner

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

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.

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

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

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

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

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. Definition.

(S.V. 1992; Meerut 75; Kanpur 81; Allahabad 79)

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

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*

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

(Meerut 1983P, 87, 93; Andhra 92; Tirupati 90; I.A.S. 85; Madras 83; Madurai 85; Nagarjuna 90; Kanpur 81; Allahabad 79)

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

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

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

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

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

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 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,

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

Solved Examples

Ex. 1. *Show that the mapping $T : V_3(\mathbf{R}) \to V_2(\mathbf{R})$ defined as*
$$T(a_1, a_2, a_3)=(3a_1-2a_2+a_3,\ a_1-3a_2-2a_3)$$
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 $\quad T(\alpha)=T(a_1, a_2, a_3)=(3a_1-2a_2+a_3,\ a_1-3a_2-2a_3)$

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

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

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

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

Ex 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. **(Nagarjuna 1990; Tirupati 90)**

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

$$T(a\alpha+b\beta)=T[a(a_1, b_1)+b(a_2, b_2)]$$
$$=T(aa_1+ba_2,\ ab_1+bb_2)$$
$$=(aa_1+ba_2+ab_1+bb_2,\ aa_1+ba_2-ab_1-bb_2,\ ab_1+bb_2)$$
$$=(a[a_1+b_1]+b[a_2+b_2],\ a[a_1-b_1]+b[a_2-b_2],\ ab_1+bb_2)$$
$$=a(a_1+b_1,\ a_1-b_1,\ b_1)+b(a_2+b_2,\ a_2-b_2,\ b_2)$$
$$=aT(\alpha)+bT(\beta)$$

$\therefore$ 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 $\quad T(1, 0)=(1+0,\ 1-0,\ 0)=(1, 1, 0)$

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

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

$$x(1, 1, 0)+y(1, -1, 1)=(0, 0, 0)$$
$$\Rightarrow (x+y,\ x-y,\ y)=(0, 0, 0)$$
$$\Rightarrow x+y=0,\ x-y=0,\ y=0 \Rightarrow x=0,\ y=0.$$

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

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

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

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

Otherwise, $\quad (a, b) \in$ null space of T
$$\Rightarrow T(a, b)=(0, 0, 0)$$
$$\Rightarrow (a+b,\ a-b,\ b)=(0, 0, 0)$$
$$\Rightarrow a+b=0,\ a-b=0,\ b=0$$
$$\Rightarrow a=0,\ b=0.$$

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

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

Example 3. *Let F be field of complex numbers and let T be the function from F^3 into F^3 defined by*
$$T(x_1, x_2, x_3)=(x_1-x_2+2x_3,\ 2x_1+x_2-x_3,\ -x_1-2x_2).$$
Verify that T is a linear transformation. Describe the null space of T. **(Meerut 1979, 85)**

Solution. Let $\alpha=(x_1, x_2, x_3)$, $\beta=(y_1, y_2, y_3) \in F^3$. Then

$T(\alpha)=(x_1-x_2+2x_3,\ 2x_1+x_2-x_3,\ -x_1-2x_2)$ and

$T(\beta)=(y_1-y_2+2y_3,\ 2y_1+y_2-y_3,\ -y_1-2y_2)$.

Also let $a, b \in F$. Then $a\alpha+b\beta \in F^3$ and

$$a\alpha+b\beta=a(x_1, x_2, x_3)+b(y_1, y_2, y_3)$$
$$=(ax_1+by_1,\ ax_2+by_2,\ ax_3+by_3).$$

Now, by the definition of T, we have

$$T(a\alpha+b\beta)=([ax_1+by_1]-[ax_2+by_2]+2[ax_3+by_3],\ 2[ax_1+by_1]+ax_2+by_2-[ax_3+by_3],\ -[ax_1+by_1]-2[ax_2+by_2])$$
$$=(a[x_1-x_2+2x_3]+b[y_1-y_2+2y_3],\ a[2x_1+x_2-x_3]+b[2y_1+y_2-y_3],\ a[-x_1-2x_2]+b[-y_1-2y_2])$$
$$=a(x_1-x_2+2x_3,\ 2x_1+x_2-x_3,\ -x_1-2x_2)+b(y_1-y_2+2y_3,\ 2y_1+y_2-y_3,\ -y_1-2y_2)$$
$$=a\,T(\alpha)+b\,T(\beta).$$

$\therefore$ T is a linear transformation from F^3 into F^3.

Now $(x_1, x_2, x_3) \in$ null space of T
$$\Leftrightarrow T(x_1, x_2, x_3)=(0, 0, 0)$$
$$\Leftrightarrow (x_1-x_2+2x_3,\ 2x_1+x_2-x_3,\ -x_1-2x_2)=(0, 0, 0)$$

Scanned by TapScanner

$$\Leftrightarrow \left.\begin{array}{l} x_1-x_2+2x_3=0, \\ 2x_1+x_2-x_3=0, \\ -x_1-2x_2+0x_3=0. \end{array}\right\} \qquad \ldots(1)$$

$\therefore$ the null space of T is the solution space of the system of linear homogeneous equations (1). Let A be the coefficient matrix of the equations (1). Then

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

$$\sim\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -5 \\ 0 & -3 & 2 \end{bmatrix}$$ performing the elementary row operations $R_2 \to R_2-2R_1$, $R_3 \to R_3+R_1$

$$\sim\begin{bmatrix} 1 & -1 & 2 \\ 0 & 3 & -5 \\ 0 & 0 & -3 \end{bmatrix}$$ by $R_3 \to R_3+R_2$.

This last matrix is in Echelon form. Its rank = the number of non-zero rows = 3. Therefore rank $A=3$ = the number of unknowns in the equations (1). Hence the equations (1) have no linearly independent solutions. Therefore $x_1=0$, $x_2=0$, $x_3=0$ is the only solution of the equations (1). Thus $(0, 0, 0)$ is the only vector which belongs to the null space of T. Hence the null space of T is the zero subspace of F^3.

Example 4. *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. **(Meerut 1982)**

Solution. If $A \in V$, then $T(A)=AB-BA \in V$ because $AB-BA$ is also an $n \times n$ matrix over the field F. Thus T is a function from V into V.

Let $A_1, A_2 \in V$ and $a, b \in F$. Then $aA_1+bA_2 \in V$ and

$$T(aA_1+bA_2)=(aA_1+bA_2)B-B(aA_1+bA_2)$$
$$=aA_1B+bA_2B-aBA_1-bBA_2=a(A_1B-BA_1)+b(A_2B-BA_2)$$
$$=aT(A_1)+bT(A_2).$$

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

Example 5. *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 $\alpha=(a_1, b_1)$, $\beta=(a_2, b_2) \in V_2(\mathbf{R})$ 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 $V_2(\mathbf{R})$ into $V_2(\mathbf{R})$.

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

We have $T(1, 0)=(0, 0)$ and $T(0, 1)=(1, 0)$.

Thus the range of T is the subspace of $V_2(\mathbf{R})$ 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 $V_2(\mathbf{R})$ spanned by the vector $(1, 0)$. Thus 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 $\dim V_2(\mathbf{R})=2$ which is even.

Example 6. *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. **(Marathwada 1971)**

Solution. If $\alpha \in U$, then $T_1(\alpha)$ and $T_2(\alpha) \in V$. Therefore $xT_1(\alpha)+yT_2(\alpha) \in V$. Thus T as defined above is a mapping from U into V. Let $\alpha, \beta \in U$ and $a, b \in F$. Then

$T(a\alpha+b\beta)=xT_1(a\alpha+b\beta)+yT_2(a\alpha+b\beta)$ [by def. of T]

$=x[aT_1(\alpha)+bT_1(\beta)]+y[aT_2(\alpha)+bT_2(\beta)]$

[$\because$ T_1 and T_2 are linear transformations]

$=a[xT_1(\alpha)+yT_2(\alpha)]+b[xT_1(\beta)+yT_2(\beta)]$

$=aT(\alpha)+bT(\beta)$ [by def. of T]

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

Scanned by TapScanner

Example 7. *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*) $T[T(\alpha)]=\mathbf{0} \Rightarrow T(\alpha)=\mathbf{0}$. **(Meerut 1979, 85)**

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

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

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

$\Rightarrow T(\alpha)=\mathbf{0}$ 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 8. *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 $T(2, -3, 5)$. **(I.A.S. 1985)**

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

$=5T(\alpha_1)-8T(\alpha_2)+5T(\alpha_3)$ $\quad [\because T$ is a linear transformation$]$

$=5(1, 0)-8\,(2, -1)+5\,(4, 3)$

$=(5, 0)-(16, -8)+(20, 15)$

$=(9, 23)$.

Exercises

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 $T : \mathbf{R}^2 \to \mathbf{R}^3$ 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.

Ans. 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 $R(T)$.

3. Let F be a subfield of the 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. **(Meerut 1983)**

4. Which of the following functions T from $\mathbf{R}^2$ into $\mathbf{R}^2$ are linear transformations ?

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

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

Ans. (a) T is not a linear transformation. (b) T is a linear transformation. (c) T is a linear transformation.

5. 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.

6. 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. **(Meerut 1981)**

Ans. (i) $\{(1, 0, 1), (2, 1, 1)\}$ is a basis of $R(T)$ and dim. $R(T)=2$; $\{(3, -1, 1)\}$ is a basis of $N(T)$ and dim. $N(T)=1$.

7. 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. **(Meerut 1992)**

§ 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

Scanned by TapScanner

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.

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.

(Andhra 1992; Meerut 78, 91; Madras 81; Kanpur 69)

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

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

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

$(T_1+T_2)(a\alpha+b\beta)=T_1(a\alpha+b\beta)+T_2(a\alpha+b\beta)$ [by (1)]

$=[aT_1(\alpha)+bT_1(\beta)]+[aT_2(\alpha)+bT_2(\beta)]$ [$\because$ T_1 and T_2 are linear transformations]

$=a[T_1(\alpha)+T_2(\alpha)]+b[T_1(\beta)+T_2(\beta)]$ [$\because$ V is a vector space]

$=a(T_1+T_2)(\alpha)+b(T_1+T_2)(\beta)$ [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:

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

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

$(cT)(a\alpha+b\beta)=cT(a\alpha+b\beta)$ [by (2)]

$=c[aT(\alpha)+bT(\beta)]$ [$\because$ T 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 $T \in L(U, V)$ 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

$[T_1+(T_2+T_3)](\alpha)=T_1(\alpha)+(T_2+T_3)(\alpha)$ [by (1) i.e., by def. of addition in $L(U, V)$]

$=T_1(\alpha)+[T_2(\alpha)+T_3(\alpha)]$ [by (1)]

$=[T_1(\alpha)+T_2(\alpha)]+T_3(\alpha)$ [$\because$ addition in V is associative]

$=(T_1+T_2)(\alpha)+T_3(\alpha)$ [by (1)]

$=[(T_1+T_2)+T_3](\alpha)$ [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

$(T_1+T_2)(\alpha)=T_1(\alpha)+T_2(\alpha)$ [by (1)]

$=T_2(\alpha)+T_1(\alpha)$ [$\because$ addition in V is commutative]

$=(T_2+T_1)(\alpha)$ [by (1)]

$\therefore$ $T_1+T_2=T_2+T_1$ [by def. of equality of two functions]

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

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

$(\hat{0}+T)(\alpha)=\hat{0}(\alpha)+T(\alpha)$ [by (1)]

$=0+T(\alpha)$ [by def. of $\hat{0}$]

$=T(\alpha)$ [0 being additive identity in V]

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

$\therefore$ $\hat{0}$ 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

$(-T+T)(\alpha)=(-T)(\alpha)+T(\alpha)$ [by def. of addition in $L(U, V)$]

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

$=0 \in V$ [by def. of $-T$]

$=\hat{0}(\alpha)$ [by def. of $\hat{0}$]

$\therefore \quad -T+T=\hat{0}$ 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

$[c(T_1+T_2)](\alpha)=c[(T_1+T_2)(\alpha)]$

[by (2) *i.e.*, by def. of scalar multiplication in $L(U, V)$]

$=c[T_1(\alpha)+T_2(\alpha)]$

$=cT_1(\alpha)+cT_2(\alpha)$ [by (1)]

[$\because$ $c \in F$ and $T_1(\alpha), T_2(\alpha) \in V$ which is a vector space]

$=(cT_1)(\alpha)+(cT_2)(\alpha)$

$=(cT_1+cT_2)(\alpha)$ [by (2)]

$\therefore \quad c(T_1+T_2)=cT_1+cT_2$. [by (1)]

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

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

$=aT(\alpha)+bT(\alpha)$ [$\because$ V is a vector space]

$=(aT)(\alpha)+(bT)(\alpha)$ [by (2)]

$=(aT+bT)(\alpha)$ [by (2)]

$\therefore \quad (a+b)T=aT+bT$. [by (1)]

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

$[(ab)T](\alpha)=(ab)T(\alpha)$

$=a[bT(\alpha)]$ [by (2)]

$=a[(bT)(\alpha)]$ [$\because$ V is a vector space]

$=[a(bT)](\alpha)$ [by (2)]

$\therefore \quad (ab)T=a(bT)$. [by (2)]

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

$(1T)(\alpha)=1T(\alpha)$

$=T(\alpha)$ [by (2)]

$\therefore \quad 1T=T$. [$\because$ V is a vector space]

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 fied F and let $\beta_1, \ldots\ldots, \beta_n$ be any vectors in V. Then there exists a unique linear transformation T from U into V such that*

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

(Meerut 1980, 81, 93, Nagarjuna 91)

Proof. Existence of T.

Let $\alpha \in U$.

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

Now to show that T is a linear transformation.

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

$$\alpha=x_1\alpha_1+\ldots+x_n\alpha_n$$

and

$$\beta=y_1\alpha_1+\ldots+y_n\alpha_n.$$

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

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

$=(ax_1+by_1)\beta_1+\ldots+(ax_n+by_n)\beta_n$ [by def. of T]

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

$=a\,T(\alpha)+b\,T(\beta)$ [by def. of T]

$\therefore$ 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

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

$=x_1T'(\alpha_1)+\ldots+x_nT'(\alpha_n)$ [$\because$ T' is a linear transformation]

$=x_1\beta_1+\ldots+x_n\beta_n$ [by def. of T']

$=T(\alpha)$. [by def. of T]

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.*

(Meerut 1970, 76, 80, 93P; Madras 81; Nagarjuna 77; Andhra 81)

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

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

$\beta_1, 0,\ldots,0$ are vectors in V.

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

$$T_{pq}(\alpha_i)=\begin{cases}0, & \text{if } i\neq q\\ \beta_p, & \text{if } i=q\end{cases} \qquad \ldots(1)$$

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

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, $1\leqslant i\leqslant n$,

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

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

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

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

$$=\sum_{p=1}^{m}a_{pi}\,\beta_p \qquad \text{[On summing with respect to } q.$$

Remember that $\delta_{iq}=1$ when $q=i$ and $\delta_{iq}=0$ when $q\neq i$]

$=T(\alpha_i)$. [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

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

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

$$\Rightarrow \sum_{p=1}^{m} \sum_{q=1}^{n} b_{pq} T_{pq} (\alpha_i) = 0 \in V$$

[$\because$ $\hat{0}$ is zero transformation]

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

$$\Rightarrow \sum_{p=1}^{m} b_{pi} \beta_p = 0$$

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

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

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

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

$\Rightarrow B_1$ 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 $L(U, V)$ 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 $U \neq \{0\}$ and $V \neq \{0\}$, then $n \geqslant 1$ and $m \geqslant 1$. Therefore $L(U, V)$ does not just consist of the element $\hat{0}$, because dimension of $L(U, V)$ is $mn \geqslant 1$.

§ 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 fro[illegible] W.

(Mee[illegible] Nagarju[illegible]

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

$(ST)(a\alpha + b\beta) = S[T(a\alpha + b\beta)]$ [by def. of product of two functions]

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

$= aS[T(\alpha)] + bS[T(\beta)]$ [$\because$ S 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 $TS \neq ST$ as is obvious from the following examples.

Example 1. *Let T_1 and T_2 be linear operators on $\mathbf{R}^2$ defined as follows :*

$$T_1(x_1, x_2) = (x_2, x_1)$$

and $$T_2(x_1, x_2) = (x_1, 0).$$

Show that $T_1 T_2 \neq T_2 T_1$.

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 2. *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*

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

and $$T(f(x))=\int_0^x f(x)\,dx \quad ...(1)$$

for every $f(x)\in V$. ...(2)

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

Solution. Let $f(x)=a_0+a_1x+a_2x^2+...\in V$.

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

$$=D\left[\int_0^x f(x)\,dx\right]=D\left[\int_0^x (a_0+a_1x+a_2x^2+...)\,dx\right]$$

$$=D\left[a_0x+\frac{a_1}{2}x^2+\frac{a_2}{3}x^3+...\right]_0^x=\frac{d}{dx}\left[a_0x+\frac{a_1}{2}x^2+...\right]$$

$$=a_0+a_1x+a_2x^2+...=f(x)=I[f(x)].$$

Thus we have

$(DT)[f(x)]=I[f(x)]\ \forall\ f(x)\in V$. **Therefore** $DT=I$.

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

$$=T\left[\frac{d}{dx}(a_0+a_1x+a_2x^2+...)\right]$$

$$=T(a_1+2a_2x+...)$$

$$=\int_0^x (a_1+2a_2x+...)\,dx$$

$$=\left[a_1x+a_2x^2+...\right]_0^x$$

$$=a_1x+a_2x^2+...$$

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

Thus $\exists\ f(x)\in V$ such that

$$(TD)[f(x)]\neq I[f(x)].$$

$$\therefore\ TD\neq I.$$

Hence $$TD\neq DT,$$

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

Example 3. *Let* $V(\mathbf{R})$ *be the vector space of all polynomials in* x *with coefficients in the field* $\mathbf{R}$. *Let* D *and* T *be two linear transformations on* V *defined as*

$$D[f(x)]=\frac{d}{dx}f(x)\ \forall\ f(x)\in V$$

and $$T[f(x)]=x f(x)\ \forall\ f(x)\in V.$$

Then show that $DT\neq TD$.

Solution. We have

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

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

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

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

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

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

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

$$(DT)(f(x))\neq(TD)(f(x))$$

$$\Rightarrow\ DT\neq TD.$$

Also we see that

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

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

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

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

(i) $A\hat{0}=\hat{0}=\hat{0}A$

(ii) $AI=A=IA$

(iii) $A(BC)=(AB)C$

(iv) $A(B+C)=AB+AC$

(v) $(A+B)C=AC+BC$

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

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

$$(A+B)(\alpha)=A(\alpha)+B(\alpha)$$

$$(cA)(\alpha)=cA(\alpha) \quad ...(1)$$

$$(AB)(\alpha)=A[B(\alpha)] \quad ...(2)$$

$\forall\ \alpha\in V$ and $\forall\ c\in F$. ...(3)

Now we shall prove the above results.

(i) We have $\forall\ \alpha\in V$,

$$(A\hat{0})(\alpha)=A[\hat{0}(\alpha)] \quad \text{[by (3)]}$$

Scanned by TapScanner

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

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

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

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

(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 $\forall\ \alpha\in V$,

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

$=A\,[B(\alpha)+C(\alpha)]$ [by (3)]

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

[$\because$ 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)]

$\therefore$ $c\,(AB)=(cA)\,B$. [by (2)]

Again $[c\,(AB)]\,(\alpha)=c\,[(AB)\,(\alpha)]$ [by (3)]

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

$=A\,[cB(\alpha)]$

[$\because$ A is linear transformation and $B(\alpha)\in V$] [by (2)]

$=A\,[(cB)\,(\alpha)]$ [by (3)]

$=[A\,(cB)]\,(\alpha)$.

$\therefore$ $c\,(AB)=A(cB)$.

§ 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.*

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 $$(ST)\,(\alpha)=S[T(\alpha)]$$

$$\forall\ S,\,T\in L(V,\,V)\ \textit{and}\ \forall\ \alpha\in V.$$

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 $\hat{0}$ 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. 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 $\qquad (\alpha+\beta)\gamma=\alpha\gamma+\beta\gamma \ \forall \ \alpha, \beta, \gamma \in V.$

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

$$T^m T^n=T^{m+n}$$

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

$$p(T)=a_0 I+a_1T+a_2T^2+\ldots+a_nT^n \in L(V, V)$$

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. 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

$T^{-1} : V \to U$ 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.*

(Meerut 1983; Andhra 92)

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

Since T is one-one and onto, therefore there exists 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

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

Scanned by TapScanner

$\therefore$ 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*

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

Proof. Let α be any element of V and let $T(\alpha)=\beta$. Then $T^{-1}(\beta)=\alpha$.

We have $T(\alpha)=\beta$

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

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

$$\Rightarrow T[T^{-1}(\beta)]=T(\alpha)$$
$$\Rightarrow (TT^{-1})(\beta)=\beta$$
$$\Rightarrow (TT^{-1})(\beta)=\beta=I(\beta)$$
$$\Rightarrow TT^{-1}=I.$$

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 $A^{-1}=B=C$. **(Meerut 1968, 69)**

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

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

$\therefore$ 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

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

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

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 $AB=I$

$$\Rightarrow A^{-1}(AB)=A^{-1}I$$
$$\Rightarrow (A^{-1}A)B=A^{-1}$$
$$\Rightarrow IB=A^{-1}$$
$$\Rightarrow B=A^{-1}.$$

Again $CA=I$

$$\Rightarrow (CA)A^{-1}=IA^{-1}$$
$$\Rightarrow C(AA^{-1})=A^{-1}$$
$$\Rightarrow CI=A^{-1}$$
$$\Rightarrow C=A^{-1}.$$

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

$$AC=I=CA.$$

We have $C(AB)=CI=C.$ ...(1)

Also $(CA)B=IB=B.$ ...(2)

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

$$C(AB)=(CA)B$$
$$\Rightarrow C=B.$$

Hence the inverse of A is unique.

Scanned by TapScanner

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

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

Singular and Non-singular transformations. 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*
$$\alpha\in U \text{ and } T(\alpha)=0 \Rightarrow \alpha=0.$$

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

Theorem 7. *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.* **(Allahabad 1976)**

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

Let $\alpha_1, \alpha_2 \in U$. Then
$$T(\alpha_1)=T(\alpha_2)$$
$$\Rightarrow T(\alpha_1)-T(\alpha_2)=0$$
$$\Rightarrow T(\alpha_1-\alpha_2)=0$$
$$\Rightarrow \alpha_1-\alpha_2=0 \qquad [\because T \text{ is non-singular}]$$
$$\Rightarrow \alpha_1=\alpha_2.$$
$\therefore$ T is one-one.

Conversely let T be one-one. We know that $T(0)=0$. Since T is one-one, therefore $\alpha\in U$ and $T(\alpha)=0=T(0)\Rightarrow\alpha=0$. Thus the null space of T consists of zero vector alone. Therefore T is non-singular.

Theorem 8. *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.*

(Meerut 1972, 78, 84 P, 90, 91, 93P)

Proof. First suppose that T is non-singular.

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

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 $a_1, a_2, \ldots, a_n \in F$ and let
$$a_1T(\alpha_1)+\ldots+a_nT(\alpha_n)=0$$
$$\Rightarrow T(a_1\alpha_1+\ldots+a_n\alpha_n)=0 \qquad [\because T \text{ is linear}]$$
$$\Rightarrow a_1\alpha_1+\ldots+a_n\alpha_n=0 \qquad [\because T \text{ is non-singular}]$$
$$\Rightarrow a_i=0,\ i=1, 2, \ldots, n \qquad [\because \alpha_1, \ldots, \alpha_n \text{ are linearly independent}]$$
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 $\alpha\neq 0\in U$. 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

$T(\alpha) \neq 0$ because the set consisting of zero vector alone is linearly dependent. Thus

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

This shows that the null space of T consists of the zero vector alone. Therefore T is non-singular.

Theorem 9. (i) *A linear transformation T on a finite dimensional vector space is invertible iff T is non-singular.*

(Kanpur 1981, Meerut 93)

(ii) *A linear transformation T on a finite dimensional vector space is invertible iff T is onto.*

Proof. (i) Let $V(F)$ be a finite dimensional vector space of dimension n. Let T be a linear transformation on V.

If T is invertible, then T must be one-one. Hence T must be non-singular.

Conversely if T is non-singular, then T must be one-one. Now T will be invertible if we prove that T is onto. For proof see example 3 page 85.

(ii) If T is invertible, then T must be onto.

Conversely let T be onto. Then T will be invertible if we prove that T is one-one. For proof see example 5 page 86.

Important Note. In the above theorem, we have proved that if T is a linear transformation on a finite dimensional vector space V, then T is one-one implies T must be onto. Also T is onto implies T must be one-one. However this theorem fails if V is not finite dimensional. This is clear from the following example :

Example. Let $V(F)$ be the vector space of all polynomials in x with coefficients in F. Let D be the differentiation operator on V. The range of D is all of V and so D is onto. However D is not one-one because all constant polynomials are sent into 0 by D.

Now let T be the linear transformation on V defined by

$$T[f(x)] = x f(x) \; \forall \; f(x) \in V.$$

Then T is one-one.

But T is not onto V. If $g(x)$ is a non-zero constant polynomial in V, then $g(x)$ will not be the image of any element in V under the function T.

Theorem 10. *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 $\{\alpha_1, \ldots\ldots, \alpha_n\}$ be a basis for U. Then $\{\alpha_1, \ldots\ldots, \alpha_n\}$ 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 $a_1, \ldots\ldots, a_n \in F$ 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

$$T(c_1\alpha_1 + \ldots + c_n\alpha_n) = \mathbf{0}$$
$$\Rightarrow c_1T(\alpha_1) + \ldots + c_nT(\alpha_n) = \mathbf{0}$$
$$\Rightarrow c_i = 0,\ 1 \leqslant i \leqslant n \text{ because } T(\alpha_1), \ldots, T(\alpha_n) \text{ are linearly independent}$$
$$\Rightarrow \alpha = \mathbf{0}.$$

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

Solved Examples

Example 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 $e_1 = (1, 0)$, $e_2 = (0, 1)$.*

Solution. Let (x_1, x_2) be any member of F^2. Then we are to find a formula for $T(x_1, x_2)$ under the given conditions that $T(1, 0) = (a, b)$, $T(0, 1) = (c, d)$.

We know that the set $\{e_1, e_2\}$ is a basis for the vector space F^2. Therefore any vector $(x_1, x_2) \in F^2$ can be expressed as a linear combination of the elements of this basis set.

Obviously $(x_1, x_2) = x_1(1, 0) + x_2(0, 1) = x_1e_1 + x_2e_2$.

$$\therefore\ T(x_1, x_2) = T(x_1e_1 + x_2e_2)$$
$$= x_1T(e_1) + x_2T(e_2), \text{ by linearity of } T$$
$$= x_1(a, b) + x_2(c, d) = (x_1a, x_1b) + (x_2c, x_2d)$$
$$= (x_1a + x_2c,\ x_1b + x_2d).$$

Example 2. *Describe explicity the linear transformation $T : \mathbf{R}^2 \to \mathbf{R}^2$ such that $T(2, 3) = (4, 5)$ 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

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

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

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

$$T(x_1, x_2) = T\left[\frac{x_2}{3}(2, 3) + \frac{3x_1 - 2x_2}{3}(1, 0)\right], \text{ by (1)}$$
$$= \frac{x_2}{3}T(2, 3) + \frac{3x_1 - 2x_2}{3}T(1, 0), \text{ by linearity of } T$$
$$= \frac{x_2}{3}(4, 5) + \frac{3x_1 - 2x_2}{3}(0, 0) = \left(\frac{4x_2}{3}, \frac{5x_2}{3}\right).$$

Example 3. *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.*

Solution. Let $(x_1, x_2) \in \mathbf{R}^2$. We can write

$$(x_1, x_2) = x_1(1, 0) + x_2(0, 1).$$

Now $T(x_1, x_2) = T[x_1(1, 0) + x_2(0, 1)] = x_1T(1, 0) + x_2T(0, 1)$

$$= x_1(1, 1) + x_2(-1, 2)$$
$$= (x_1 - x_2,\ x_1 + 2x_2). \qquad \ldots(1)$$

(1) gives the required formula for T. Now let the given vertices of the square be A, B, C, D respectively and let A', B', C', D' be their T-images. We have

$A' = T(A) = T(0, 0) = (0, 0)$, on putting $x_1 = 0$, $x_2 = 0$ in (1)
$B' = T(B) = T(1, 0) = (1, 1)$, on putting $x_1 = 1$, $x_2 = 0$ in (1)
$C' = T(C) = T(1, 1) = (0, 3)$, on putting $x_1 = 1$, $x_2 = 1$ in (1)
$D' = T(D) = T(0, 1) = (-1, 2)$, on putting $x_1 = 0$, $x_2 = 1$ in (1).

Now $A'B' = \sqrt{2} = C'D'$. Also $A'D' = \sqrt{5} = B'C'$. Hence $A'B'C'D'$ is a parallelogram.

Example 4. *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 (**R**) into V_3 (**R**) such that

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

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 (**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 (**R**). Then we can write

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

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

Example 5. *Let T be a linear operator on V_3 (**R**) defined by*

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

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

(Meerut 1982, 87, 89; Nagarjuna 91)

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

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

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

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

$\therefore$ T is one-one.

Now T is a linear transformation on a finite dimensional vector space V_3 (**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 $T^{-1}(p,q,r)=(a,b,c)$.

Now $T(a,b,c)=(p,q,r)$

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

$$\begin{aligned} \Rightarrow a=\frac{p}{3},\ b=\frac{p}{3}-q,\ c=r-2a-b &= r-\frac{2p}{3}-\frac{p}{3}+q \\ &= r-p+q. \end{aligned}$$

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

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

Example 6. *For the linear operator T of Ex. 5, prove that*

$$(T^2-I)(T-3I)=\hat{0}.$$

Solution. We have

$$\begin{aligned} (T-3I)(a,b,c) &= T(a,b,c)-3I(a,b,c) \\ &= (3a,a-b,2a+b+c)-3(a,b,c), \quad \text{by def. of } T \text{ and } I \\ &= (0,a-4b,2a+b-2c). \qquad \ldots(1) \end{aligned}$$

$$\begin{aligned} \therefore\quad (T^2-I)(T-3I)(a,b,c) &= (T^2-I)[(T-3I)(a,b,c)] \\ &= (T^2-I)(0,a-4b,2a+b-2c), \text{ by (1)} \\ &= T^2(A,B,C)-I(A,B,C), \qquad \ldots(2) \end{aligned}$$

where $A=0,\ B=a-4b,\ C=2a+b-2c$.

Now $T(A,B,C)=(3A,A-B,2A+B+C)$, by def. of T.

$$\begin{aligned} \therefore\quad T^2(A,B,C) &= T(3A,A-B,2A+B+C) \\ &= T(l,m,n), \text{ say} \\ &= (3l,l-m,2l+m+n), \text{ by def. of } T \\ &= (9A,3A-A+B,6A+A-B+2A+B+C), \\ &\qquad \text{on putting the values of } l,m,n \\ &= (9A,2A+B,9A+C) \\ &= (0,a-4b,2a+b-2c). \end{aligned}$$

Also $I(A,B,C)=(A,B,C)=(0,a-4b,2a+b-2c)$.

Hence from (2), we have

$$\begin{aligned} (T^2-I)(T-3I)(a,b,c) &= T^2(A,B,C)-I(A,B,C) \\ &= (0,a-4b,2a+b-2c)-(0,a-4b,2a+b-2c) \\ &= (0,0,0)=\hat{0}(a,b,c)\ \forall\ (a,b,c)\in V_3(\mathbf{R}). \end{aligned}$$

Therefore, by def. of zero transformation, we have

$$(T^2-I)(T-3I)=\hat{0}.$$

Example 7. *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 $\alpha\delta-\beta\gamma\neq 0$.*

(Nagarjuna 1977)

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 8. *Find two linear operators T and S on V_2 (R) such that*

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

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

$T(a, b)=(a, 0) \; \forall \; (a, b) \in V_2$ **(R)**

and $S(a, b)=(0, a) \; \forall \; (a, b) \in V_2$ **(R)**.

We have $(TS)(a, b)=T[S(a, b)]=T(0, a)=(0, 0)$

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

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

Again $(ST)(a, b)=S[T(a, b)]=S(a, 0)=(0, a)$

$\neq \hat{0}(a, b) \; \forall \; (a, b) \in V_2$ **(R)**.

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

Example 9. *Let V be a vector space over the field F and T a linear operator on V. If $T^2=\hat{0}$, 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 $T^2=\hat{0}$ but $T \neq \hat{0}$.*

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

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

$\Rightarrow T[T(\alpha)]=0 \; \forall \; \alpha \in V$

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

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

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

For the second part of the question, consider the linear transformation T on V_2 (R) defined by

$T(a, b)=(0, a) \; \forall \; (a, b) \in V_2$ **(R)**.

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

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

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

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

Example 10. *Let T : $\mathbf{R}^3 \to \mathbf{R}^3$ be defined as $T(a, b, c)=(0, a, b)$. Show that $T \neq \hat{0}$, $T^2 \neq \hat{0}$ but $T^3=\hat{0}$.*

Solution. We have $T(a, b, c)=(0, a, b)$.

Therefore $T(a, b, c) \neq (0, 0, 0) \; \forall \; (a, b, c) \in \mathbf{R}^3$.

Hence $T \neq \hat{0}$.

Again $T^2(a, b, c)=T[T(a, b, c)]=T(0, a, b)=(0, 0, a)$.

Thus $T^2(a, b, c) \neq (0, 0, 0) \; \forall \; (a, b, c) \in \mathbf{R}^3$. Hence $T^2 \neq \hat{0}$.

Finally $T^3(a, b, c)=T^2[T(a, b, c)]=T^2(0, a, b)=T[T(0, a, b)]$ $=T(0, 0, a)=(0, 0, 0)$. Thus $T^3(a, b, c)=(0, 0, 0) \; \forall \; (a, b, c) \in \mathbf{R}^3$. In other words $T^3(a, b, c)=\hat{0}(a, b, c) \; \forall \; (a, b, c) \in \mathbf{R}^3$. Hence $T^3=\hat{0}$.

Example 11. *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 $T\alpha_1=T\alpha_2$.* **(Meerut 1977)**

Solution. Let α_1 be the zero vector of U. Then $T\alpha_1=\mathbf{0}$.

Since $\ker T \neq \mathbf{0}$, therefore there exists a non-zero vector, say α_2, in U such that $T\alpha_2=\mathbf{0}$.

Now we see that α_1, α_2 are vectors in U such that $\alpha_1 \neq \alpha_2$ and $T\alpha_1=T\alpha_2$.

Example 12. *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=rank (T)+nullity (T)

=dim of range of T+dim of null space of T.

$\therefore$ 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 $T \leqslant 3$ 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 13. *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 $(AB)(B^{-1}A^{-1})=I=(B^{-1}A^{-1})(AB)$.

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

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

$\therefore$ A is one-one. $\qquad$ [$\because$ BA is one-one]

A **is onto.**

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

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

Thus $\beta \in V$ [illegible] $B($ $) \in V$ su[illegible] $A[B(\alpha)]=\beta$.

$\therefore$ A [illegible]

$\therefore$ A is invertible.

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

Example 14. *Let T be a linear transformation from $V_3(\mathbf{R})$ into $V_2(\mathbf{R})$, and let S be a linear transformation from $V_2(\mathbf{R})$ into $V_3(\mathbf{R})$. Prove that the transformation ST is not invertible.*

Solution. As proved in Ex. 12, we can prove that T cannot be non-singular i.e. T cannot be one-one.

Let ST be invertible. Then ST is one-one.

Let $\alpha_1, \alpha_2 \in V_3(\mathbf{R})$. Then

$$T(\alpha_1)=T(\alpha_2)$$
$$\Rightarrow S[T(\alpha_1)]=S[T(\alpha_2)] \Rightarrow (ST)(\alpha_1)=(ST)(\alpha_2)$$

[$\because$ ST is one-one]

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

$\therefore$ T is one-one which is not possible.

$\therefore$ ST cannot be invertible.

Example 15. *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. **(Meerut 1983P)**

Solution. First we shall show that B is invertible.

B is one-one. Let $\alpha_1, \alpha_2 \in V$.

Then $\qquad B(\alpha_1)=B(\alpha_2)$

$$\Rightarrow A[B(\alpha_1)]=A[B(\alpha_2)] \Rightarrow (AB)(\alpha_1)=(AB)(\alpha_2)$$
$$\Rightarrow I(\alpha_1)=I(\alpha_2) \Rightarrow \alpha_1=\alpha_2.$$

$\therefore$ B is one-one.

Now B is a linear transformation on a finite dimensional vector space V.

Therefore B is one-one implies that B must be onto. Consequently B is invertible.

Now $\qquad AB=I$

$$\Rightarrow (AB)B^{-1}=IB^{-1} \Rightarrow A(BB^{-1})=B^{-1}$$
$$\Rightarrow AI=B^{-1} \Rightarrow A=B^{-1}.$$

But $\qquad B^{-1}B=I=BB^{-1}$

$$\Rightarrow AB=I=BA$$
$$\Rightarrow A \text{ is invertible and } A^{-1}=B.$$

Example. Let $V(\mathbf{R})$ [illegible] the vector space of all polynomials in x with elements in $\mathbf{R}$. Consider the linear transformations D and T on V defined as follows :

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

and $$T[f(x)]=\int_0^x f(x)\,dx \ \forall\, f(x) \in V.$$

Here $\qquad DT=I.$

But D is not invertible because D is not one-one.

Example 16. *If A is a linear transformation on a vector space V such that*
$$A^2-A+I=\hat{0},$$
then A is invertible. **(Meerut 1969, 88)**

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

First we shall prove that A is one-one.

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

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

$\Rightarrow A[A(\alpha_1)]=A[A(\alpha_2)]$ ...(1)

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

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

$\Rightarrow (A^2-A)(\alpha_1)=(A^2-A)(\alpha_2)$ [From (2) and (1)]

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

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

$\therefore$ A is one-one.

Now to prove that A is onto.

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

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

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

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

$=\alpha.$

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 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.* **(Kanpur 1980)**

Solution. Given that A has a unique right inverse B *i.e.* $AB=I$ and B is unique.

We have $A(BA+B-I)$

$=A(BA)+AB-AI=(AB)A+AB-A$

$=IA+I-A=A+I-A=I.$

$\therefore$ $BA+B-I$ is a right inverse of A. But B is the unique right inverse of A.

$\therefore$ $BA+B-I=B$

$\Rightarrow$ $BA=I.$

Thus $AB=I=BA.$

$\therefore$ A is invertible.

Example 18. *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=rank (T)+nullity (T)

and dim V=rank (T^2)+nullity (T^2).

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

nullity (T)=nullity (T^2)

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

Now $T(\alpha)=0$

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

$\therefore$ $\alpha \in$ 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 Ex. 7 page 118]

Example. 19. *Let V be a finite dimensional vector space over the field F.*

Let $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ *and* $\{\beta_1, \beta_2, ..., \beta_n\}$

be two ordered bases for V. Show that there exists a unique invertible linear transformation T on V such that
$$T(\alpha_i)=\beta_i,\ 1 \leqslant i \leqslant n.$$

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

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

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

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

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

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

Scanned by TapScanner

$\Rightarrow a_i = 0$ for each $1 \leqslant i \leqslant n$

$\Rightarrow \alpha = 0$. [$\because$ $\beta_1, \ldots, \beta_n$ are linearly independent]

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

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

Solution. Let dim $V = n$.

Since $\{\alpha_1, \ldots, \alpha_k\}$ is a linearly independent subset of V, therefore it can be extended to form a basis for V. Let

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

be a basis for V. Similarly let $\{\beta_1, \ldots, \beta_k, \beta_{k+1}, \ldots, \beta_n\}$ be a basis for V.

Now there exists a unique linear transformation T on V such that $T(\alpha_j) = \beta_j,\ j = 1, 2, \ldots, k, \ldots, n$.

Also T is invertible because T maps a basis of V onto a basis of V and V is finite dimensional.

Thus there exists an invertible linear transformation T on V such that $T(\alpha_i) = \beta_i,\ i = 1, \ldots, k$.

Example 21. *Let T_1 be a linear transformation on a vector space $V(F)$. Prove that the set of all linear transformations S on V for which $T_1 S = \hat{0}$ is a subspace of the vector space of all linear transformations.*

Solution. Let L denote the vector space of all linear transformations on the vector space V. Let $W = \{S : S$ is a linear transformation on V and $T_1 S = \hat{0}\}$. To prove that W is a subspace of L.

Let $a, b \in F$ and $S_1, S_2 \in W$. Then by def. of W, we have $T_1 S_1 = \hat{0}$ and $T_2 T_2 = \hat{0}$. We shall show that

$$T_1(aS_1 + bS_2) = \hat{0}.$$

Let $\alpha \in V$. Then

$[T_1(aS_1 + bS_2)](\alpha) = T_1[(aS_1 + bS_2)(\alpha)]$,

by def. of product of linear transformations

$= T_1[(aS_1)(\alpha) + (bS_2)(\alpha)] = T_1[aS_1(\alpha) + bS_2(\alpha)]$

$= aT_1[S_1(\alpha)] + bT_1[S_2(\alpha)]$, since T_1 is a linear transformation

$= a(T_1 S_1)(\alpha) + b(T_1 S_2)(\alpha) = a\hat{0}(\alpha) + b\hat{0}(\alpha)$

$= a\,0 + b\,0 = 0 = \hat{0}(\alpha)$.

Thus $[T_1(aS_1 + bS_2)](\alpha) = \hat{0}(\alpha)\ \forall\ \alpha \in V$.

Therefore $T_1(aS_1 + bS_2) = \hat{0}$. Consequently by def. of W, $aS_1 + bS_2 \in W$. Thus $a, b \in F$ and $S_1, S_2 \in W \Rightarrow aS_1 + bS_2 \in W$. Hence W is a subspace of L.

Exercises

1. Fill up the blanks in the following statements :
   (i) A linear operator T on $\mathbf{R}^2$ defined by $T(x, y) = (ax + by, cx + dy)$ will be invertible iff...... **(Meerut 1976)**
   (ii) If T is a linear operator on $\mathbf{R}^2$ defined by $T(x, y) = (x - y, y)$, then $T^2(x, y) = \ldots\ldots$, **(Meerut 1976)**

   Ans. (i) $ad - bc \neq 0$; (ii) $(x - 2y, y)$.
2. State whether the following statements are true or false :
   (i) For two linear operators T and U on $\mathbf{R}^2$,
   $TU = \hat{0} \Rightarrow UT = \hat{0}$. **(Meerut 1976)**
   (ii) If S and T are linear operators on a vector space U, then $(S + T)^2 = S^2 + 2ST + T^2$. **(Meerut 1977)**

   Ans. (i) False; (ii) False, because it is not necessary that $ST = TS$.
3. If T is a linear transformation from a vector space V into a vector space W, then obtain conditions for T^{-1} to be a linear transformation from W to V. **(Meerut 1976)**

   Ans T should be one-one and onto.
4. Show that the identity operator on a vector space is always invertible.
5. 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. **(Meerut 1976, 80)**
6. A linear operator T on a vector space V is said to be **nilpotent** if there exists a positive integer r such that $T^r = \hat{0}$.
   If V is the space of all polynomials of degree less than or equal to n over a field F, prove that the differentiation operator on V is nilpotent. **(Meerut 1976)**
7. Describe explicitly a linear transformation from $V_3(\mathbf{R})$ into $V_4(\mathbf{R})$ which has its range the subspace spanned by the vectors $(1, 2, 0, -4), (2, 0, -1, 3)$.

   Ans. $T(a, b, c) = (a + 2b, 2a, -b, -4a - 3b)$.

Scanned by TapScanner

8. 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.
Ans $T^{-1}(a, b)=(b, a-b)$.

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

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.
Hint The vectors $e_1=(1, 0, 0)$, $e_2=(0, 1, 0)$, $e_3=(0, 0, 1)$, form a basis for $\mathbf{C}^3$. Show that $T(e_1)$, $T(e_2)$, $T(e_3)$, are linearly dependent vectors. Consequently they do not form a basis for $\mathbf{C}^3$. Since T does not map a basis of $\mathbf{C}^3$ onto a basis of $\mathbf{C}^3$, therefore T is not invertible.

11. 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 $L(W, W)$.

12. Give an example of a one-one linear transformation of an infinite dimensional vector space which is not an isomorphism. **(Poona 1970)**

13. Show that if two linear transformations of a finite dimensional vector space coincide on a basis of that vector space, then they are identical. **(Poona 1970)**

14. 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 $T(\alpha)=\mathbf{0}$. **(Banaras 1972)**

15. Let V and W be n-dimensional vector spaces over the field F. Show that for any linear transformation $T: V \to W$, the following statements are all equivalent :
(i) T is invertible.
(ii) $T(V)=W$.
(iii) There is a basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ for V such that $\{T\alpha_1, T\alpha_2, \ldots, T\alpha_n\}$ is a basis for W. **(Meerut 1973)**

16. Let T be a linear transformation from a finite dimensional vector space V into a finite dimensional vector space W. Show that
(i) $T(\mathbf{0})=\mathbf{0}$.
(ii) If U is a subspace of V, then the image of U under T is also a subspace of W.
(iii) If dim V = dim W and if $T(V)=W$, then T is invertible. **(Meerut 1974)**

17. 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}. **(Meerut 1980)**
Ans. $T^{-1}(x, y, z)=(\tfrac{1}{2}x+\tfrac{1}{2}y, z, \tfrac{1}{2}x-\tfrac{1}{2}y)$.

§ 12. 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
$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

In a compact form the above matrix is represented by $A=[a_{ij}]_{m \times n}$. The element a_{ij} is called the $(i, j)^{th}$ 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 $m \times n$.

Scanned with CamScanner

Equality of two matrices. Definition.

Let $A=[a_{ij}]_{m\times n}$ and $B=[b_{ij}]_{m\times n}$. 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 $(i, k)^{th}$ 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 inter-changing 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 $(i, j)^{th}$ element of A^T is the $(j, i)^{th}$ element of A. If A is an $m\times n$ matrix and B is an $n\times p$ matrix, it can be shown that $(AB)^T=B^TA^T$. 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 (adj A)=(adj A) $A=(\det A)\, I$ where I is 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 $\qquad B=A^{-1}$.

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

$$A^{-1}=\frac{1}{\det A}(\text{adj. } A).$$

Scanned with CamScanner

(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 $\alpha_1,\ldots\ldots\ldots,\alpha_m\in V_n(F)$ defined by $\alpha_i=(a_{i1},a_{i2},\ldots,a_{in}),\ 1\leqslant i\leqslant m$.

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 $\beta_j=(a_{1j},a_{2j},\ldots,a_{mj}),\ 1\leqslant j\leqslant n$.

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.

§ 13. Representation of transformations by matrices

Matrix of a linear transformation.

(Meerut 1980, 83; I.A.S. 88)

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

$[T]=[T;B;B']=$ matrix of T relative to ordered bases B and B'

$$=[a_{ij}]_{m\times n}$$

where $\quad T(\alpha_j)=\sum_{i=1}^{m}a_{ij}\beta_i$, for each $j=1,2,\ldots,n$ ...(1)

The coordinates of $T(\alpha_1)$ in ordered basis B' form the first column of this matrix, the coordinates of $T(\alpha_2)$ in ordered basis B' form the second column of this matrix and so on.

Scanned with CamScanner

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 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

$[T]_B$ or $[T; B]$ = matrix of T relative to ordered basis B

$$= [a_{ij}]_{n \times n},$$

where $T(\alpha_j) = \sum_{i=1}^{n} a_{ij} \alpha_i$, for each $j = 1, 2, \ldots, n$.

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 $T(\alpha_1) = T(1, 0) = (1, 0)$.

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 $T(\alpha_2) = T(0, 1) = (0, 0) = 0\,(1, 0) + 0\,(0, 1)$.

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

$$= [T]_B \text{ or } [T; B] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

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*

$$[T; B; B'] = [a_{ij}]_{m \times n}. \qquad \textbf{(I.A.S. 1988)}$$

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 page 123 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 :

Scanned with CamScanner

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*

$$[T(\alpha)]_{B'} = A\ [\alpha]_B \text{ where}$$

$[\alpha]_B$ 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'.

(Meerut 1980)

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

$$T(\alpha_j)=\sum_{i=1}^{m} a_{ij}\,\beta_i,\ j=1, 2, \ldots, n. \qquad \ldots(1)$$

If $\alpha = x_1\alpha_1 + \ldots + x_n\alpha_n$ is a vector in U, then

$$T(\alpha)=T\left(\sum_{j=1}^{n} x_j\,\alpha_j\right)$$

$$=\sum_{j=1}^{n} x_j\,T(\alpha_j) \qquad [\because T \text{ is a linear transformation}]$$

$$=\sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\,\beta_i \qquad [\text{From } (1)]$$

$$=\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}\,x_j\right)\beta_i. \qquad \ldots(2)$$

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

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

$\therefore\ [T(\alpha)]_{B'}=AX=A\ [\alpha]_B=[U;\ B;\ B']\ [\alpha]_B.$

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 (ii) $[\hat{0};\ B]=$ *null matrix of the type $n \times n$.*

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

$$[I;\ B]=[\delta_{ij}]_{n \times n}=I \text{ i.e. unit matrix of order } n.$$

(ii) We have $\hat{0}(\alpha_j)=0,\ j=1, 2, \ldots, n$

$$=0\alpha_1+0\alpha_2+\ldots+0\alpha_n$$

$$=\sum_{i=1}^{n} o_{ij}\,\alpha_i, \text{ where each } o_{ij}=0.$$

$\therefore$ By def. of matrix of a linear transformation, we have

$[\hat{0};\ B]=[o_{ij}]_{n \times n}=$ null matrix of the type $n \times n$.

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

$$B'=\{\beta_1, \beta_2, \ldots, \beta_m\}.$$

Scanned with CamScanner

Let $[a_{ij}]_{m\times n}$ be the matrix of T relative to B, B'. Then

$$T(\alpha_j)=\sum_{i=1}^{m} a_{ij}\,\beta_i,\ j=1, 2, \dots, n.$$

Also let $[b_{ij}]_{m\times n}$ be the matrix of S relative to B, B'. Then

$$S(\alpha_j)=\sum_{i=1}^{m} b_{ij}\,\beta_i,\ j=1, 2, \dots, n.$$

(i) We have

$$(T+S)\,(\alpha_j)=T(\alpha_j)+S(\alpha_j),\ j=1, 2, \dots, n$$

$$=\sum_{i=1}^{m} a_{ij}\,\beta_i+\sum_{i=1}^{m} b_{ij}\,\beta_i=\sum_{i=1}^{m}(a_{ij}+b_{ij})\,\beta_i.$$

$\therefore$ 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 $(cT)\,(\alpha_j)=cT\,(\alpha_j),\ j=1, 2, \dots, n$

$$=c\sum_{i=1}^{m} a_{ij}\,\beta_i=\sum_{i=1}^{m}(ca_{ij})\,\beta_i.$$

$\therefore$ $[cT;\ B;\ B']=$ matrix of cT relative to B, B'
$=[ca_{ij}]_{m\times n}=c\,[a_{ij}]_{m\times n}=c\,[T;\ B;\ B']$.

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 $C=DA$.*

(Meerut 1975)

Proof. Let dim $U=n$, dim $V=m$ and dim $W=p$. Further let
$B=\{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $B'=\{\beta_1, \beta_2, \dots, \beta_m\}$
and $B''=\{\gamma_1, \gamma_2, \dots, \gamma_p\}$.
Let $A=[a_{ij}]_{m\times n}$, $D=[d_{ki}]_{p\times m}$ and $C=[c_{kj}]_{p\times n}$. Then

$$T(\alpha_j)=\sum_{i=1}^{m} a_{ij}\,\beta_i,\ j=1, 2, \dots, n, \qquad \dots(1)$$

$$S(\beta_i)=\sum_{k=1}^{p} d_{ki}\,\gamma_k,\ i=1, 2, \dots, m, \qquad \dots(2)$$

$$\text{and } (ST)\,(\alpha_j)=\sum_{k=1}^{p} c_{kj}\,\gamma_k,\ j=1, 2, \dots, n. \qquad \dots(3)$$

We have $(ST)\,(\alpha_j)=S\,[T\,(\alpha_j)],\ j=1, 2, \dots, n$

$$=S\left(\sum_{i=1}^{m} a_{ij}\,\beta_i\right) \qquad \text{[From (1)]}$$

$$=\sum_{i=1}^{m} a_{ij}\,S\,(\beta_i) \qquad [\because\ S \text{ is linear}]$$

$$=\sum_{i=1}^{m} a_{ij}\sum_{k=1}^{p} d_{ki}\,\gamma_k \qquad \text{[From (2)]}$$

$$=\sum_{k=1}^{p}\left(\sum_{i=1}^{m} d_{ki}\,a_{ij}\right)\gamma_k. \qquad \dots(4)$$

Therefore from (3) and (4), we have

$$c_{kj}=\sum_{i=1}^{m} d_{ki}\,a_{ij},\ j=1, 2, \dots, n\ ;\ k=1, 2, \dots, p.$$

$$\therefore\ [c_{kj}]_{p\times n}=\left[\sum_{i=1}^{m} d_{ki}\,a_{ij}\right]_{p\times n}$$

$=[d_{ki}]_{p\times m}\,[a_{ij}]_{m\times n}$, 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

$C=DA$ *i.e.* $[ST]_B=[S]_B\,[T]_B$. **(Banaras 1972)**

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

$$T(\alpha_j)=\sum_{i=1}^{n} a_{ij}\,\alpha_i,\ j=1, 2, \dots, n, \qquad \dots(1)$$

$$S(\alpha_i)=\sum_{k=1}^{n} d_{ki}\,\alpha_k,\ i=1, 2, \dots, n, \qquad \dots(2)$$

Scanned with CamScanner

and $(ST)(\alpha_j)=\sum_{k=1}^{n} c_{kj}\,\alpha_k,\ j=1,2,\ldots,n.$ ...(3)

We have $(ST)(\alpha_j)=S[T(\alpha_j)]$

$$=S\left(\sum_{i=1}^{n} a_{ij}\,\alpha_i\right)=\sum_{i=1}^{n} a_{ij}\,S(\alpha_i)=\sum_{i=1}^{n} a_{ij}\sum_{k=1}^{n} d_{ki}\,\alpha_k$$

$$=\sum_{k=1}^{n}\left(\sum_{i=1}^{n} d_{ki}\,a_{ij}\right)\alpha_k. \qquad \ldots(4)$$

$\therefore$ from (3) and (4), we have $c_{kj}=\sum_{i=1}^{n} d_{ki}\,a_{ij}$.

$$\therefore\quad [c_{kj}]_{n\times n}=\left[\sum_{i=1}^{n} d_{ki}\,a_{ij}\right]_{n\times n}=[d_{ki}]_{n\times n}\,[a_{ij}]_{n\times n}$$

$\therefore\ C=DA.$

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.* (Meerut 19[illegible])

Proof. Let $B=\{\alpha_1,\ldots,\alpha_n\}$ and

$B'=\{\beta_1,\ldots,\beta_m\}$.

Let M be the vector space of all $m\times n$ matrices over the field F. Let

$\psi: L(U,V)\longrightarrow M$ such that

$\psi(T)=[T;B;B']\ \forall\ T\in L(U,V)$.

Let $T_1, T_2\in L(U,V)$; and let

$[T_1;B;B']=[a_{ij}]_{m\times n}$ and $[T_2;B;B']=[b_{ij}]_{m\times n}$.

Then $T_1(\alpha_j)=\sum_{i=1}^{m} a_{ij}\,\beta_i,\ j=1,2,\ldots,n$

and $T_2(\alpha_j)=\sum_{i=1}^{m} b_{ij}\,\beta_i,\ j=1,2,\ldots,n.$

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 $i=1,\ldots,m$ and $j=1,\ldots,n$

$\Rightarrow \sum_{i=1}^{m} a_{ij}\,\beta_i=\sum_{i=1}^{m} b_{ij}\,\beta_i$ for $j=1,\ldots,n$

$\Rightarrow T_1(\alpha_j)=T_2(\alpha_j)$ for $j=1,\ldots,n$

$\Rightarrow T_1=T_2$ [$\because$ 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 $[T;B;B']=[c_{ij}]_{m\times n}$

$\Rightarrow \psi(T)=[c_{ij}]_{m\times n}$.

$\therefore$ ψ is onto.

ψ is a linear transformation.

If $a,b\in F$, then

$\psi(aT_1+bT_2)=[aT_1+bT_2;B;B']$ [by def. of ψ]

$=[aT_1;B;B']+[bT_2;B;B']$ [by theorem 4]

$=a\,[T_1;B;B']+b\,[T_2;B;B']$ [by theorem 4]

$=a\psi(T_1)+b\psi(T_2)$, by def. of ψ.

$\therefore$ 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.*,

$\psi(T_1T_2)=\psi(T_1)\,\psi(T_2)$

and $\psi(I)=I$ *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

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

$\Rightarrow [T^{-1}T]_B=[I]_B=[TT^{-1}]_B$

$\Rightarrow [T^{-1}]_B\,[T]_B=I=[T]_B\,[T^{-1}]_B$

$\Rightarrow [T]_B$ 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 $CA = I = AC$

$\Rightarrow [S]_B\,[T]_B = I = [T]_B\,[S]_B$

$\Rightarrow [ST]_B = [I]_B = [TS]_B$

$\Rightarrow ST = I = TS$

$\Rightarrow T$ is invertible.

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'.

Theorem 8. *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) $[\alpha]_B = A[\alpha]_{B'}$

(2) $[\alpha]_{B'} = A^{-1}[\alpha]_B$

for every vector α in V. **(Meerut 1984P, 93P)**

Solution. 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 $[T]_B = A = [a_{ij}]_{n\times n}$. Then

$$T(\alpha_j) = \sum_{i=1}^{n} a_{ij}\,\alpha_i,\ j = 1, 2, \ldots, n. \qquad \ldots(2)$$

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

$\therefore\ x_i = \sum_{j=1}^{n} a_{ij}\,y_j$ because the **expression for α 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 $[\alpha]_B = x_i = \sum_{j=1}^{n} a_{ij}\,y_j$

$= i^{th}$ entry of $A\,[\alpha]_{B'}$.

$\therefore\ [\alpha]_B = A\,[\alpha]_{B'}$

$\Rightarrow A^{-1}\,[\alpha]_B = A^{-1}A\,[\alpha]_{B'}$

$\Rightarrow A^{-1}\,[\alpha]_B = I\,[\alpha]_{B'}$

$\Rightarrow A^{-1}\,[\alpha]_B = [\alpha]_{B'}$.

Note. The matrix $A = [T]_B$ is called the **transition matrix** from B to B'. It 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

$$\beta_1 = a_{11}\alpha_1 + a_{21}\alpha_2 + \ldots + a_{n1}\alpha_n$$
$$\beta_2 = a_{12}\alpha_1 + a_{22}\alpha_2 + \ldots + a_{n2}\alpha_n$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots$$
$$\beta_n = a_{1n}\alpha_1 + a_{2n}\alpha_2 + \ldots + a_{nn}\alpha_n.$$

Then the matrix $A = [a_{ij}]_{n\times n}$ *i.e.* A is the transpose of the matrix of coefficients in the above relations. Thus

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Now suppose α is any vector in V. If $[\alpha]_B$ is the coordinate matrix of α relative to the basis B and $[\alpha]_{B'}$ its coordinate matrix relative to the basis B' then

$$[\alpha]_B=A\,[\alpha]_{B'}$$

and
$$[\alpha]_{B'}=A^{-1}\,[\alpha]_B.$$

Theorem 9. *Let $B=\{\alpha_1,\dots,\alpha_n\}$ and $B'=\{\beta_1,\dots,\beta_n\}$ be two ordered bases for an n-dimensional vector space $V(F)$. If $(x_1,\dots,x_n)$ 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

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

Proof. We have $T(\alpha)=T\left(\sum_{i=1}^{n} x_i\,\alpha_i\right)$

$$=\sum_{i=1}^{n} x_i\,T(\alpha_i) \qquad [\because\ T \text{ is linear}]$$

$$=\sum_{i=1}^{n} x_i\beta_i=\beta.$$

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* $B=C^{-1}AC.$ **(Meerut 1976)**

Theorem 10. *The relation of similarity is an equivalence relation in the set of all $n\times n$ matrices over the field F.*

(Meerut 1969, 76; Kanpur 81)

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

$$A=P^{-1}BP$$
$$\Rightarrow\ PAP^{-1}=P\,(P^{-1}BP)\,P^{-1}$$
$$\Rightarrow\ PAP^{-1}=B$$
$$\Rightarrow\ B=PAP^{-1}$$
$$\Rightarrow\ B=(P^{-1})^{-1}AP^{-1}$$

[$\because$ P is invertible means P^{-1} is invertible and $(P^{-1})^{-1}=P$]

$\Rightarrow$ B is similar to A.

Transitive. Let A be similar to B and B be similar to C. Then

$$A=P^{-1}BP$$

and
$$B=Q^{-1}CQ,$$

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 11. *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 12. *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

$$A=PBP^{-1}$$
$$\Rightarrow P^{-1}AP=P^{-1}(PBP^{-1})P$$
$$\Rightarrow P^{-1}AP=B \Rightarrow B=P^{-1}AP$$
$$\Rightarrow B=P^{-1}A(P^{-1})^{-1} \Rightarrow B \text{ is similar to } A.$$

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 13. *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.* **(Andhra 1992)**

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

$$T(\alpha_j)=\sum_{i=1}^{n} a_{ij}\,\alpha_i,\ j=1, 2, \ldots, n \qquad \ldots(1)$$

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
$$T(\beta_j)=T[S(\alpha_j)] \qquad \text{[From (3)]}$$
$$=T\left(\sum_{k=1}^{n} p_{kj}\,\alpha_k\right) \qquad \text{[From (4), on replacing } i \text{ by } k \text{ which is immaterial]}$$
$$=\sum_{k=1}^{n} p_{kj}\,T(\alpha_k) \qquad [\because T \text{ is linear}]$$
$$=\sum_{k=1}^{n} p_{kj}\sum_{i=1}^{n} a_{ik}\,\alpha_i \qquad \text{[From (1), on replacing } j \text{ by } k]$$
$$=\sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}\,p_{kj}\right)\alpha_i. \qquad \ldots(5)$$

Also
$$T(\beta_j)=\sum_{k=1}^{n} c_{kj}\,\beta_k \qquad \text{[From (2), on replacing } i \text{ by } k]$$
$$=\sum_{k=1}^{n} c_{kj}\,S(\alpha_k) \qquad \text{[From (3)]}$$
$$=\sum_{k=1}^{n} c_{kj}\sum_{i=1}^{n} p_{ik}\,\alpha_i \qquad \text{[From (4), on replacing } j \text{ by } k]$$
$$=\sum_{i=1}^{n}\left(\sum_{k=1}^{n} p_{ik}\,c_{kj}\right)\alpha_i. \qquad \ldots(6)$$

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]
$$\Rightarrow AP=PC$$
$$\Rightarrow P^{-1}AP=P^{-1}PC \qquad [\because P^{-1} \text{ exists}]$$
$$\Rightarrow P^{-1}AP=IC \Rightarrow P^{-1}AP=C$$
$\Rightarrow$ C is similar to A.

Scanned with CamScanner

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 14. *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

$$T_1(\alpha_j)=\sum_{i=1}^{n} a_{ij}\alpha_i,\ j=1,2,\ldots,n, \quad \ldots(1)$$

and $$T_2(\beta_j)=\sum_{i=1}^{n} a_{ij}\beta_i,\ j=1,2,\ldots,n. \quad \ldots(2)$$

Let S be the linear operator on V defined by

$$S(\alpha_j)=\beta_j,\ j=1,2,\ldots,n. \quad \ldots(3)$$

Since S maps a basis of V onto a basis of V, therefore S is invertible.

We have $$T_2(\beta_j)=T_2[S(\alpha_j)] \quad \text{[From (3)]}$$
$$=(T_2S)(\alpha_j). \quad \ldots(4)$$

Also $$T_2(\beta_j)=\sum_{i=1}^{n} a_{ij}\beta_i \quad \text{[From (2)]}$$
$$=\sum_{i=1}^{n} a_{ij}S(\alpha_i) \quad \text{[From (3)]}$$
$$=S\left(\sum_{i=1}^{n} a_{ij}\alpha_i\right) \quad [\because S \text{ is linear}]$$
$$=S[T_1(\alpha_j)] \quad \text{[From (1)]}$$
$$=(ST_1)(\alpha_j). \quad \ldots(5)$$

From (4) and (5), we have

$$(T_2S)(\alpha_j)=(ST_1)(\alpha_j),\ j=1,2,\ldots,n.$$

Since T_2S and ST_1 agree on a basis for V, therefore we have

$$T_2S=ST_1$$

$$\Rightarrow T_2SS^{-1}=ST_1S^{-1} \Rightarrow T_2I=ST_1S^{-1}$$
$$\Rightarrow T_2=ST_1S^{-1} \Rightarrow T_2 \text{ is similar to } T_1.$$

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

$$[T]_B \text{ and } [T]_{B'}$$

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. (Kanpur 1981). *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

$$\text{tr } A=\sum_{i=1}^{n} a_{ii}=a_{11}+a_{22}+\ldots+a_{nn}.$$

In the following two theorems we have given some fundamental properties of the trace function.

Theorem 15. *Let A and B be two square matrices of order n over a field F and $\lambda\in F$. Then*

(1) $tr(\lambda A)=\lambda\ tr\ A$;

(2) $tr(A+B)=tr\ A+tr\ B$;

(3) $tr(AB)=tr(BA)$. **(Poona 1970)**

Proof. Let $A=[a_{ij}]_{n\times n}$ and $B=[b_{ij}]_{n\times n}$.

(1) We have $\lambda A=[\lambda a_{ij}]_{n\times n}$, by def. of multiplication of a matrix by a scalar.

Scanned with CamScanner

$\therefore \quad \operatorname{tr}(\lambda A)=\sum_{i=1}^{n} \lambda a_{ii}=\lambda \sum_{i=1}^{n} a_{ii}=\lambda \operatorname{tr} A.$

(2) We have $A+B=[a_{ij}+b_{ij}]_{n\times n}$.

$\therefore \quad \operatorname{tr}(A+B)=\sum_{i=1}^{n}(a_{ii}+b_{ii})=\sum_{i=1}^{n} a_{ii}+\sum_{i=1}^{n} b_{ii}=\operatorname{tr} A+\operatorname{tr} B.$

(3) We have $AB=[c_{ij}]_{n\times n}$ where $c_{ij}=\sum_{k=1}^{n} a_{ik}\, b_{kj}$.

Also $BA=[d_{ij}]_{n\times n}$ where $d_{ij}=\sum_{k=1}^{n} b_{ik}\, a_{kj}$.

Now $\operatorname{tr}(AB)=\sum_{i=1}^{n} c_{ii}=\sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}\, b_{ki}\right)$

$=\sum_{k=1}^{n}\sum_{i=1}^{n} a_{ik}\, b_{ki}$, interchanging the order of summation in the last sum

$=\sum_{k=1}^{n}\left(\sum_{i=1}^{n} b_{ki}\, a_{ik}\right)=\sum_{k=1}^{n} d_{kk}$

$=d_{11}+d_{22}+\ldots+d_{nn}=\operatorname{tr}(BA).$

Theorem 16. *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 $C^{-1}A=D$.

Then $\operatorname{tr} B=\operatorname{tr}(DC)$

$=\operatorname{tr}(CD)$ [by theorem 15]

$=\operatorname{tr}(CC^{-1}A)=\operatorname{tr}(IA)=\operatorname{tr} A$

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

$$[T]_B \text{ and } [T]_{B'}$$

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 of 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.

Solved Examples

Example 1. *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*) $B'=\{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$.

(Andhra 1992, Nagarjuna 90; Tirupati 90)

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

$\therefore$ 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. $x=c$, $y=b-c$, $z=a-b$. ...(1)

Putting $a=3$, $b=-3$, and $c=3$ in (1), we get

$x=3$, $y=-6$ and $z=6$.

$\therefore \quad T(1, 1, 1)=(3, -3, 3)=3\,(1, 1, 1)-6\,(1, 1, 0)+6\,(1, 0, 0)$.

Also $T(1, 1, 0)=(2, -3, 3)$.

Putting $a=2$, $b=-3$ and $c=3$ in (1), we get

$T(1, 1, 0)=(2, -3, 3)=3\,(1, 1, 1)-6\,(1, 1, 0)+6\,(1, 0, 0)$.

Finally, $T(1, 0, 0)=(0, 1, 3)$.

Putting $a=0$, $b=1$ and $c=3$ in (1), we get

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

Scanned with CamScanner

$$\therefore \quad [T]_{B'}=\begin{bmatrix} 3 & 3 & 3 \\ -6 & -6 & -2 \\ 6 & 5 & -1 \end{bmatrix}.$$

Example 2. *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 $\alpha_1=(1, 0, 1)$, $\alpha_2=(-1, 2, 1)$ and $\alpha_3=(2, 1, 1)$?
(Meerut 1972, 77, 90, 91)

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 get
$$x=\frac{-a-3b+5c}{4},\ y=\frac{b+c-a}{4},\ z=\frac{b-c+a}{2}. \quad \ldots(1)$$
Putting $a=4$, $b=-2$, $c=3$ in (1), we get
$$x=\tfrac{17}{4},\ y=-\tfrac{3}{4},\ z=-\tfrac{1}{2}.$$
$$\therefore \quad T(\alpha_1)=\tfrac{17}{4}\alpha_1-\tfrac{3}{4}\alpha_2-\tfrac{1}{2}\alpha_3.$$
Also $T(\alpha_2)=T(-1, 2, 1)=(-2, 4, 9)$. Putting $a=-2$, $b=4$, $c=9$ in (1), we get $x=\frac{35}{4}$, $y=\frac{15}{4}$, $z=\frac{-7}{2}$.
$$\therefore \quad T(\alpha_2)=\frac{35}{4}\alpha_1+\frac{15}{4}\alpha_2-\frac{7}{2}\alpha_3.$$
Finally $T(\alpha_3)=T(2, 1, 1)=(7, -3, 4)$. Putting $a=7$, $b=-3$, $c=4$ in (1), we get $x=\frac{11}{2}$, $y=-\frac{3}{2}$, $z=0$.
$$\therefore \quad T(\alpha_3)=\frac{11}{2}\alpha_1-\frac{3}{2}\alpha_2+0\alpha_3.$$
$$\therefore \quad [T]_B=\begin{bmatrix} \frac{17}{4} & \frac{35}{4} & \frac{11}{2} \\ -\frac{3}{4} & \frac{15}{4} & -\frac{3}{2} \\ \frac{1}{2} & -\frac{7}{2} & 0 \end{bmatrix}.$$

Example 3. *Let T be a linear operator on* $\mathbf{R}^3$ *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}.* (Meerut 1976, 93)

Solution. Suppose B is the standard ordered basis for R^3. Then $B=\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. Let $A=[T]_B$ *i.e.* let A be the matrix of T with respect to B. First we shall compute A. We have
$$T(1, 0, 0)=(3, -2, -1),$$
$$T(0, 1, 0)=(0, 1, 2),$$
and
$$T(0, 0, 1)=(1, 0, 4).$$
$$\therefore \quad A=[T]_B=\begin{bmatrix} 3 & 0 & 1 \\ -2 & 1 & 0 \\ -1 & 2 & 4 \end{bmatrix}.$$
Now T will be invertible if the matrix $[T]_B$ is invertible. [See theorem 7 on page 165].

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
$$\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix},\ -\begin{vmatrix} 3 & 1 \\ -2 & 0 \end{vmatrix},\ \begin{vmatrix} 3 & 0 \\ -2 & 1 \end{vmatrix} \text{ i.e. } -1, -2, 3.$$
$$\therefore \quad \text{Adj. } A=\text{transpose of the matrix} \begin{bmatrix} 4 & 8 & -3 \\ 2 & 13 & -6 \\ -1 & -2 & 3 \end{bmatrix}$$
$$=\begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}.$$
$$\therefore \quad A^{-1}=\frac{1}{\det A}\text{ Adj. } A=\frac{1}{9}\begin{bmatrix} 4 & 2 & -1 \\ 8 & 13 & -2 \\ -3 & -6 & 3 \end{bmatrix}.$$
Now $[T^{-1}]_B=([T]_B)^{-1}=A^{-1}$. [See theorem 7 page 165]

We shall now find a formula for T^{-1}. Let $\alpha=(a, b, c)$ be any vector belonging to $\mathbf{R}^3$. Then

Scanned with CamScanner

$[T^{-1}(\alpha)]_B=[T^{-1}]_B\,[\alpha]_B$ [See Note on page 161]

$$=\frac{1}{9}\begin{bmatrix}4 & 2 & -1\\ 8 & 13 & -2\\ -3 & -6 & 3\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\frac{1}{9}\begin{bmatrix}4a+2b-c\\ 8a+13b-2c\\ -3a-6b+3c\end{bmatrix}.$$

Since B is the standard ordered basis for $\mathbf{R}^3$,

$$\therefore\ T^{-1}(\alpha)=T^{-1}(a, b, c)=\frac{1}{9}(4a+2b-c,\ 8a+13b-2c,\ -3a-6b+3c).$$

Example 4. *Let T be the linear operator on $\mathbf{R}^3$ 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 $\mathbf{R}^3$?

(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 Ex. 3]

(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 $[T]_{B'}=P^{-1}[T]_B P$. [See note on page 172]

In order to compute the matrix P^{-1}, we find that det $P=-4$.

Therefore $P^{-1}=\dfrac{1}{\det P}$ 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 Ex. 2].

Example 5. *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 $\alpha_1=(1, 1)$, $\alpha_2=(-1, 0)$. **(Meerut 1976, 93P)**

Solution. By def. of T, we have

$$T(\alpha_1)=T(1, 1)=(2, 3).$$

Now our aim is to express (2, 3) as a linear combination of the vectors in the basis $\{\alpha_1, \alpha_2\}$.

Let $(a, b)=x\alpha_1+y\alpha_2=x(1, 1)+y(-1, 0)=(x-y, x)$.

Then $x-y=a$, $x=b$.

Solving these equations, we get

$$x=b,\ y=b-a. \quad \text{...(1)}$$

Putting $a=2$, $b=3$ in (1), we get $x=3$, $y=1$.

$$\therefore\ T(\alpha_1)=3\alpha_1+1\alpha_2. \quad \text{...(2)}$$

Again $T(\alpha_2)=T(-1, 0)=(-4, -2)$. Putting $a=-4$, $b=-2$ in (1), we get $x=-2$, $y=2$.

$$\therefore\ T(\alpha_2)=-2\alpha_1+2\alpha_2. \quad \text{...(3)}$$

From the relations (2) and (3), we see that the matrix of T relative to the basis $\{\alpha_1, \alpha_2\}$ is $=\begin{bmatrix}3 & -2\\ 1 & 2\end{bmatrix}$.

Example 6. *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)}. **(Meerut 1980, 85, 89; S.V.U. Tirupati 93P)**

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

$$p=-5a+2b,\ q=3a-b. \quad \text{...(1)}$$

Scanned with CamScanner

Putting $a=6$, $b=0$ in (1), we get $p=-30$, $q=18$.

$\therefore \quad T(\alpha_1)=(6, 0)=-30\alpha_1+18\alpha_2$ ...(2)

Again putting $a=10$, $b=1$ in (1), we get

$p=-48$, $q=29$.

$\therefore \quad T(\alpha_2)=(10, 1)=-48\alpha_1+29\alpha_2$. ...(3)

From the relations (2) and (3), we see that the matrix of T relative to the basis $\{\alpha_1, \alpha_2\}$ is $=\begin{bmatrix} -30 & -48 \\ 18 & 29 \end{bmatrix}$

Example 7 *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'.

Solution. (i) We have $T(1, 0)=(4, 2)$ and $T(0, 1)=(-2, 1)$.

Since B is the standard ordered basis for $\mathbf{R}^2$, therefore

$$A=[T]_B=\begin{bmatrix} 4 & -2 \\ 2 & 1 \end{bmatrix}.$$

(ii) Since B is the standard ordered basis for $\mathbf{R}^2$, therefore the transition matrix P from B to B' can be immediately written as

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

Now $[T]_{B'}=P^{-1}[T]_B P$.

We have $\det P=1\times 0-(1\times -1)=1$. The cofactors of the elements of the first row of P are $0, -1$. Also the cofactors of the elements of the second row of P are $-(-1), 1$ *i.e.*, are $1, 1$. Therefore

$$P^{-1}=\frac{1}{\det P} \text{ Adj } P=\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}.$$

$$\therefore \quad [T]_{B'}=\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$$

$$=\begin{bmatrix} 2 & 1 \\ -2 & 3 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}=\begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix}.$$

Example 8. *Let T be the linear operator on $\mathbf{R}^3$ defined by*

$$T(x_1, x_2, x_3)=(x_1+x_2+x_3, -x_1-x_2-4x_3, 2x_1-x_3).$$

What is the matrix of T in the ordered basis $\{\alpha_1, \alpha_2, \alpha_3\}$ where $\alpha_1=(1, 1, 1)$, $\alpha_2=(0, 1, 1)$, $\alpha_3=(1, 0, 1)$?

Note. Calculate the required matrix in two ways and check your answer.

Example 9. *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 $$T(\alpha_4)=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$=0\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}+1\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}+0\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}+1\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$\therefore \quad [T]_B=\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Example 10. *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

$$[T]_B=\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

$\therefore \quad T(1, 0)=1\,(1, 0)+1\,(0, 1)=(1, 1)$,

and $\quad T(0, 1)=1\,(1, 0)+1\,(0, 1)=(1, 1)$.

Scanned with CamScanner

If $(a, b) \in V_2$ (C), then we can write

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

$$\therefore \quad T(a, b)=aT(1, 0)+bT(0, 1)$$
$$=a\,(1, 1)+b\,(1, 1)=(a+b, a+b).$$

This is the explicit expression for T.

Now let us find the matrix of T with respect to B'.

We have $\quad T(1, 1)=(2, 2)$.

Let $\quad (2, 2)=x\,(1, 1)+y\,(1, -1)=(x+y, x-y)$

Then $\quad x+y=2,\ x-y=2$

$$\Rightarrow x=2,\ y=0.$$

$\therefore \quad (2, 2)=2\,(1, 1)+0\,(1, -1)$.

Also $\quad T(1, -1)=(0, 0)=0\,(1, 1)+0\,(1, -1)$.

$$\therefore \quad [T]_{B'}=\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

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 11. *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* $\alpha_1, \alpha_2, \alpha_3$.

[Meerut 1981, 84P, 93P]

Solution. Let a, b, c be scalars *i.e.*, real numbers such that

$$a\alpha_1+b\alpha_2+c\alpha_3=0$$

i.e. $\quad a\,(1, 0, -1)+b\,(1, 2, 1)+c\,(0, -3, 2)=(0, 0, 0)$

i.e. $\quad (a+b+0c,\ 0a+2b-3c,\ -a+b+2c)=(0, 0, 0)$

i.e.
$$\left.\begin{array}{r} a+b+0c=0, \\ 0a+2b-3c=0, \\ -a+b+2c=0. \end{array}\right\} \qquad \ldots(1)$$

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
$$\alpha_1=(1, 0, -1)=1e_1+0e_2-1e_3$$
$$\alpha_2=(1, 2, 1)=1e_1+2e_2+1e_3$$
$$\alpha_3=(0, -3, 2)=0e_1-3e_2+2e_3.$$

If P is the transition matrix from the basis B to the basis B', then

$$P=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & -3 \\ -1 & 1 & 2 \end{bmatrix}.$$

Let us find the matrix P^{-1}. For this let us first find Adj. P.

The cofactors of the elements of the first row of P are

$$\begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix},\ -\begin{vmatrix} 0 & -3 \\ -1 & 2 \end{vmatrix},\ \begin{vmatrix} 0 & 2 \\ -1 & 1 \end{vmatrix} \quad i.e.\ 7, 3, 2.$$

The cofactors of the elements of the second row of P are

$$-\begin{vmatrix} 1 & 0 \\ 1 & 2 \end{vmatrix},\ \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix},\ -\begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} \quad i.e.\ -2, 2, -2.$$

The cofactors of the elements of the third row of P are

$$\begin{vmatrix} 1 & 0 \\ 2 & -3 \end{vmatrix},\ -\begin{vmatrix} 1 & 0 \\ 0 & -3 \end{vmatrix},\ \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} \quad i.e.\ -3, 3, 2.$$

$$\therefore \quad \text{Adj } P=\text{transpose of the matrix} \begin{bmatrix} 7 & 3 & 2 \\ -2 & 2 & -2 \\ -3 & 3 & 2 \end{bmatrix}$$

$$=\begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}.$$

$$\therefore \quad P^{-1}=\frac{1}{\det P}\text{ Adj } P=\frac{1}{10}\begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}.$$

Now $\quad e_1=1e_1+0e_2+0e_3$.

$\therefore$ 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'

$$=\Big[\ e_1\ \Big]_{B'}=P^{-1}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$=\frac{1}{10}\begin{bmatrix} 7 & -2 & -3 \\ 3 & 2 & 3 \\ 2 & -2 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Scanned with CamScanner

$$=\frac{1}{10}\begin{bmatrix}7\\3\\2\end{bmatrix}=\begin{bmatrix}7/10\\3/10\\2/10\end{bmatrix}.$$

$$\therefore\quad e_1=\frac{7}{10}\alpha_1+\frac{3}{10}\alpha_2+\frac{2}{10}\alpha_3.$$

Also $\left[e_2\right]_B=\begin{bmatrix}0\\1\\0\end{bmatrix}$ and $\left[e_3\right]_B=\begin{bmatrix}0\\0\\1\end{bmatrix}$.

$$\therefore\quad \left[e_2\right]_{B'}=P^{-1}\begin{bmatrix}0\\1\\0\end{bmatrix}\text{ and }\left[e_3\right]_{B'}=P^{-1}\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

Thus $\left[e_2\right]_{B'}=\frac{1}{20}\begin{bmatrix}-2\\2\\-2\end{bmatrix}$, $\left[e_3\right]_{B'}=\frac{1}{10}\begin{bmatrix}-3\\3\\2\end{bmatrix}$.

$$\therefore\quad e_2=-\tfrac{2}{10}\alpha_1+\tfrac{2}{10}\alpha_2-\tfrac{2}{10}\alpha_3$$

and $\quad e_3=-\tfrac{3}{10}\alpha_1+\tfrac{3}{10}\alpha_2+\tfrac{2}{10}\alpha_3.$

Example 12. *Let A be an $m\times n$ matrix with real entries. Prove that $A=0$ (null matrix) if and only if trace $(A' A)=0$.* **(Meerut 1981)**

Solution. Let $A=[a_{ij}]_{m\times n}$. Then $A'=[b_{ij}]_{n\times m}$, where $b_{ij}=a_{ji}$.

Now $A' A$ is a matrix of the type $n\times n$.

Let $A' A=[c_{ij}]_{n\times n}$. Then

$c_{ii}=$ the sum of the products of the corresponding elements of the i^{th} row of A' 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 trace $(A' A)=\sum_{i=1}^{n} c_{ii}$

$$=\sum_{i=1}^{n}\{a_{1i}^2+a_{2i}^2+\ldots+a_{mi}^2\}$$

$=$ the sum of the squares of all the elements of A.

Now the elements of A are all real numbers. Therefore trace $(A' A)=0\Rightarrow$ 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 $A' A$ is also a null matrix and so trace $(A' A)=0$.

Hence trace $(A' A)=0$ iff $A=0$.

Example 13. *Show that the only matrix similar to the identity matrix I is I itself.* **(Meerut 1976)**

Solution. The identity matrix I is invertible and we can write

$I=I^{-1}II$. Therefore I is similar to I. Further let B be a matrix similar to I. Then there exists an invertible matrix P such that

$$B=P^{-1}IP$$

$$\Rightarrow B=P^{-1}P\qquad[\because\ P^{-1}I=P^{-1}]$$

$$\Rightarrow B=I.$$

Hence the only matrix similar to I is I itself.

Example 14. *If two linear transformations A and B on $V(F)$ are similar, then show that A^2 and B^2 are also similar and if A, B are invertible, then A^{-1}, B^{-1} are also similar.*

Solution Since A and B are similar, therefore there exists an invertible linear transformation C on V such that

$$A=CBC^{-1}\qquad\ldots(1)$$

We have $A^2=AA=(CBC^{-1})(CBC^{-1})=CBC^{-1}CBC^{-1}$

$$=CBIBC^{-1}=CBBC^{-1}=CB^2C^{-1}.$$

$\therefore\ A^2$ is similar to B^2.

If A and B are invertible, then from (1), we have

$$A^{-1}=(CBC^{-1})^{-1}=(C^{-1})^{-1}B^{-1}C^{-1}=CB^{-1}C^{-1}.$$

$\therefore\ A^{-1}$ is similar to B^{-1}.

Example 15. *If A and B are linear transformations on the same vector space and if at least one of them is invertible, then AB and BA are similar.*

Solution. Let A be invertible.

We have $A(BA)A^{-1}=ABAA^{-1}=ABI=AB$.

Thus $\qquad AB=A(BA)A^{-1}$.

$\therefore\ AB$ is similar to BA.

Now let B be invertible.

We have $B(AB)B^{-1}=BABB^{-1}=BAI=BA$.

$\therefore\ BA$ is similar to AB.

Example 16. *Let T and S be linear operators on the finite dimensional vector space $V(F)$, prove that*

(*i*) *det* $(TS)=($*det* $T)$ $($*det* $S)$;

(*ii*) *T is invertible iff det $T\neq 0$.*

Solution. (i) Let B be any ordered basis for V.

We have $[TS]_B=[T]_B\,[S]_B$.

$\therefore$ det $[TS]_B=$ det $([T]_B\,[S]_B)=($det $[T]_B)$ $($det $[S]_B)$

[$\because$ determinant of the product of two matrices is equal to the product of their determinants].

Now the determinant of a linear transformation is equal to the determinant of its matrix with respect to any ordered basis.

$\therefore$ det $(TS)=($det $T)$ $($det $S)$.

Scanned with CamScanner

(ii) Suppose T is invertible. Then there exists a linear transformation T^{-1} on V such that $T^{-1}T=I=TT^{-1}$.

$\therefore$ det $(TT^{-1})=$ det I

$\Rightarrow$ (det T) (det T^{-1})$=$det $[I]_B$, where B is any ordered basis for V

$\Rightarrow$ (det T) (det T^{-1})$=1$. [$\because$ $[I]_B$ is unit matrix and the determinant of a unit matrix is equal to 1].

Now det T and det T^{-1} are elements of F. In a field the product of two elements can be 0 iff at least one of them is 0.

$\therefore$ (det T) (det T^{-1})$=1$

$\Rightarrow$ det $T\neq 0$.

Conversely suppose that det $T\neq 0$.

Then det $[T]_B\neq 0$, where B is any ordered basis for V.

Now det $[T]_B\neq 0$ implies that the matrix $[T]_B$ is invertible.

$\therefore$ T is also invertible.

Example 17. *If T and S are linear transformations on a finite dimensional vector space V such that*

$$TS=\hat{0},\ T\neq\hat{0},\ S\neq\hat{0} \text{ then } \det T=\det S=0.$$

Solution. Let det $T\neq 0$.

Then T is invertible and T^{-1} exists.

$$\therefore\quad TS=\hat{0}$$

$\Rightarrow T^{-1}(TS)=T^{-1}\hat{0}$

$\Rightarrow (T^{-1}T)\,S=\hat{0} \Rightarrow IS=\hat{0}$

$\Rightarrow S=\hat{0}$ which is contradictory to the hypothesis that $S\neq\hat{0}$

$\therefore$ det T must be equal to 0.

Again let det $S\neq 0$. Then S is invertible.

$$\therefore\quad TS=\hat{0}$$

$\Rightarrow (TS)\,S^{-1}=\hat{0}\,S^{-1}$

$\Rightarrow T=\hat{0}$ which is contradictory to the hypothesis that $T\neq\hat{0}$.

$\therefore$ det S must be equal to zero.

Example 18. *If $\{\alpha_1, \ldots, \alpha_n\}$ and $\{\beta_1, \ldots, \beta_n\}$ are bases in the same finite dimensional vector space $V(F)$, and if T is a linear transformation such that*

$$T(\alpha_i)=\beta_i,\ i=1, \ldots, n, \text{ then } \det T\neq 0.$$

Solution. Since T maps a basis for V onto a basis for V, therefore T is invertible.

Now T is invertible implies that det $T\neq 0$. For proof see Ex. 16.

Example 19. *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.

Exercises

1. 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'.

Ans. $[T]_B=\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}$; $P=\begin{bmatrix}1 & 2\\ 1 & 1\end{bmatrix}$;

$$[T]_{B'}=\begin{bmatrix}-1 & -2\\ 1 & 2\end{bmatrix}.$$

2. Find the matrix relative to the basis

$\alpha_1=(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$, $\alpha_2=(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3})$, $\alpha_3'=(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$ 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

$$\begin{bmatrix}2 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & 3\end{bmatrix}.$$

Ans. $$\begin{bmatrix}3 & -\frac{2}{3} & -\frac{2}{3}\\ -\frac{2}{3} & \frac{10}{3} & 0\\ -\frac{2}{3} & 0 & \frac{8}{3}\end{bmatrix}.$$

3. Find the co-ordinates of the vector $(2, 1, 3, 4)$ of $\mathbf{R}^4$ relative to the basis vectors

$\alpha_1=(1, 1, 0, 0)$, $\alpha_2=(1, 0, 1, 1)$, $\alpha_3=(2, 0, 0, 2)$, $\alpha_4=(0, 0, 2, 2)$.

Ans. $(2, 1, 3, 4)=\alpha_1+\frac{1}{2}\alpha_3+\frac{3}{2}\alpha_4$.

4. Explain what is meant by the matrix of a linear transformation on V relative to a basis of V. Let F be a field and V, the set of all polynomials in x over F of degree $\leqslant 5$. If $D: V\to V$

Scanned with CamScanner

is defined by $D[f(x)]=f''(x)$ where $f'(x)$ is the derivative of $f(x)$, show that D is a linear transformation on V. Find the matrix of D in the basis $\{1, x, x^2, x^3, x^4\}$.

Ans. $\begin{bmatrix} 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 0 & 12 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

5. Let V be the vector space of those polynomial functions from the reals into itself which have degree $\leqslant 3$. Let
$$B=\{f_1, f_2, f_3, f_4\} \text{ where } f_i(x)=x^{i-1} \ (1 \leqslant i \leqslant 4).$$
Show that B forms a basis for V. For any real number t let $g_i(x)=(x+t)^{i-1}$. Show that $B'=\{g_1, g_2, g_3, g_4\}$ is also a basis for V. If D is the differentiation operator on V, write the matrices of D in the ordered bases B and B'. **(Meerut 1974)**

Ans. $[D]_B=[D]_{B'}=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

6. If A and B are $n \times n$ complex matrices, show that $AB-BA=I$ is impossible.

7. 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
$$\text{trace}\,(T)=2\ \text{trace}\,(P)$$

8. Show that the only matrix similar to the zero matrix is the zero matrix itself. **(Meerut 1976)**

§ 14. Linear Functionals. 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.**

Example 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
$$f(\alpha)=x_1a_1+x_2a_2+\ldots+x_na_n.$$
Let $\beta=(b_1, b_2, \ldots, b_n) \in V_n(F)$. If $a, b \in F$, we have
$$\begin{aligned} f(a\alpha+b\beta) &= f[a(a_1, \ldots, a_n)+b(b_1, \ldots, b_n)] \\ &= f(aa_1+bb_1, \ldots, aa_n+bb_n) \\ &= x_1(aa_1+bb_1)+\ldots+x_n(aa_n+bb_n) \\ &= a(x_1a_1+\ldots+x_na_n)+b(x_1b_1+\ldots+x_nb_n) \\ &= af(a_1, \ldots, a_n)+bf(b_1, \ldots, b_n) \\ &= af(\alpha)+bf(\beta). \end{aligned}$$
$\therefore$ f is a linear functional on $V_n(F)$.

Example 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.* **(Meerut 1977)**

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
$$\text{tr}\,(aA+bB)=\text{tr}\,(a[a_{ij}]_{n \times n}+b[b_{ij}]_{n \times n})=\text{tr}\,([aa_{ij}+bb_{ij}]_{n \times n})$$
$$=\sum_{i=1}^{n}(aa_{ii}+bb_{ii})=a\sum_{i=1}^{n} a_{ii}+b\sum_{i=1}^{n} b_{ii}=a\,(\text{tr}\, A)+b\,(\text{tr}\, B).$$

Example 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
$$f_i(\beta)=b_i.$$

Scanned with CamScanner

If $a, b \in F$, we have

$$f_i(a\alpha+b\beta)=f_i[a(a_1\alpha_1+\ldots+a_n\alpha_n)+b(b_1\alpha_1+\ldots+b_n\alpha_n)]$$
$$=f_i[(aa_1+bb_1)\alpha_1+\ldots+(aa_n+bb_n)\alpha_n]$$
$$=aa_i+bb_i=af_i(\alpha)+bf_i(\beta).$$

Hence f_i is a linear functional on V.

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 $\alpha, \beta \in V$ and $a, b \in F$. We have

$$f(a\alpha+b\beta)=0 \quad \text{(by def. of } f)$$
$$=a0+b0=af(\alpha)+bf(\beta).$$

$\therefore$ 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.

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

$$(-f)(a\alpha+b\beta)=-[f(a\alpha+b\beta)] \quad \text{[by def. of } -f]$$
$$=-[af(\alpha)+bf(\beta)] \quad [\because f \text{ is a linear functional}]$$
$$=a[-f(\alpha)]+b[-f(\beta)]$$
$$=a[(-f)\alpha]+b[(-f)(\beta)].$$

$\therefore$ $-f$ is a linear functional on V.

Properties of a linear functional.

Theorem. *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. Let $\alpha\in V$. Then $f(\alpha)\in F$.

We have
$$f(\alpha)+0=f(\alpha) \quad [\because 0 \text{ is zero element of } F]$$
$$=f(\alpha+\mathbf{0}) \quad [\because \mathbf{0} \text{ is zero element of } V]$$
$$=f(\alpha)+f(\mathbf{0}) \quad [\because f \text{ is a linear functional}]$$

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.

(iii) We have $f[\alpha+(-\alpha)]=f(\alpha)+f(-\alpha)$ $[\because f$ is a linear functional]

But $f[\alpha+(-\alpha)]=f(\mathbf{0})=0$ [by (i)]

Thus in F, we have

$$f(\alpha)+f(-\alpha)=0$$
$$\Rightarrow f(-\alpha)=-f(\alpha).$$

§ 15. Dual Spaces.

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 :

$$(f_1+f_2)(\alpha)=f_1(\alpha)+f_2(\alpha) \ \forall\ \alpha\in V. \quad \ldots(1)$$

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

$$(f_1+f_2)(a\alpha+b\beta)=f_1(a\alpha+b\beta)+f_2(a\alpha+b\beta) \quad \text{[by (1)]}$$
$$=[af_1(\alpha)+bf_1(\beta)]+[af_2(\alpha)+bf_2(\beta)]$$
$$[\because f_1 \text{ and } f_2 \text{ are linear functionals}]$$
$$=a[f_1(\alpha)+f_2(\alpha)]+b[f_1(\beta)+f_2(\beta)]$$
$$=a[(f_1+f_2)(\alpha)]+b[(f_1+f_2)(\beta)] \quad \text{[by (1)]}$$

$\therefore$ 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 :

$$(cf)(\alpha)=cf(\alpha) \ \forall\ \alpha \in V. \quad \ldots(2)$$

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

Scanned with CamScanner

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

$(cf)(a\alpha+b\beta)=cf(a\alpha+b\beta)$ [by (2)]

$=c\,[af(\alpha)+bf(\beta)]$ [$\because$ f is linear functional]

$=c\,[af(\alpha)]+c\,[bf(\beta)]$ [$\because$ F is a field]

$=(ca)f(\alpha)+(cb)f(\beta)$

$=(ac)f(\alpha)+(bc)f(\beta)$

$=a\,[cf(\alpha)]+b\,[cf(\beta)]$

$=a\,[(cf)(\alpha)]+b\,[(cf)(\beta)]$.

$\therefore$ cf is a linear functional on V. Thus

$f\in V'$ 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

$[f_1+(f_2+f_3)](\alpha)=f_1(\alpha)+(f_2+f_3)(\alpha)$ [by (1)]

$=f_1(\alpha)+[f_2(\alpha)+f_3(\alpha)]$ [by (1)]

$=[f_1(\alpha)+f_2(\alpha)]+f_3(\alpha)$ [$\because$ addition in F is associative]

$=(f_1+f_2)(\alpha)+f_3(\alpha)$ [by (1)]

$=[(f_1+f_2)+f_3](\alpha)$ [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

$(f_1+f_2)(\alpha)=f_1(\alpha)+f_2(\alpha)$ [by (1)]

$=f_2(\alpha)+f_1(\alpha)$ [$\because$ addition in F is commutative]

$=(f_2+f_1)(\alpha)$ [by (1)]

$\therefore$ $f_1+f_2=f_2+f_1$.

Existence of additive identity in V'. Let $\hat{0}$ be the zero linear functional on V *i.e.*

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

Then $\hat{0}\in V'$. If $f\in V'$ and $\alpha\in V$, we have

$(\hat{0}+f)(\alpha)=\hat{0}(\alpha)+f(\alpha)$ [by (1)]

$=0+f(\alpha)$ [by def. of $\hat{0}$]

$=f(\alpha)$ [0 being additive identity in F]

$\therefore$ $\hat{0}+f=f \;\forall\; f\in V'$.

$\therefore$ $\hat{0}$ 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

$(-f+f)(\alpha)=(-f)(\alpha)+f(\alpha)$ **[by (1)]**

$=-f(\alpha)+f(\alpha)$ **[by def. of $-f$]**

$=0$

$=\hat{0}(\alpha)$ **[by def. of $\hat{0}$]**

$\therefore$ $-f+f=\hat{0}$ 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

$[c(f_1+f_2)](\alpha)=c\,[(f_1+f_2)(\alpha)]$ [by (2)]

$=c\,[f_1(\alpha)+f_2(\alpha)]$ [by (1)]

$=cf_1(\alpha)+cf_2(\alpha)$

$=(cf_1)(\alpha)+(cf_2)(\alpha)$ [by (2)]

$=(cf_1+cf_2)(\alpha)$ [by (1)]

$\therefore$ $c(f_1+f_2)=cf_1+cf_2$.

(ii) Let $a, b\in F$ and $f\in V'$. If $\alpha\in V$, we have

$[(a+b)f](\alpha)=(a+b)f(\alpha)$ [by (2)]

$=af(\alpha)+bf(\alpha)$ [$\because$ F is a field]

$=(af)(\alpha)+(bf)(\alpha)$ [by (2)]

$=(af+bf)(\alpha)$ [by (1)]

$\therefore$ $(a+b)f=af+bf$.

(iii) Let $a, b\in F$ and $f\in V'$. If $\alpha\in V$, we have

$[(ab)f](\alpha)=(ab)f(\alpha)$ [by (2)]

$=a\,[bf(\alpha)]$ [$\because$ multiplication in F is associative]

$=a\,[(bf)(\alpha)]$ [by (2)]

$=[a(bf)]$ [illegible] [by (2)]

$\therefore$ $(ab)f=$ [illegible]).

(iv) Let 1 be the multiplicative identity of F and $f\in V'$. If $\alpha\in V$, we have

$(1f)(\alpha)=1f(\alpha)$ [by (2)]

$=f(\alpha)$ [$\because$ F is a field]

$\therefore$ $1f=f$.

Hence V' is a vector space over the field F.

Scanned with CamScanner

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.* **(Meerut 1970, 72, 83; Allahabad 78)**

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.

§ 16. Dual bases.

Theorem 1. *Let V be an n-dimensional vector space over the field F and let $B=\{\alpha_1,...,\alpha_n\}$ be an ordered basis for V. If $\{x_1,...,x_n\}$ is any ordered set of n scalars, then there exists a unique linear functional f on V such that $f(\alpha_i)=x_i$, $i=1, 2,..., n$.*

Proof. Existence of f. Let $\alpha \in V$.

Since $B=\{\alpha_1, \alpha_2,..., \alpha_n\}$ is a basis for V, therefore there exist unique scalars $a_1, a_2,..., a_n$ such that

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

For this vector α, let us define

$$f(\alpha)=a_1x_1+...+a_nx_n.$$

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

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

Therefore according to our definition of f, we have

$$f(\alpha_i)=0x_1+0x_2+...+1x_i+0x_{i+1}+...+0x_n$$

i.e. $$f(\alpha_i)=x_i,\ i=1, 2,..., n.$$

Now to show that f is a linear functional.

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

$$\alpha=a_1\alpha_1+...+a_n\alpha_n,$$

and $$\beta=b_1\alpha_1+...+b_n\alpha_n.$$ Then

$$f(a\alpha+b\beta)=f[a(a_1\alpha_1+...+a_n\alpha_n)+b(b_1\alpha_1+...+b_n\alpha_n)]$$
$$=f[(aa_1+bb_1)\alpha_1+...+(aa_n+bb_n)\alpha_n]$$
$$=(aa_1+bb_1)x_1+...+(aa_n+bb_n)x_n \quad \text{[by def. of } f]$$
$$=a(a_1x_1+...+a_nx_n)+b(b_1x_1+...+b_nx_n)=af(\alpha)+bf(\beta).$$

$\therefore$ f is a linear functional on V. Thus there exists a linear functional f on V such that $f(\alpha_i)=x_i$, $i=1, 2,..., n$.

Uniqueness of f. Let g be a linear functional on V such that

$$g(\alpha_i)=x_i,\ i=1, 2,..., n.$$

For any vector $\alpha=a_1\alpha_1+...+a_n\alpha_n \in V$, we have

$$g(\alpha)=g(a_1\alpha_1+...+a_n\alpha_n)$$
$$=a_1g(\alpha_1)+...+a_ng(\alpha_n) \quad [\because g \text{ is linear}]$$
$$=a_1x_1+...+a_nx_n \quad \text{[by def. of } g]$$
$$=f(\alpha). \quad \text{[by def. of } f]$$

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 $f(\alpha)=g(\alpha)\ \forall\ \alpha \in V$ *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 $B=\{\alpha_1,..., \alpha_n\}$ be a basis for V. Then there is a uniquely determined basis $B'=\{f_1,..., f_n\}$ for V' such that $f_i(\alpha_j)=\delta_{ij}$. Consequently the dual space of an n-dimensional space is n-dimensional.* **(Meerut 1974, 77, 88, 93; Poona 70; Allahabad 78)**

The basis B' is called the **dual basis** of B.

Proof. $B=\{\alpha_1,..., \alpha_n\}$ is an ordered basis for V. Therefore by theorem 1, there exists a unique linear functional f_1 on V such that

$$f_1(\alpha_1)=1, f_1(\alpha_2)=0,..., f_1(\alpha_n)=0$$

where $\{1, 0,..., 0\}$ is an ordered set of n scalars.

In fact, for each $i=1, 2,..., n$ 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. $$f_i(\alpha_j)=\delta_{ij}, \quad ...(1)$$

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 $B'=\{f_1,..., f_n\}$. 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.

Scanned with CamScanner

Let $c_1 f_1+c_2 f_2+\ldots+c_n f_n=\hat{0}$

$\Rightarrow (c_1 f_1+\ldots+c_n f_n)(\alpha)=\hat{0}(\alpha) \ \forall \ \alpha\in V$

$\Rightarrow c_1 f_1(\alpha)+\ldots+c_n f_n(\alpha)=0 \ \forall \ \alpha\in V$ [$\because \ \hat{0}(\alpha)=0$]

$\Rightarrow \sum_{i=1}^{n} c_i f_i(\alpha)=0 \ \forall \ \alpha\in V$

$\Rightarrow \sum_{i=1}^{n} c_i f_i(\alpha_j)=0, j=1, 2,\ldots, n$

[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

$$\left[\sum_{i=1}^{n} a_i f_i\right](\alpha_j)=\sum_{i=1}^{n} a_i f_i(\alpha_j)$$

$$=\sum_{i=1}^{n} a_i \delta_{ij} \qquad \text{[from (1)]}$$

$=a_j$, on summing with respect to i and remembering that $\delta_{ij}=1$ when $i=j$ and $\delta_{ij}=0$ when $i\neq j$

$=f(\alpha_j)$ [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. *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

$$\alpha=\sum_{i=1}^{n} f_i(\alpha)\,\alpha_i. \qquad \textbf{(Meerut 1972, 79, 85)}$$

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 $f(\alpha_j)=\left(\sum_{i=1}^{n} c_i f_i\right)(\alpha_j)=\sum_{i=1}^{n} c_i f_i(\alpha_j)$

$$=\sum_{i=1}^{n} c_i \delta_{ij} \qquad \text{[From (1)]}$$

$=c_j, j=1, 2,\ldots, n.$

$\therefore \ f=\sum_{i=1}^{n} f(\alpha_i) f_i.$

Now let α be any vector in V. Let

$$\alpha=x_1\alpha_1+\ldots+x_n\alpha_n. \qquad \ldots(2)$$

Then $f_i(\alpha)=f_i\left(\sum_{j=1}^{n} x_j \alpha_j\right)$ $\left[\text{From (2), } \alpha=\sum_{j=1}^{n} x_j \alpha_j\right]$

$$= \sum_{j=1}^{n} x_j f_i(\alpha_j) \qquad [\because f_i \text{ is linear functional}]$$

$$= \sum_{j=1}^{n} x_j \delta_{ij} \qquad [\text{From (1)}]$$

$$= x_i.$$

$\therefore \quad \alpha = f_1(\alpha)\,\alpha_1 + \ldots + f_n(\alpha)\,\alpha_n = \sum_{i=1}^{n} f_i(\alpha)\,\alpha_i.$

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 $f(\alpha) \neq 0$.* **(Poona 1970)**

Proof Since $\alpha \neq 0$, 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 $f(\alpha)=0 \;\forall\; f \in V'$, 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

$$f(\alpha-\beta) \neq 0$$
$$\Rightarrow f(\alpha)-f(\beta) \neq 0$$
$$\Rightarrow f(\alpha) \neq f(\beta).$$

Hence the result.

§ 17. Reflexivity.

Second 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** 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.

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 $\alpha \to L_\alpha$ is an isomorphism of V onto V''.

(Meerut 1973, 76, 77, 78, 82, 83; 88, 90, 93P; Kanpur 69)

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 $\quad a, b \in F$ and $f, g \in V'$. Then

$L_\alpha(af+bg)=(af+bg)(\alpha) \qquad$ [From (1)]

$=(af)(\alpha)+(bg)(\alpha)$

$=af(\alpha)+bg(\alpha)$ [by scalar multiplication of linear functionals]

$=a\,[L_\alpha(f)]+b\,[L_\alpha(g)]. \qquad$ [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

$\psi(\alpha)=\psi(\beta)$

$\Rightarrow L_\alpha=L_\beta \Rightarrow L_\alpha(f)=L_\beta(f) \;\forall\; f \in V'$

$\Rightarrow f(\alpha)=f(\beta) \;\forall\; f \in V' \qquad$ [From (1)]

$\Rightarrow f(\alpha)-f(\beta)=0 \;\forall\; f \in V' \Rightarrow f(\alpha-\beta)=0 \;\forall\; f \in V'$

$\Rightarrow \alpha-\beta=\mathbf{0} \qquad$ [$\because$ by theorem 4 of § 16, 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 $\mathbf{0}$]

$\Rightarrow \alpha=\beta.$

$\therefore \quad \psi$ is one-one.

ψ is a linear transformation.

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

Scanned with CamScanner

$$\psi(a\alpha+b\beta)=L_{a\alpha+b\beta}$$ [by def. of ψ]

For every $f\in V'$, we have

$L_{a\alpha+b\beta}(f)=f(a\alpha+b\beta)$

$=af(\alpha)+bf(\beta)$ [From (1)]

$=aL_\alpha(f)+bL_\beta(f)$

$=(aL_\alpha)(f)+(bL_\beta)(f)=(aL_\alpha+bL_\beta)(f)$. [From (1)]

$\therefore\quad L_{a\alpha+b\beta}=aL_\alpha+bL_\beta=a\psi(\alpha)+b\psi(\beta)$

Thus $\psi(a\alpha+b\beta)=a\psi(\alpha)+b\psi(\beta)$.

$\because\quad \psi$ 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** betweed 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

$$L_i(f_j)=\delta_{ij}. \quad \ldots(1)$$

By previous theorem, for each i there is a vector α_i in V such that $L_i=L_{\alpha_i}$ where $L_{\alpha_i}(f)=f(\alpha_i)\ \forall\ f\in V'$. ...(2)

The correspondence $\alpha\leftrightarrow L_\alpha$ is an isomorphism of V onto V''. Under an isomorphism of 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

$f_j(\alpha_i)=L_{\alpha_i}(f_j)=L_i(f_j)$

$=\delta_{ij}$. [From (1)]

$\therefore\quad B'=\{f_1,\ldots,f_n\}$ 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* $B''=(B')'=B$.

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 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 $B''=B$.

Solved Examples

Example 1. *Find the dual basis of the basis set*

$$B=\{(1,-1,3),(0,1,-1),(0,3,-2)\}$$

for $V_3(\mathbf{R})$,

Solution. Let $\alpha_1=(1,-1,3)$, $\alpha_2=(0,1,-1)$, $\alpha_3=(0,3,-2)$.

Then $$B=\{\alpha_1,\alpha_2,\alpha_3\}.$$

Scanned with CamScanner

If $B'=\{f_1, f_2, f_3\}$ is dual basis of B, then

$$f_1(\alpha_1)=1, f_1(\alpha_2)=0, f_1(\alpha_3)=0,$$
$$f_2(\alpha_1)=0, f_2(\alpha_2)=0, f_2(\alpha_3)=0,$$
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 $$(a, b, c)=x\,(1, -1, 3)+y\,(0, 1, -1)+z\,(0, 3, -2) \quad \ldots(1)$$
$$=x\alpha_1+y\alpha_2+z\alpha_3$$

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
$$x=(xy^{-1})\,y=(xy^{-1})\,f(\beta)$$
$$=f[(xy^{-1})\,\beta] \qquad [\because\ f \text{ is linear functional}]$$

Thus there exists $\alpha=(xy^{-1})\,\beta\in V$ such that $f(\alpha)=x$.

Note. If f is a non-zero linear functional on $V(F)$, then f is necessarily a function from V onto F.

Important Note. 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 $0\in N$. 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 $\alpha, \beta\in N \Rightarrow a\alpha+b\beta\in N$.

$\therefore$ 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
$$\dim V=\dim N+0$$
$$\Rightarrow n=\dim N.$$

(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^1 is 1.

$\therefore$ In this case we have
$$\dim V=\dim N+1$$
$$\Rightarrow \dim N=n-1.$$

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 $f(\alpha)=x$.

Scanned with CamScanner

We have $f(\alpha)=x$

$\Rightarrow f(\alpha)=(xy^{-1})\,y$ [$\because$ $0\neq y\in F \Rightarrow y^{-1}$ exists]

$\Rightarrow f(\alpha)=cy$ where $c=xy^{-1}\in F$

$\Rightarrow f(\alpha)=c\,f(\alpha_0)$

$\Rightarrow f(\alpha)=f(c\alpha_0)$ [$\because$ f is a linear functional]

$\Rightarrow f(\alpha)-f(c\alpha_0)=0$

$\Rightarrow f(\alpha-c\alpha_0)=0$

$\Rightarrow \alpha-c\alpha_0\in N$

$\Rightarrow \alpha-c\alpha_0=\beta$ for some $\beta\in N$

$\Rightarrow \alpha=c\alpha_0+\beta.$

If possible, let

$\alpha=c'\alpha_0+\beta'$ where $c'\in F$ and $\beta'\in N$.

Then $c\alpha_0+\beta=c'\alpha_0+\beta'$

$\Rightarrow (c-c')\,\alpha_0+(\beta-\beta')=0$ ...(1)

$\Rightarrow f\,[(c-c')\,\alpha_0+(\beta-\beta')]=f(0)$

$\Rightarrow (c-c')\,f(\alpha_0)+f\,(\beta-\beta')=0$

$\Rightarrow (c-c')\,f(\alpha_0)=0$ [$\because$ $\beta,\beta'\in N\Rightarrow\beta-\beta'\in N$ and thus $f(\beta-\beta')=0$]

$\Rightarrow (c-c')=0$ [$\because$ $f(\alpha_0)$ is a non-zero element of F]

$\Rightarrow c=c'$.

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 $k\in F$.* **(Meerut 1979, 84)**

Solution. It is given that $f(\alpha)=0\Rightarrow g(\alpha)=0$. Therefore if α belongs to null space of f, then α also belongs to 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

$$g=kf\ \forall\ k\in F.$$

(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 $$k=\frac{g(\alpha_0)}{f(\alpha_0)}.$$

If $\alpha\in V$, then we can write

$\alpha=c\alpha_0+\beta$ where $c\in F$ and $\beta\in$ null space of f.

We have $g(\alpha)=g\,(c\alpha_0+\beta)=cg\,(\alpha_0)+g\,(\beta)$

$=cg\,(\alpha_0)$

[$\because$ $\beta\in$ null space of $f\Rightarrow f(\beta)=0$ and so $g(\beta)=0$]

Also $(kf)\,(\alpha)=kf(\alpha)=kf(c\alpha_0+\beta)$

$=k\,[cf\,(\alpha_0)+f\,(\beta)]$

$=kc\,f(\alpha_0)$ [$\because$ $f(\beta)=0$]

$=\dfrac{g(\alpha_0)}{f(\alpha_0)}\,cf(\alpha_0)=cg(\alpha_0)$.

Thus $g(\alpha)=(kf)\,(\alpha)\ \forall\ \alpha\in V$.

$\therefore\ g=kf$.

Exercises

1. 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. **(Meerut 1973)**

2. Find the dual basis of the basis set
$$B=\{(1,0,0),(0,1,0),(0,0,1)\}$$
for $V_3(\mathbf{R})$.
Ans. $B'=\{f_1,f_2,f_3\}$
where $f_1(a,b,c)=a,\ f_2(a,b,c)=b,\ f_3(a,b,c)=c$.

3. Find the dual basis of the basis set
$$B=\{(1,-2,3),(1,-1,1),(2,-4,7)\}\text{ of }V_3(\mathbf{R}).$$
Ans. $B'=\{f_1,f_2,f_3\}$
where $f_1(a,b,c)=-3a-5b-2c$,
$f_2(a,b,c)=2a+b,\quad f_3(a,b,c)=a+2b+c$.

§ 18. Annihilators.

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*

$$f(\alpha)=0\ \forall\ \alpha\in S.$$

(Meerut 1970, 71, 76, 92; Marathwada 71; S.V.U. Tirupati 90)

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 $S^0=V'$. **(Meerut 1976)**

If $S=V$, then $S^0=V^0=$zero subspace of V'. **(Meerut 1976)**

Scanned with CamScanner

If V is finite dimensional and S contains a non-zero vector, then $S^0 \neq V'$. If $0 \neq \alpha \in S$, 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'.* **(Poona 1970; Meerut 79, 89, 92; Kanpur 81)**

Proof. First we see that S^0 is a non-empty subset of V' because at least $\hat{0} \in S^0$. We have

$$\hat{0}(\alpha) = 0 \ \forall \ \alpha \in S.$$

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

$(af + bg)(\alpha) = (af)(\alpha) + (bg)(\alpha) = af(\alpha) + bg(\alpha) = a0 + b0 = 0.$

$\therefore \quad af + bg \in S^0.$

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*

$$\dim W + \dim W^0 = \dim V.$$

(Meerut 1980, 81, 82, 84, 87, 91, 92; Marathwada 71, 93; S.V U. Tirupati 90)

Proof. If W is zero subspace of V, then $W^0 = V'$.

$\therefore \quad \dim W^0 = \dim V' = \dim V.$

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 $f \in W^0 \Rightarrow f(\alpha) = 0 \ \forall \ \alpha \in W$

$\Rightarrow f(\alpha_j) = 0$ for each $j = 1, \ldots, m$ $\quad [\because \ \alpha_1, \ldots, \alpha_m$ are in $W]$

$\Rightarrow \left(\sum_{i=1}^{n} x_i f_i\right)(\alpha_j) = 0$ $\quad$ [From (1)]

$\Rightarrow \sum_{i=1}^{n} x_i f_i(\alpha_j) = 0 \Rightarrow \sum_{i=1}^{n} x_i \delta_{ij} = 0$

$\Rightarrow x_j = 0$ for each $j = 1, \ldots, m$.

Putting $x_1 = 0, x_2 = 0, \ldots, x_m = 0$ in (1), we get

$f = x_{m+1} f_{m+1} + \ldots + x_n f_n$

$=$ a linear combination of the elements of S.

$\therefore \quad f \in L(S).$

Thus $\quad f \in W^0 \Rightarrow f \in L(S)$.

$\therefore \quad W^0 \subseteq L(S)$.

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 $g(\alpha) = g\left(\sum_{j=1}^{m} c_j \alpha_j\right)$ $\quad$ [From (3)]

$= \sum_{j=1}^{m} c_j \, g(\alpha_j)$ $\quad [\because \ g$ is linear functional]

$= \sum_{j=1}^{m} c_j \left(\sum_{k=m+1}^{n} y_k f_k\right)(\alpha_j)$ $\quad$ [From (2)]

$= \sum_{j=1}^{m} c_j \sum_{k=m+1}^{n} y_k f_k(\alpha_j) = \sum_{j=1}^{m} c_j \sum_{k=m+1}^{n} y_k \delta_{kj},$

$= \sum_{j=1}^{m} c_j\, 0$ [$\because$ $\delta_{kj}=0$ if $k \neq j$ which is so for each $k=m+1,\ldots,n$ and for each $j=1,\ldots,m$]

$=0.$

Thus $g(\alpha)=0 \ \forall \ \alpha \in W$. Therefore $g \in W^0$.

Thus $g \in L(S) \Rightarrow g \in W^0$.

$\therefore \quad L(S) \subseteq W^0$.

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 $\dim W' = \dim W = m$.

Now $\dim V'/W^0 = \dim V' - \dim W^0$

$= \dim V - (\dim V - \dim W) = \dim W = m.$

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 $W^{00}=W$.*

(Meerut 1968, 78, 90, 91)

Proof. We have

$$W^0 = \{f \in V' : f(\alpha)=0 \ \forall \ \alpha \in W\} \quad \ldots(1)$$

and $$W^{00} = \{\alpha \in V : f(\alpha)=0 \ \forall \ f \in W^0\}. \quad \ldots(2)$$

Let $\alpha \in W$. Then from (1), $f(\alpha)=0 \ \forall \ f \in W^0$ and so from (2), $\alpha \in W^{00}$.

$\therefore \quad \alpha \in W \Rightarrow \alpha \in W^{00}$.

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 $\dim W + \dim W^0 = \dim V$. [by theorem (2)]

Applying the same theorem for vector space V' and its subspace W^0, we get

$\dim W^0 + \dim W^{00} = \dim V' = \dim V.$

$\therefore \quad \dim W = \dim V - \dim W^0 = \dim V - [\dim V - \dim W^{00}]$

$= \dim W^{00}.$

Since W is a subspace of W^{00} and $\dim W = \dim W^{00}$, therefore $W = W^{00}$.

Solved Examples

Example 1. *If S_1 and S_2 are two subsets of a vector space V such that $S_1 \subseteq S_2$, then show that $S_2^0 \subseteq S_1^0$.*

Solution. Let $f \in S_2^0$. Then

$f(\alpha)=0 \ \forall \ \alpha \in S_2$

$\Rightarrow f(\alpha)=0 \ \forall \ \alpha \in S_1$ [$\because$ $S_1 \subseteq S_2$]

$\Rightarrow f \in S_1^0.$

$\therefore \quad S_2^0 \subseteq S_1^0.$

Example 2. *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 $S \subseteq L(S)$.

$\therefore \quad [L(S)]^0 \subseteq S^0. \quad \ldots(1)$

Now let $f \in S^0$. Then $f(\alpha)=0 \ \forall \ \alpha \in S$.

If β is any element of $L(S)$, then

$$\beta = \sum_{i=1}^{n} x_i \alpha_i \text{ where each } \alpha_i \in S.$$

We have $f(\beta) = \sum_{i=1}^{n} x_i f(\alpha_i)$

$=0$, since each $f(\alpha_i)=0$.

Thus $f(\beta)=0 \ \forall \ \beta \in L(S)$.

$\therefore \quad f \in (L(S))^0.$

Therefore $\quad S^0 \subseteq (L(S))^0 \quad \ldots(2)$

From (1) and (2), we conclude that $S^0 = (L(S))^0$.

Example 3. *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 Ex. 2] ...(1)

$\therefore \quad S^{00} = (L(S))^{00}.$

But V is finite-dimensional and $L(S)$ is a subspace of V. Therefore by theorem 3, $\quad (L(S))^{00} = L(S)$.

$\therefore$ from (1), we have $\quad S^{00} = L(S)$.

Example 4. *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 $W_1 = W_2$.* **(Meerut 1979)**

Scanned with CamScanner

Solution We have $W_1=W_2$

$\Rightarrow W_1^0=W_2^0$.

Conversely, let $W_1^0=W_2^0$

Then $W_1^{00}=W_2^{00}$

$\Rightarrow W_1=W_2$.

Example 5. *Let W_1 and W_2 be subspaces of a finite dimensional vector space V.*

(a) *Prove that* $(W_1+W_2)^0=W_1^0\cap W_2^0$.

(I.A.S. 1985; Meerut 76, 77, 88, 91, 93P)

(b) *Prove that* $(W_1\cap W_2)^0=W_1^0+W_2^0$.

(Meerut 1970, 73, 75, 77, 88, 91, 93P)

Solution. (a) 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

$\alpha=\alpha_1+\alpha_2$ 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 $W_2\subseteq W_1+W_2$.

$\therefore\ (W_1+W_2)^0\subseteq W_2^0$. ...(3)

From (2) and (3), we have

$(W_1+W_2)^0\subseteq W_1^0\cap W_2^0$. ...(4)

From (1) and (4), we have

$$(W_1+W_2)^0=W_1^0\cap W_2^0.$$

(b) 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

$(W_1^0+W_2^0)^0=W_1^{00}\cap W_2^{00}$

$\Rightarrow (W_1^0+W_2^0)^0=W_1\cap W_2$ [$\because$ $W_1^{00}=W$ etc.]

$\Rightarrow (W_1^0+W_2^0)^{00}=(W_1\cap W_2)^0$

$\Rightarrow W_1^0+W_2^0=(W_1\cap W_2)^0$.

Example 6. *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) $W_1^0\cap W_2^0=\{\hat{0}\}$

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 $W_1^0\cap W_2^0=\{\hat{0}\}$.

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

$\alpha=\alpha_1+\alpha_2$ where $\alpha_1\in W_1$, $\alpha_2\in W_2$.

We have $f(\alpha)=f(\alpha_1+\alpha_2)$

$=f(\alpha_1)+f(\alpha_2)$ [$\because$ f is linear functional]

$=0+0$ [$\because$ $f\in W_1^0$ and $\alpha_1\in W_1\Rightarrow f(\alpha_1)=0$ and similarly $f(\alpha_2)=0$]

$=0$.

Thus $f(\alpha)=0\ \forall\ \alpha\in V$.

$\therefore\ f=\hat{0}$.

$\therefore\ W_1^0\cap W_2^0=\{\hat{0}\}$.

(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

$\alpha=\alpha_1+\alpha_2$ 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

$f_1(\alpha)=f_1(\alpha_1+\alpha_2)=f(\alpha_2)$ ...(1)

and $f_2(\alpha)=f_2(\alpha_1+\alpha_2)=f(\alpha_1)$. ...(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

$f_1(a\alpha+b\beta)=f_1\ [a(\alpha_1+\alpha_2)+b(\beta_1+\beta_2)]$

$=f_1\ [(a\alpha_1+b\beta_1)+(a\alpha_2+b\beta_2)]$

$=f(a\alpha_2+b\beta_2)$

[$\because$ $a\alpha_1+b\beta_1\in W_1$, $a\alpha_2+b\beta_2\in W_2$]

$=af(\alpha_2)+bf(\beta_2)$ [$\because$ f is linear functional]

$=af_1(\alpha)+bf_1(\beta)$ [From (1)]

$\therefore\ f_1$ is linear functional on V *i.e.* $f_1\in V'$.

Scanned with CamScanner

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
$\alpha_1 = \alpha_1 + 0$, where $\alpha_1 \in W_1$, $0 \in W_2$.

$\therefore$ from (1), we have

$$f_1(\alpha_1) = f_1(\alpha_1 + 0) = f(0) = 0.$$

Thus $f_1(\alpha_1) = 0 \ \forall \ \alpha_1 \in W_1$.

$\therefore \ f_1 \in W_1^0$.

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 $(f_1 + f_2)(\alpha) = f_1(\alpha) + f_2(\alpha)$

$= f(\alpha_2) + f(\alpha_1)$ [From (1) and (2)]

$= f(\alpha_1) + f(\alpha_2)$

$= f(\alpha_1 + \alpha_2)$ [$\because$ f is linear functional]

$= f(\alpha)$.

Thus $(f_1 + f_2)(\alpha) = f(\alpha) \ \forall \ \alpha \in V$.

$\therefore \ f = f_1 + f_2$.

Thus $f \in V' \Rightarrow f = f_1 + f_2$ where $f_1 \in W_1^0$, $f_2 \in W_2^0$.

$\therefore \ V' = W_1^0 + W_2^0$.

Hence $V' = W_1^0 \oplus W_2^0$.

Example 7. *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.*

Solution. Let dim $V = n$, dim $W_1 = m$.

Then dim $W_2 = n - m$.

We have dim $W_1' = \dim W_1 = m$.

Also dim $W_2^0 = \dim V - \dim W_2 = n - (n-m) = m$.

$\therefore$ dim $W_1' = \dim W_2^0$

$\Rightarrow$ W_1' is isomorphic to W_2^0.

Again dim $W_2' = \dim W_2 = n - m$

Also dim $W_1^0 = \dim V - \dim W_1 = n - m$.

$\therefore$ dim $W_2' = \dim W_1^0$

$\Rightarrow$ $W_2' \cong W_1^0$.

§ 19. 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 $\alpha \in W \Rightarrow T(\alpha) \in W$.* **[Meerut 1976]**

Example 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.

Example 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 memeber of the set.

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 $V = n$.

Scanned with CamScanner

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 $1\leqslant j\leqslant m$, then α_j is in W. But W is invariant under T. Therefore if $1\leqslant j\leqslant m$, 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

$$T(\alpha_j)=\sum_{i=1}^{m} a_{ij}\,\alpha_i,\ 1\leqslant j\leqslant m. \qquad \ldots(2)$$

In other words in the relation (1), the scalars a_{ij} are all zero if $1\leqslant j\leqslant m$ and $m+1\leqslant i\leqslant n$.

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.

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.

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

$$\alpha=\alpha_1+\ldots+\alpha_k$$

$$\Rightarrow T(\alpha)=T(\alpha_1+\ldots+\alpha_k)$$

$$\Rightarrow T(\alpha)=T(\alpha_1)+\ldots+T(\alpha_k) \qquad [\because\ T \text{ is linear}]$$

$$\Rightarrow T(\alpha)=T_1(\alpha_1)+\ldots+T_k(\alpha_k) \qquad [\because\ \text{if } \alpha_i\in W_i, \text{ then by def. of } T_i, \text{ we have } T(\alpha_i)=T_i(\alpha_i)]$$

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.

Matrix representation of reducibility. If T is a linear operator on a finite dimensional vector space V and T 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

$$M=[T_1]_{B_1},\ \text{and}\ N=[T_2]_{B_2}.$$

Solved Examples

Example 1. *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 $T(W)\subseteq W$.*

Scanned with CamScanner

Solution. We have, by definition

$$T(W)=\{T(\alpha) : \alpha \in W\}.$$

Since $0 \in W$ and $T(0)=0$, therefore $T(W)$ is not empty because at least $0 \in T(W)$.

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

$aT(\alpha_1)+bT(\alpha_2)=T(a\alpha_1+b\alpha_2)$, because T 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

$a, b \in F$ and $T(\alpha_1), T(\alpha_2) \in T(W)$

$$\Rightarrow aT(\alpha_1)+bT(\alpha_2) \in T(W).$$

$\therefore$ $T(W)$ 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 $T(\alpha) \in W$. Thus $T(\alpha) \in T(W) \Rightarrow T(\alpha) \in W$.

Therefore $T(W) \subseteq W$.

Conversely suppose that $T(W) \subseteq W$.

Then $T(\alpha) \in W \ \forall \ \alpha \in W$. Therefore W is invariant under T.

Example 2. *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

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

If $\beta \in N(T)$, then $T(\beta)=0 \in N(T)$ 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 $\beta \in V \Rightarrow T(\beta) \in R(T)$.

Thus $\beta \in R(T) \Rightarrow T(\beta) \in R(T)$. Therefore $R(T)$ is invariant under T.

Example 3. *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.*

Solution. We have

$$\alpha \in W \Rightarrow \alpha \in W_i \text{ for each } i$$

$\Rightarrow T(\alpha) \in W_i$ for each i [$\because$ each W_i is invariant under T]

$\Rightarrow T(\alpha) \in \bigcap W_i \Rightarrow T(\alpha) \in W.$

$\therefore$ W is invariant under T.

Example 4. *Prove that the subspace spanned by two subspaces each of which is invariant under some linear operator T, is itself invariant under T.* **(Meerut 1987)**

Solution. Let W_1 and W_2 be two subspaces of a vector space V. Let W be the subspace of V spanned by $W_1 \cup W_2$. Then we know that

$$W=W_1+W_2.$$

Now it is given that both W_1 and W_2 are invariant under a linear operator T and we are to prove that W is also invariant under T.

Let $\alpha \in W$. Then

$$\alpha=\alpha_1+\alpha_2, \text{ where } \alpha_1 \in W_1, \alpha_2 \in W_2.$$

We have $T(\alpha)=T(\alpha_1+\alpha_2)$

$=T(\alpha_1)+T(\alpha_2)$ because T is linear.

Now $T(\alpha_1) \in W_1$ since W_1 is invariant under T and $\alpha_1 \in W_1$.

Similarly $T(\alpha_2) \in W_2$

Thus $T(\alpha_1)+T(\alpha_2) \in W_1+W_2$

i.e. $T(\alpha)=T(\alpha_1)+T(\alpha_2) \in W.$

Thus $\alpha \in W \Rightarrow T(\alpha) \in W$.

$\therefore$ W is invariant under T.

Example 5. *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*

Solution. If $f(t)$ is a polynomial in the indeterminate t over the field F, then we know that $f(T)$ is a linear operator on V where T is a linear operator on V.

Now W is the null space of $f(T)$. Therefore

$$\alpha \in W \Rightarrow f(T)(\alpha)=0. \qquad \ldots(1)$$

We are to show that W is invariant under T

i.e. $\alpha \in W \Rightarrow T(\alpha) \in W.$

Obviously $[f(T)]\,T=T\,f(T)$ because $t\,f(t)=f(t)\,t$ and polynomials in T behave like ordinary polynomials.

Scanned with CamScanner

$\therefore \quad \alpha \in W \Rightarrow [(f(T))\, T]\, (\alpha) = [T\, f(T)]\, (\alpha)$
$\Rightarrow f(T)\, [T(\alpha)] = T\, [f(T)\, (\alpha)]$
$\Rightarrow f(T)\, [T(\alpha)] = T(0)$ [From (1)]
$\Rightarrow f(T)\, [T(\alpha)] = 0$
$\Rightarrow T(\alpha) \in W$ since W is null space of $f(T)$.

$\therefore$ W is invariant under T.

Example 6. *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 $T(\alpha) \in W$.

$\therefore \quad T(\alpha) = c\alpha$ for some $c \in \mathbf{R}$
$\Rightarrow T(\alpha) = cI(\alpha)$ where I is identity operator on V
$\Rightarrow [T - cI](\alpha) = 0$
$\Rightarrow T - cI$ is singular [$\because \alpha \neq 0$]
$\Rightarrow T - cI$ 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.

$\therefore \begin{bmatrix} -c & -1 \\ 1 & -c \end{bmatrix}$ *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 $V_2(\mathbf{R})$ can be invariant under T.

Example 7. *Show that the space generated by $(1, 1, 1)$ and $(1, 2, 1)$ is an invariant sub-space of $\mathbf{R}^3$ under T, where*

$$T(x, y, z) = (x + y - z,\ x + y,\ x + y - z).$$ **(Meerut 1977)**

Solution. Let W be the subspace of $\mathbf{R}^3$ generated by the vectors $(1, 1, 1)$ and $(1, 2, 1)$. T is a linear transformation on $\mathbf{R}^3$ defined by $T(x, y, z) = (x + y - z,\ x + y,\ x + y - z)$.

Now W will be invariant under T if $\alpha \in W \Rightarrow T(\alpha) \in W$. If α is an arbitrary vector in W, then $\alpha = a\,(1, 1, 1) + b\,(1, 2, 1)$ for some $a, b \in \mathbf{R}$. Since T is a linear transformation, therefore $T(\alpha) = aT\,(1, 1, 1) + bT\,(1, 2, 1)$.

Now $T(\alpha)$ will be in W if we show that $T(1, 1, 1)$ and $T(1, 2, 1)$ are both in W. We have $T(1, 1, 1) = (1 + 1 - 1,\ 1 + 1,\ 1 + 1 - 1) = (1, 2, 1)$ which is a vector belonging to a set generating W. Therefore $T(1, 1, 1)$ is in W. Also $T(1, 2, 1) = (1 + 2 - 1,\ 1 + 2,\ 1 + 2 - 1) = (2, 3, 2) = (1, 1, 1) + (1, 2, 1)$. Thus $T(1, 2, 1)$ is a linear combination of the vectors $(1, 1, 1)$ and $(1, 2, 1)$ which generate W. Therefore $T(1, 2, 1)$ is also in W.

Hence $\alpha \in W \Rightarrow T(\alpha) \in W$. Therefore W is invariant under T.

Exercises

1. 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.

2. 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. **(Meerut 1976, 80)**

(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.

§ 20 Projections. Definition *Suppose a vector space V is the direct sum of its subspaces W_1 and W_2. Then every vector α in V can be uniquely written as $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 \in W_1$ and $\alpha_2 \in W_2$. The projection on W_1 along W_2 is the linear transformation E on V defined by* $E(\alpha) = \alpha_1$. **(Meerut 1975)**

In order to make the definition sensible, we shall show that the correspondence E as defined in it is a linear transformation on V.

Obviously E is a function from V into V.

Scanned with CamScanner

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 $E(\alpha) = \alpha_1$, $E(\beta) = \beta_1$. Also W_1 is a subspace and therefore $a\alpha_1 + b\beta_1 \in W_1$. Similarly $a\alpha_2 + b\beta_2 \in W_2$.

Now $E(a\alpha + b\beta) = E[a(\alpha_1 + \alpha_2) + b(\beta_1 + \beta_2)]$

$= E[(a\alpha_0 + b\beta_1) + (a\alpha_2 + b\beta_2)]$

$= a\alpha_1 + b\beta_1$

[by def. of E, since $a\alpha_1 + b\beta_1 \in W_1$ and $a\alpha_2 + b\beta_2 \in W_2$]

$= aE(\alpha) + bE(\beta)$ [$\because$ $E(\alpha) = \alpha_1$, $E(\beta) = \beta_1$]

$\therefore$ E is a linear transformation on V.

Theorem 1. *A linear transformation E on V is a projection on some subspace if and only if it is idempotent i.e. $E^2 = E$*

(Meerut 1975, 78, 85, 88, 91, 93, 93P)

Proof. Let $V = W_1 \oplus W_2$ and let E be the projection on W_1 along W_2. Then to prove that $E^2 = E$.

Let α be any vector in V.

Then $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 \in W_1$, $\alpha_2 \in W_2$.

By def. of projection, we have

$E(\alpha) = \alpha_1$. ...(1)

Now $E^2(\alpha) = E[E(\alpha)]$

$= E(\alpha_1)$

$= E(\alpha_1 + 0)$ where $\alpha_1 \in W_1$ and $0 \in W_2$

$= \alpha_1$ [by def of projection since the decomposition for $\alpha_1 \in V$ is $\alpha_1 + 0$]

$= E(\alpha)$ [From (1)]

Thus $E^2(\alpha) = E(\alpha)$ $\forall$ $\alpha \in V$ Therefore $E^2 = E$.

Conversely, let $E^2 = E$.

Let $W_1 = \{\alpha \in V : E(\alpha) = \alpha\}$,

and $W_2 = \{\alpha \in V : E(\alpha) = 0\}$.

W_2 is a subspace of V because it is null space of E.

Also W_1 is a subspace of V as shown below :

Let $a, b \in F$ and $\alpha, \beta \in W_1$. Then $E(\alpha) = \alpha$, $E(\beta) = \beta$.

We have $E(a\alpha + b\beta) = aE(\alpha) + bE(\beta)$ [$\because$ E is linear]

$= a\alpha + b\beta$.

Now $a\alpha + b\beta \in W_1$ and therefore W_1 is a subspace of V.

Now we shall prove that E is the projection on W_1 along W_2. First we shall prove that $V = W_1 \oplus W_2$. For this we are to prove that

(i) $V = W_1 + W_2$

and (ii) W_1 and W_2 are disjoint.

Proof of (i). Let $\alpha \in V$. Then α can be written as

$\alpha = E(\alpha) + [\alpha - E(\alpha)]$.

Let $\alpha_1 = E(\alpha)$ and $\alpha_2 = \alpha - E(\alpha)$.

We have $E(\alpha_1) = E[E(\alpha)] = E^2(\alpha)$

$= E(\alpha)$ [$\because$ $E = E^2$]

$= \alpha_1$.

$\therefore$ $\alpha_1 \in W_1$.

Also $E(\alpha_2) = E[\alpha - E(\alpha)] = E(\alpha) - E^2(\alpha)$

$= E(\alpha) - E(\alpha)$ [$\because$ $E = E^2$]

$= 0$.

$\therefore$ $\alpha_2 \in W_2$.

Thus $\alpha \in V$ can be written as $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 \in W_1$, $\alpha_2 \in W_2$. Therefore $V = W_1 + W_2$.

Proof of (ii). Let $\alpha \in W_1 \cap W_2$. Then $\alpha \in W_1$, $\alpha \in W_2$.

Now $\alpha \in W_1 \Rightarrow E(\alpha) = \alpha$.

Also $\alpha \in W_2 \Rightarrow E(\alpha) = 0$.

$\therefore$ $\alpha \in W_1, W_2 \Rightarrow \alpha = 0$.

$\therefore$ W_1 and W_2 are disjoint.

Thus $V = W_1 \oplus W_2$.

Now let $\alpha \in V$ and $\alpha = \alpha_1 + \alpha_2$ where $\alpha_1 \in W_1$, $\alpha_2 \in W_2$.

Then $E(\alpha) = E(\alpha_1 + \alpha_2)$

$= E(\alpha_2 + \alpha_2)$ [$\because$ E is linear]

$= \alpha_1 + 0$ [$\because$ $\alpha_1 \in W_1 \Rightarrow E(\alpha_1) = \alpha_1$ and $\alpha_2 \in W_2 \Rightarrow E(\alpha_2) = 0$]

$= \alpha_1$.

Hence E is a projection on W_1 along W_2

Theorem 2. *Suppose the vector space V is the direct sum of its subspaces W_1 and W_2. If E is the projection on W_1 along W_2, then W_1 and W_2 are respectively, the set of all solutions of the equations $E(\alpha) = \alpha$, and $E(\alpha) = 0$.*

Proof. Let V be the vector space. Let

$M = \{\alpha \in V : E(\alpha) = \alpha\}$

and $N = \{\alpha \in V : E(\alpha) = 0\}$.

Then to prove that

(i) $M = W_1$

and (ii) $N = W_2$.

Proof of (i). Let $\alpha \in W_1$. Then $\alpha \in V$ and it can be written as $\alpha = \alpha + 0$ where $\alpha \in W_1$, $0 \in W_2$.

$\therefore$ by def. of projection, we have

$E(\alpha)=\alpha$

$\Rightarrow \alpha \in M$.

$\therefore \quad W_1 \subseteq M$.

Now let $\alpha \in M$. Then $E(\alpha)=\alpha$.

Let $\alpha=\alpha_1+\alpha_2$ where $\alpha_1 \in W_1, \alpha_2 \in W_2$.

We have $\quad E(\alpha)=\alpha_1$ [by def of projection]

$\Rightarrow \alpha=\alpha_1 \in W_1$ [$\because$ $E(\alpha)=\alpha$]

$\therefore \quad M \subseteq W_1$.

Hence $M=W_1$.

Proof of (ii). Let $\alpha \in W_2$. Then $\alpha \in V$ can be written as $\alpha=0+\alpha$ where $0 \in W_1, \alpha \in W_2$.

$\therefore$ by def. of projection, we have

$E(\alpha)=0$

$\Rightarrow \alpha \in N$.

$\therefore \quad W_2 \subseteq N$.

Now let $\alpha \in N$. Then $E(\alpha)=0$.

Let $\alpha=\alpha_1+\alpha_2$ where $\alpha_1 \in W_1, \alpha_2 \in W_2$.

We have $\quad E(\alpha)=\alpha_1$ [by def. of projection]

$\Rightarrow 0=\alpha_1$ [$\because$ $E(\alpha)=0$]

$\Rightarrow \alpha=\alpha_2$ [$\because$ $\alpha=\alpha_1+\alpha_2$]

$\Rightarrow \alpha \in W_2$ because $\alpha_2 \in W_2$.

$\therefore \quad N \subseteq W_2$.

Hence $\quad N=W_2$.

Theorem 3. *Suppose the vector space V is the direct sum of its subspaces W_1 and W_2. If E is the projection on W_1 along W_2, then*

(i) the range of E, i.e. $R(E)=W_1$

and (ii) the null space of E i.e. $N(E)=W_2$.

Proof. (i) We have

$R(E)=\{\alpha \in V : \alpha=E(\beta) \text{ for some } \beta \in V\}$.

To prove that $R(E)=W_1$.

Let $\alpha \in W_1$. Then $\alpha=\alpha+0$ where $\alpha \in W_1, 0 \in W_2$

$\therefore \quad E(\alpha)=\alpha$ [by def of projection]

$\Rightarrow \alpha \in R(E)$ because α is the image of α under E.

$\therefore \quad W_1 \subseteq R(E)$.

Now let $\alpha \in R(E)$. Then there exists $\beta \in V$ such that $E(\beta)=\alpha$.

$\Rightarrow E^2(\beta)=E(\alpha)$

$\Rightarrow E(\beta)=E(\alpha)$ [$\because$ $E^2=E$, E being a projection]

$\Rightarrow E(\alpha)=\alpha$ [$\because$ $E(\beta)=\alpha$]

$\Rightarrow \alpha \in W_1$.

$\therefore \quad R(E) \subseteq W_1$.

Hence $\quad R(E)=W_1$.

(ii) We have

$N(E)=\{\alpha \in V : E(\alpha)=0\}$.

To prove that $\quad W_2=N(E)$.

For proof, see theorem 2, part (ii).

Theorem 4. *A linear transformation E is a projection if and only if $I-E$ is a projection; if E is the projection on W_1 along W_2, then $I-E$ is the projection on W_2 along W_1.* **(Meerut 1989)**

Proof. We recall that the set of all linear operators on a vector space V form a ring with unity element I with respect to addition and multiplication of transformations.

Suppose E is the projection. Then $E^2=E$.

$(I-E)$ will be projection if $(I-E)^2=I-E$.

We have $(I-E)^2=(I-E)(I-E)=II-IE-EI+E^2$

$=I-E-E+E$ [$\because$ $E^2=E$]

$=I-E$.

$\therefore \quad I-E$ is a projection.

Conversely, let $I-E$ be a projection. Then

$(I-E)^2=I-E$

$\Rightarrow (I-E)(I-E)=I-E \Rightarrow I-E-E+E^2=I-E$

$\Rightarrow E^2-E=\hat{0} \Rightarrow E^2=E$

$\Rightarrow E$ is a projection.

Now let E be the projection on W_1 along W_2.

Then $\quad V=W_1 \oplus W_2=W_2 \oplus W_1$.

Let $\alpha=\alpha_1+\alpha_2 \in V$ where $\alpha_1 \in W_1, \alpha_2 \in W_2$.

Then $E(\alpha)=\alpha_1$, by def. of projection.

Now $\quad (I-E)(\alpha)=I(\alpha)-E(\alpha)=\alpha-\alpha_1=\alpha_2$.

$\therefore \quad I-E$ is a projection on W_2 along W_1.

Theorem 5. *If $V=W_1 \oplus \ldots \oplus W_k$, then there exist k linear operators $E_1, \ldots, E_k$ on V such that*

(a) each E_i is a projection ($E_i^2=E_i$);

(b) $E_iE_j=\hat{0}$; if $i \neq j$;

(c) $I=E_1+\ldots+E_k$;

(d) the range of E_i is W_i. **(Meerut 1979, 83P)**

Scanned with CamScanner

Conversely, if $E_1, \ldots, E_k$ are k linear operators on V which satisfy conditions (a), (b) and (c), and if we let W_i be the range of E_i, then V is the direct sum of $W_1, \ldots, W_k$.

Proof. (a) Since V is the direct sum of $W_1, \ldots, W_k$, therefore if $\alpha \in V$ then we can uniquely write

$$\alpha = \alpha_1 + \ldots + \alpha_k \text{ where each } \alpha_i \in W_i,\ 1 \leqslant i \leqslant k.$$

Let E_i be a function from V into V defined by the rule

$$E_i(\alpha) = \alpha_i.$$

E_i is a linear transformation on V as shown below:

Let $\alpha, \beta \in V$. Then

$$\alpha = \alpha_1 + \ldots + \alpha_k,\ \alpha_i \in W_i,$$
$$\beta = \beta_1 + \ldots + \beta_k,\ \beta_i \in W_i.$$

If $a, b \in F$, we have

$$E_i(a\alpha + b\beta) = E_i[a(\alpha_1 + \ldots + \alpha_k) + b(\beta_1 + \ldots + \beta_k)]$$
$$= E_i[(a\alpha_1 + b\beta_1) + \ldots + (a\alpha_k + b\beta_k)]$$
$$= a\alpha_i + b\beta_i \quad \text{[by def. of } E_i]$$
$$= aE_i(\alpha) + bE_i(\beta).$$

$\therefore$ E_i is a linear transformation on V.

We have $E_i^2(\alpha) = E_i[E_i(\alpha)] = E_i(\alpha_i)$

$$= E_i(0 + \ldots + \alpha_i + 0 + \ldots + 0), \text{ where } \alpha_i \in W_i$$
$$= \alpha_i \quad \text{[by def. of } E_i]$$
$$= E_i(\alpha).$$

Thus $E_i^2(\alpha) = E_i(\alpha)\ \forall\ \alpha \in V$.

$\therefore$ $E_i^2 = E_i$ i.e. E_i is a projection.

Thus there exist k linear operators E_i, $1 \leqslant i \leqslant k$ on V such that

$$E_i^2 = E_i.$$

(b) Let $i \neq j$.

Let $\alpha \in V$. Then $\alpha = \alpha_1 + \ldots + \alpha_k$, where $\alpha_i \in W_i$.

We have $(E_i E_j)(\alpha) = E_i[E_j(\alpha)]$

$$= E_i(\alpha_j) \quad \text{[by def. of } E_j]$$
$$= 0$$

[$\because$ $i \neq j$ means that in the decomposition of α_j as the sum of the vectors of $W_1, \ldots, W_k$, the vector belonging to W_i will be 0]

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

$\therefore$ $E_i E_j = \hat{0}$ if $i \neq j$.

(c) Let $\alpha \in V$. Then $\alpha = \alpha_1 + \ldots + \alpha_k$ where each $\alpha_i \in W_i$. We have $(E_1 + \ldots + E_k)(\alpha) = E_1(\alpha) + \ldots + E_k(\alpha)$

$$= \alpha_1 + \ldots + \alpha_k = \alpha = I(\alpha).$$

$\therefore$ $E_1 + \ldots + E_k = I$.

(d) Let $R(E_i)$ denote the range of E_i.

Let $\alpha \in R(E_i)$. Then $\exists\ \beta \in V$ such that $E_i(\beta) = \alpha$.

Let $\beta = \beta_1 + \ldots + \beta_k$ where $\beta_i \in W_i$.

Then $E_i(\beta) = \beta_i$.

$\therefore$ $\alpha = \beta_i \in W_i$.

Thus $\alpha \in R(E_i) \Rightarrow \alpha \in W_i$.

$\therefore$ $R(E_i) \subseteq W_i$.

Now let $\alpha \in W_i$. Then $E_i(\alpha) = \alpha$.

$\therefore$ $\alpha \in R(E_i)$.

Thus $\alpha \in W_i \Rightarrow \alpha \in R(E_i)$.

$\therefore$ $W_i \subseteq R(E_i)$.

Hence $R(E_i) = W_i$.

Converse. Suppose $E_1, \ldots, E_k$ are linear operators on V which satisfy the first three conditions. Let W_i be the range of E_i.

From (c), we have

$$I = E_1 + \ldots + E_k$$
$$\Rightarrow I(\alpha) = (E_1 + \ldots + E_k)(\alpha),\ \forall\ \alpha \in V$$
$$\Rightarrow \alpha = E_1(\alpha) + \ldots + E_k(\alpha),\ \forall\ \alpha \in V \quad \ldots(1)$$

Since W_i is the range of E_i, therefore

$$E_i(\alpha) \in W_i.$$

Thus from (1), we see that if $\alpha \in V$, then $\alpha \in W_1 + \ldots + W_k$.

Therefore $V = W_1 + \ldots + W_k$.

Now to show that the expression (1) for α as the sum of the vectors belonging to $W_1, \ldots, W_k$ is unique. Let

$$\alpha = \alpha_1 + \ldots + \alpha_k \text{ where } \alpha_i \in W_i.$$

Since W_i is the range of E_i, therefore

let $\alpha_i = E_i(\beta_i)$ where $\beta_i \in V$, $1 \leqslant i \leqslant k$.

We have $E_j(\alpha) = E_j(\alpha_1 + \ldots + \alpha_k) = E_j(\alpha_1) + \ldots + E_j(\alpha_k)$

$$= \sum_{i=1}^{k} E_j(\alpha_i) = \sum_{i=1}^{k} E_j[E_i(\beta_i)] = \sum_{i=1}^{k} E_j E_i(\beta_i)$$

$$= E_j^2(\beta_j) \quad [\because\ E_i E_j = \hat{0} \text{ if } i \neq j]$$
$$= E_j(\beta_j) \quad [\because\ E_j^2 = E_j]$$
$$= \alpha_j.$$

Hence the expression for α is unique.

$\therefore$ V is the direct sum of $W_1, \ldots, W_k$.

Scanned with CamScanner

§ 21. **Projections and Invariance.**

Theorem 1. *If a subspace W_1 of a vector space V is invariant under the linear operator T on V, then $ETE=TE$ for every projection E on W_1. Conversely, if $ETE=TE$ for some projection E on W_1, then W_1 is invariant under T.* **(Meerut 1979, 88)**

Proof. T is a linear operator on V and W_1 is a subspace of V invariant under T. Let

$$V=W_1\oplus W_2 \text{ for some } W_2.$$

Let E be the projection on W_1 along W_2. Then to prove that

$$ETE=TE.$$

Let $\alpha \in V$.

Then we can write

$$\alpha=\alpha_1+\alpha_2 \text{ where } \alpha_1 \in W_1,\ \alpha_2 \in W_2.$$

We have $(ETE)(\alpha)=(ET)\,[E(\alpha)]$

$=(ET)\,(\alpha_1)$ [by def. of projection E]

$=E\,[T(\alpha_1)]$

$=T\,(\alpha_1)$

[$\alpha_1 \in W_1$ and W_1 is invariant under T. So $T(\alpha_1) \in W_1$ and consequently $E\,[T(\alpha_1)]=T(\alpha_1)$]

$=T\,[E(\alpha)]$ [$\because$ $E(\alpha)=\alpha_1$]

$=(TE)\,(\alpha)$.

Thus $(ETE)\,(\alpha)=(TE)\,(\alpha)\ \forall\ \alpha \in V$.

$\therefore$ $ETE=TE$.

Conversely, let $V=W_1\oplus W_2$ and $ETE=TE$ for the projection E on W_1 along W_2. Then to prove that W_1 is invariant under T.

Let $\alpha \in W_1$. Since $ETE=TE$, therefore

$(ETE)\,(\alpha)=(TE)\,(\alpha)$

$\Rightarrow (ET)\,[E\,(\alpha)]=T\,[E\,(\alpha)]$

$\Rightarrow (ET)\,(\alpha)=T\,(\alpha)$ [$\because$ $\alpha \in W_1 \Rightarrow E\,(\alpha)=\alpha$]

$\Rightarrow E\,[T\,(\alpha)]=T\,(\alpha)$

$\Rightarrow T\,(\alpha) \in W_1$ [Since E is projection on W_1 along W_2, therefore $E\,[T\,(\alpha)]=T\,(\alpha) \Rightarrow T\,(\alpha) \in W_1$]

Thus $\alpha \in W_1 \Rightarrow T(\alpha) \in W_1$.

$\therefore$ W_1 is invariant under T.

Note. The above theorem can be also stated in a slightly different form :

Let E be a projection on V and let T be a linear operator on V. Prove that the range of T is invariant under T iff $ETE=TE$.

Proof. Let the range of $E=W_1$ and the null space of $E=W_2$. Then $V=W_1\oplus W_2$ and E is the projection on W_1 along W_2.

Now proceed as in the above theorem.

Theorem 2. *If W_1 and W_2 are subspaces with $V=W_1\oplus W_2$, then a necessary and sufficient condition that the linear transformation T be reduced by the pair (W_1, W_2) is that $ET=TE$, where E is the projection on W_1 along W_2.*

Proof. $V=W_1\oplus W_2$ and E is the projection on W_1 along W_2. T is a linear operator on V and T is reduced by the pair (W_1, W_2).

$\therefore$ W_1 and W_2 are both invariant under T.

To prove that $ET=TE$.

Let $\alpha \in V$. We can write

$$\alpha=\alpha_1+\alpha_2 \text{ where } \alpha_1 \in W_1,\ \alpha_2 \in W_2.$$

W_1 and W_2 are both invariant under T. Therefore $T(\alpha_1)\in W_1$, $T(\alpha_2) \in W_2$. E is a projection on W_1 along W_2. Therefore $E[T(\alpha_1)]=T(\alpha_1)$ and $E[T\,(\alpha_2)]=\mathbf{0}$.

Now we have $(ET)\,(\alpha)=E\,[T\,(\alpha)]=E[T\,(\alpha_1+\alpha_2)]$

$=E[T\,(\alpha_1)+T(\alpha_2)]=E[T\,(\alpha_1)]+E[T\,(\alpha_2)]=T(\alpha_1)+\mathbf{0}$

$=T(\alpha_1)=T[E\,(\alpha)]=(TE)(\alpha)$.

$\therefore$ $ET=TE$

Conversely, suppose V is the direct sum of W_1 and W_2, E is the projection on W_1 along W_2 and $ET=TE$ where T is a linear operator on V. Then to prove that T is reduced by the pair (W_1, W_2), *i.e.* both W_1 and W_2 are invariant under T.

Let $\alpha \in W_1$. Since $ET=TE$, therefore

$(ET)(\alpha)=(TE)(\alpha)$

$\Rightarrow E[T\,(\alpha)]=T[E\,(\alpha)]$

$\Rightarrow E[T\,(\alpha)]=T(\alpha)$ [$\because$ $\alpha\in W_1 \Rightarrow E\,(\alpha)=\alpha$ if E is projection on W_1 along W_2]

$\Rightarrow T(\alpha) \in W_1$ [$\because$ if E is projection on W_1 along W_2, then $E\,(\alpha)=\alpha \Rightarrow \alpha \in W_1$]

Thus $\alpha \in W_1 \Rightarrow T(\alpha) \in W_1$.

$\therefore$ W_1 is invariant under T.

Now let $\beta\in W_2$. Since $ET=TE$, therefore

$(ET)\,(\beta)=(TE)\,(\beta)$

$\Rightarrow E[T\,(\beta)]=T[E(\beta)]$

$\Rightarrow E[T\,(\beta)]=T(\mathbf{0})$ [$\because$ $\beta\in W_2 \Rightarrow E(\beta)=\mathbf{0}$ if E is projection on W_1 along W_2]

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

$\Rightarrow E\,[T(\beta)]=\mathbf{0}$

$\Rightarrow T(\beta)\in W_2$ [$\because$ if E is projection on W_1 along W_2, then $E(\beta)=\mathbf{0} \Rightarrow \beta \in W_2$]

$\therefore$ W_2 is invariant under T.

Hence T is reduced by the pair (W_1, W_2).

Note. The above theorem may also be stated as below :

Let E be the projection on V and let T be a linear operator on V. Prove that both the range and null space of E are invariant under T iff ET=TE. (Meerut 1973)

Solved Examples

Example 1. *Let V be the direct sum of its subspaces W_1 and W_2. If E_1 is the projection on W_1 along W_2 and E_2 is the projection on W_2 along W_1, prove that*

(i) $E_1+E_2=I$,

and (ii) $E_1E_2=\hat{0}$, $E_2E_1=\hat{0}$.

Solution. (i) Let $\alpha \in V$. Then

$\alpha=\alpha_1+\alpha_2$ where $\alpha_1 \in W_1$, $\alpha_2 \in W_2$.

Since E_1 is projection on W_1 along W_2, therefore $E_1(\alpha)=\alpha_1$. Also E_2 is projection on W_2 along W_1. Therefore $E_2(\alpha)=\alpha_2$.

We have $(E_1+E_2)(\alpha)=E_1(\alpha)+E_2(\alpha)$

$=\alpha_1+\alpha_2=\alpha=I(\alpha)$.

$\therefore$ $E_1+E_2=I$.

(ii) We have $E_1E_2=E_1(I-E_1)$ $[\because E_1+E_2=I \Rightarrow E_2=I-E_1]$

$=E_1I-E_1^2$

$=E_1-E_1$ $[\because E_1$ is projection $\Rightarrow E_1^2=E_1]$

$=\hat{0}$.

Similarly $E_2E_1=E_2(I-E_2)=E_2I-E_2^2$

$=E_2-E_2=\hat{0}$.

Example 2. *Let $E_1, \ldots, E_k$ be linear operators on a vector space V such that $E_1+\ldots+E_k=I$. Prove that if*

$E_iE_j=0$ for $i\neq j$, then $E_i^2=E_i$ for each i.

Solution. We have

$E_i^2=E_iE_i=E_i(I-E_1-E_2-\ldots-E_{i-1}-E_{i+1}-\ldots-E_k)$

$=E_iI-E_iE_1-E_iE_2-\ldots-E_iE_{i-1}-E_iE_{i+1}-\ldots-E_iE_k$

$=E_i$. $[\because E_iE_j=\hat{0}$ for $i\neq j]$.

Example 3. *Let E be an idempotent linear operator on a vector space V i.e., $E^2=E$. If W_1 is the range of E and W_2 is the null space of E, show that*

(i) *α is in W_1 iff $E(\alpha)=\alpha$;*

(ii) *V is the direct sum of W_1 and W_2;*

(iii) *E is the projection on W_1 along W_2.* (Meerut 1985)

Solution. (i) Let $\alpha \in W_1$ where W_1 is the range of E.

Then $\exists\, \beta \in V$ such that $E(\beta)=\alpha$.

Now $E(\beta)=\alpha$

$\Rightarrow E[E(\beta)]=E(\alpha) \Rightarrow E^2(\beta)=E(\alpha)$

$\Rightarrow E(\beta)=E(\alpha)$

$\Rightarrow \alpha=E(\alpha)$. $[\because E^2=E]$

Now let $\alpha \in V$ be such that $E(\alpha)=\alpha$. Since α is the image of α under E, therefore $\alpha \in$ the range of E *i.e.*, $\alpha \in W_1$.

(ii) Let $\alpha \in V$. We can write

$\alpha=E(\alpha)+[\alpha-E(\alpha)]$

$=\alpha_1+\alpha_2$ where $\alpha_1=E(\alpha)$ and $\alpha_2=\alpha-E(\alpha)$.

Since $\alpha_1=E(\alpha)$, therefore α_1 is in the range of E *i.e.*, α_1 is in W_1.

Also $E(\alpha_2)=E[\alpha-E(\alpha)]=E(\alpha)-E^2(\alpha)$

$=E(\alpha)-E(\alpha)$

$=0$. $[\because E^2=E]$

$\therefore$ $\alpha_2 \in$ the null space of E *i.e.* $\alpha_2 \in W_2$.

Thus $\alpha=\alpha_1+\alpha_2$ where $\alpha_1 \in W_1$, $\alpha_2 \in W_2$.

$\therefore$ $V=W_1+W_2$.

Now let $\alpha \in W_1 \cap W_2$. Then $\alpha \in W_1$, W_2. We have

$\alpha \in W_1 \Rightarrow E(\alpha)=\alpha$.

Also since $\alpha \in W_2$, therefore $E(\alpha)=0$.

$\therefore$ $\alpha=0$ and thus W_1 and W_2 are disjoint.

$\therefore$ V is the direct sum of W_1 and W_2.

(iii) Let $\alpha \in V$. Then $\alpha=\alpha_1+\alpha_2$ where $\alpha_1 \in W_1$, $\alpha_2 \in W_2$.

We have $E(\alpha)=E(\alpha_1)+E(\alpha_2)=\alpha_1+0=\alpha_1$.

$\therefore$ E is the projection on W_1 along W_2.

Example 4. *V is an n-dimensional vector space over a field F and E is a linear operator on V which is idempotent. Show that $V=R\oplus N$ where R is the range space of E and N is the null space of E.* (Meerut 1977, 90)

Solution. In order to prove that $V=R\oplus N$, we are to prove that

(i) $V=R+N$, and (ii) $R\cap N=\{0\}$.

For proof of (i) see Ex. 3.

Now we shall prove (ii). Let $\alpha \in R \cap N$.

Then $\alpha \in R$, and $\alpha \in N$.

By def. of N, we have

$\alpha \in N \Rightarrow E(\alpha)=0$.

Scanned with CamScanner

Also by def of R, we have $\alpha \in R \Rightarrow \exists\, \beta \in V$ such that
$E(\beta)=\alpha$.

Now $E(\beta)=\alpha$

$\Rightarrow E[E(\beta)]=E(\alpha) \Rightarrow E^2(\beta)=E(\alpha)$

$\Rightarrow E(\beta)=E(\alpha)$ $[\because E^2=E]$

$\Rightarrow \alpha=E(\alpha)$ $[\because E(\beta)=\alpha]$

$\Rightarrow \alpha=\mathbf{0}$. $[\because E(\alpha)=0]$

$\therefore$ $R \cap N=\{\mathbf{0}\}$. Hence $V=R \oplus N$.

Example 5. *Prove that if T is a linear transformation on V such that $T^2(I-T)=T(I-T)^2=\hat{0}$, then T is a projection on V.*

Solution. We have $T^2(I-T)=T(I-T)^2=\hat{\mathbf{0}}$

$$\Rightarrow T^2I-T^3=T[(I-T)(I-T)]=\mathbf{0}$$

$$\Rightarrow T^2-T^3=T[I-T-T+T^2]=\hat{\mathbf{0}}$$

$$\Rightarrow T^2-T^3=T-T^2-T^2+T^3=\hat{\mathbf{0}}$$

$$\Rightarrow T^3=T^2 \text{ and } T^3=T^2+T^2-T$$

$$\Rightarrow T^2=T^2+T^2-T \Rightarrow T^2=T.$$

$\therefore$ T is a projection on V.

Example 6. *If T is a linear transformation on V, E is a projection on V and $F=I-E$, then $T=ETE+ETF+FTE+FTF$.*

Solution. We have

$$ETE+ETF+FTE+FTF$$

$=ETE+ET(I-E)+FTE+FT(I-E)$ $[\because F=I-E]$

$=ETE+ETI-ETE+FTE+FTI-FTE$

$=(ETE-ETE)+ET+FT+(FTE-FTE)$

$=\hat{\mathbf{0}}+ET+FT+\hat{\mathbf{0}}=ET+FT=ET+(I-E)T$

$=ET+IT-ET=T$

Example 7. *If E_1 and E_2 are projections on V and if $E_1E_2=E_2E_1$, then $E_1+E_2-E_1E_2$ is a projection.*

Solution. Since E_1, E_2 are projections, therefore

$$E_1^2=E_1,\ E_2^2=E_2.$$

Also it is given that $E_1E_2=E_2E_1$.

Therefore $(E_1+E_2-E_1E_2)^2$

$=(E_1+E_2-E_1E_2)(E_1+E_2-E_1E_2)$

$=(E_1^2+E_1E_2-E_1^2E_2+E_2E_1+E_2^2-E_2E_1E_2-E_1E_2E_1$
$\qquad -E_1E_2^2+E_1E_2E_1E_2$

$=E_1+E_1E_2-E_1E_2+E_2E_1+E_2-E_1E_2E_2-E_1E_1E_2$
$\qquad -E_1E_2+E_1E_1E_2E_2$

$=E_1+E_2E_1+E_2-E_1E_2^2-E_1^2E_2-E_1E_2+E_1^2E_2^2$

$=E_1+E_2+E_1E_2-E_1E_2-E_1E_2-E_1E_2+E_1E_2=E_1+E_2-E_1E_2$.

$\therefore$ $E_1+E_2-E_1E_2$ is idempotent and therefore is a projection.

Example 8. *Two projections E and F have the same range iff $EF=F$ and $FE=E$.*

Solution. E and F are two projections having W_1 and W_2 as their ranges respectively. Also $EF=F$ and $FE=E$. Then to prove that $W_1=W_2$.

Let $\alpha \in W_1$. Then

$E(\alpha)=\alpha$ $[\because W_1$ is range of $E]$

$\Rightarrow F[E(\alpha)]=F(\alpha)$

$\Rightarrow (FE)(\alpha)=F(\alpha)$

$\Rightarrow E(\alpha)=F(\alpha)$ $[\because FE=E]$

$\Rightarrow \alpha=F(\alpha)$

$\Rightarrow \alpha \in W_2$

$[\because F(\alpha)=\alpha \Rightarrow \alpha \in$ the range of F *i.e.* $\alpha \in W_2]$

$\therefore$ $W_1 \subseteq W_2$

Now let $\beta \in W_2$. Then

$F(\beta)=\beta$

$\Rightarrow E[F(\beta)]=E(\beta)$

$\Rightarrow (EF)(\beta)=E(\beta)$

$\Rightarrow F(\beta)=E(\beta)$ $[\because EF=F]$

$\Rightarrow \beta=E(\beta)$

$\Rightarrow \beta \in W_1$ $[\because W_1$ is the range of $E]$

$\therefore$ $W_2 \subseteq W_1$.

Hence $W_1=W_2$.

Conversely, suppose E and F are two projections having the same range. Then to prove that $EF=F$ and $FE=E$

Let $\alpha \in V$. We have

$(EF)(\alpha)=E[F(\alpha)]$

$=F(\alpha)$

$[\because F(\alpha) \in$ the range of $F \Rightarrow F(\alpha) \in$ the range of $E]$

$\therefore$ $EF=F$.

Also $(FE)(\alpha)=F[E(\alpha)]$

$=E(\alpha)$

$[\because E(\alpha) \in$ the range of $E \Rightarrow E(\alpha) \in$ the range of $F]$.

$\therefore$ $FE=E$.

Scanned with CamScanner

Example 3. *Suppose that E and F are projections on a vector space V over the field F whose characteristic is not equal to 2 i.e. F is such that* $1+1\neq 0$. *Then prove that* $E+F$ *is a projection if and only if* $EF=FE=\hat{0}$.

Solution. Let $E+F$ be a projection. Then

$$(E+F)^2=E+F$$

$$\Rightarrow (E+F)(E+F)=E+F$$

$$\Rightarrow E^2+EF+FE+F^2=E+F$$

$$\Rightarrow E+EF+FE+F=E+F \quad [\because E^2=E,\ F^2=F]$$

$$\Rightarrow EF+FE=\hat{0}. \qquad \ldots(1)$$

Multiplying (1) on both left and right by E, we get

$$E^2F+EFE=\hat{0}$$

i.e. $$EF+EFE=\hat{0}, \qquad \ldots(2)$$

and $$EFE+FE^2=\hat{0}$$

i.e. $$EFE+FE=\hat{0}. \qquad \ldots(3)$$

Subtracting (2) and (3), we get

$$EF-FE=\hat{0}$$

$$\Rightarrow EF=FE.$$

Putting $FE=EF$ in (1), we get

$$EF+EF=\hat{0}$$

$$\Rightarrow (1+1)\,EF=\hat{0}$$

$$\Rightarrow EF=\hat{0} \qquad [\because 1+1\neq 0]$$

Thus $EF=FE=\hat{0}$.

Conversely, suppose that $EF=FE=\hat{0}$.

We have $(E+F)^2=E^2+EF+FE+F^2$.

$$=E+\hat{0}+\hat{0}+F=E+F.$$

$\therefore$ $E+F$ is a projection.

Example 10. *Let E be a polynomial on a finite-dimensional vector space V. Show that by choosing a suitable basis B for V, the matrix of E with respect to B can be put into a particular simple form.*

Solution. Let $V=W_1\oplus W_2$. Suppose E is the projection on W_1 along W_2. Let $\dim V=n$, $\dim W_1=m$.

Then $\dim W_2=n-m$. 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

$$B=\{\alpha_1,\ldots,\alpha_m,\alpha_{m+1},\ldots,\alpha_n\} \text{ is a basis for } V.$$

We have, for $1\leqslant i\leqslant m$

$E(\alpha_i)=\alpha_i$ $[\because \alpha_i\in W_1$ and E is projection on W_1 along $W_2]$.

Also for $m+1\leqslant i\leqslant n$,

$$E(\alpha_i)=0 \qquad [\because \alpha_i\in W_2 \text{ in this case}]$$

$\therefore$ the matrix of E with respect to the ordered basis B is

$$[E]_B=\begin{bmatrix} I & O \\ O & O \end{bmatrix}$$

where I is unit matrix of order m and O's are null matrices of suitable sizes.

Exercises

1. Let E be a projection on a vector space V and let T be a linear operator on V. Prove that the range of E is invariant under T if and only if $ETE=TE$.
2. Let E be the projection on a vector space V and let T be a linear operator on V. Prove that both range and null space of E are invariant under T if and only if
$$ET=TE.$$
3. If $V=W_1\oplus W_2\oplus W_3$, prove that there exist three linear operators E_1, E_2, E_3 on V such that
(i) each E_i is a projection $(E_i^2=E_i)$;
(ii) $E_iE_j=0$ if $i\neq j$;
(iii) the range of E_i is W_i. **(Meerut 1971, 91)**
4. Let E_1, E_2 be two linear operators on a vector space V such that $E_1+E_2=I$, $E_1^2=E_1$, $E_2^2=E_2$.
Prove that $E_1E_2=\hat{0}=E_2E_1$.
5. Let E_1 and E_2 be linear operators on a vector space V such that $E_1+E_2=I$. Prove that
$E_iE_j=\hat{0}$ for $i\neq j$ if and only if $E_i^2=E_i$ for each i. **(Meerut 1974)**

§ 22 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

Scanned with CamScanner

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 $T_1T_2\alpha$ 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 have

$$[\alpha, f]=f(\alpha),\ [\alpha, g]=g(\alpha).$$

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^*.

(Allahabad 1977)

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

$$f(\alpha)=g[T(\alpha)] \ \forall\ \alpha \in U \qquad \text{...(1)}$$

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 $f=T'(g)$ on U, defined by (1).

Thus $\quad T': V' \to U'$ such that

$T'(g)=f \ \forall\ g \in V'$ 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

$$[\alpha, T'g]=[T\alpha, g] \qquad \text{...(2)}$$

$$\forall\ g \in V' \text{ and } \forall\ \alpha \in U.$$

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

$$T'(ag_1+bg_2)=aT'g_1+bT'g_2 \qquad \text{...(3)}$$

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']

$=a\,[T\alpha, g_1]+b\,[T\alpha, g_2]$

[by def. of scalar multiplication in V']

$=a\,[\alpha, T'g_1]+b\,[\alpha, T'g_2]$ [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

$$[\alpha, T_1g]=[T\alpha, g] \ \forall\ g \in V' \text{ and } \alpha \in U. \qquad \text{...(4)}$$

Then from (2) and (4), we get

$$[\alpha, T_1g]=[\alpha, T'g] \ \forall\ \alpha \in U,\ \forall\ g \in V'$$

$$\Rightarrow T_1g=T'g \ \forall\ g \in V'$$

$$\Rightarrow T_1=T'.$$

Scanned with CamScanner

$\therefore$ 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.*

$[R(T)]^0=N(T')$. **(Marathwada 1971; Meerut 83P, 87, 88)**

If in addition U and V are finite dimensional, then

(ii) $\rho(T')=\rho(T)$ **(Meerut 1983P, 88; Poona 70)**

and (iii) the range of T' is the annihilator of the null space of T i.e.

$R(T')=[N(T)]^0$. **(Marathwada 1971; Meerut 88)**

Proof. (i) If $g \in V'$, then by definition of T', we have

$[\alpha, T' g]=[T\alpha, g] \ \forall \ \alpha \in U$, ...(1)

Let $g \in N(T')$ which is a subspace of V'. Then

$Tg=\hat{0}$ where $\hat{0}$ is zero element of U' i.e., $\hat{0}$ is zero functional on U. Therefore from (1), we get

$[T\alpha, g]=[\alpha, \hat{0}] \ \forall \ \alpha \in U$

$\Rightarrow [T\alpha, g]=0 \ \forall \ \alpha \in U$ $\quad [\because \ \hat{0}(\alpha)=0 \ \forall \ \alpha \in U]$

$\Rightarrow g(\beta)=0 \ \forall \ \beta \in R(T)$ $\quad [\because \ R(T)=\{\beta \in V : \beta=T(\alpha)$ for some $\alpha \in U\}]$

$\Rightarrow g \in [R(T)]^0$.

$\therefore \ N(T') \subseteq [R(T)]^0$.

Now let $g \in [R(T)]^0$ which is a subspace of V'. Then

$g(\beta)=0 \ \forall \ \beta \in R(T)$

$\Rightarrow [T\alpha, g]=0 \ \forall \ \alpha \in U$ $\quad [\because \ \forall \ \alpha \in U, T\alpha \in R(T)]$

$\Rightarrow [\alpha, T' g]=0 \ \forall \ \alpha \in U$ $\quad$ [From (1)]

$\Rightarrow T'g=\hat{0}$ (zero functional on U)

$\Rightarrow g \in N(T')$.

$\therefore \ [R(T)]^0 \subseteq N(T')$.

Hence $[R(T)]^0=N(T')$.

(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

$\dim R(T)+\dim [R(T)]^0=\dim V$. [See Th. 2 Page 206]

$\therefore \ \dim [R(T)]^0=\dim V-\dim R(T)$

$=\dim V-r=m-r$.

By part (i) of this theorem $[R(T)]^0=N(T')$.

$\therefore \ \dim N(T')=m-r$

$\Rightarrow$ nullity of $T'=m-r$.

But T' is a linear transformation from V' into U'.

$\therefore \ \rho(T')+\nu(T')=\dim V'$

or $\rho(T')=\dim V'-\nu(T')$

$=\dim V-$ nullity of $T'=m-(m-r)=r$.

$\therefore \ \rho(T)=\rho(T')=r$.

(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 $f \in R(T')$. Then $f=T'g$ 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 $f(\alpha)=0 \ \forall \ \alpha \in N(T)$. Therefore $f \in [N(T)]^0$.

$\therefore \ R(T') \subseteq [N(T)]^0$

$\Rightarrow R(T')$ is a subspace of $[N(T)]^0$.

Now $\dim N(T)+\dim [N(T)]^0=\dim U$. [Theorem 2 page 206]

$\therefore \ \dim [N(T)]^0=\dim U-\dim N(T)$

$=\dim R(T)$ $\quad [\because \ \dim U=\dim R(T)+\dim N(T)]$

$=\rho(T)$

$=\rho(T')=\dim R(T')$.

Thus $\dim R(T')=\dim [N(T)]^0$ and $R(T') \subseteq [N(T)]^0$.

$\therefore \ R(T')=[N(T)]^0$.

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.*

Scanned with CamScanner

Proof. Let dim $U=n$, dim $V=m$.

Let $B=\{\alpha_1, \ldots, \alpha_n\}$, $B'=\{f_1, \ldots, f_n\}$,

$B_1=\{\beta_1, \ldots, \beta_m\}$, $B_1'=\{g_1, \ldots, g_m\}$.

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

$$T(\alpha_j) \text{ or simply } T\alpha_j=\sum_{i=1}^{m} a_{ij}\beta_i,\ j=1, 2, \ldots, n. \quad \ldots(1)$$

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

$$T'(g_i) \text{ or simply } T' g_i=\sum_{j=1}^{n} c_{ji} f_j,\ i=1, 2, \ldots, m \quad \ldots(2)$$

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

$$f=\sum_{j=1}^{n} f(\alpha_j) f_j. \qquad \text{[See theorem 3 page 197]}$$

Applying this formula for $T' g_i$ in place of f, we get

$$T' g_i=\sum_{j=1}^{n} \{(T' g_i)(\alpha_j)\} f_j \quad \ldots(3)$$

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']

$=g_i\left(\sum_{k=1}^{m} a_{kj}\beta_k\right)$ [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]

$=a_{ij}$ [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

$$T' g_i=\sum_{j=1}^{n} a_{ij} f_j. \quad \ldots(4)$$

Since $f_1, \ldots, f_n$ 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

$B=\{\alpha_1, \ldots, \alpha_n\}$ and $B_1=\{\beta_1, \ldots, \beta_m\}$

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 $T(\alpha_1), \ldots, T(\alpha_n)$ are nothing but the column vectors of the matrix A. Also these vectors span the range of T because $\alpha_1, \ldots, \alpha_n$ 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 liner 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 $V(F)$:*

Scanned with CamScanner

(i) $\hat{0}'=\hat{0}$;

(ii) $I'=I$;

(iii) $(T_1+T_2)'=T_1'+T_2'$; (Meerut 1975, 90)

(iv) $(T_1 T_2)'=T_2' T_1'$; (Meerut 1975, 90; Allahabad 77)

(v) $(aT)'=aT'$ *where* $a \in F$;

(vi) $(T^{-1})'=(T')^{-1}$ *if T is invertible*; (Meerut 1975)

(vii) $(T')'=T''=T$ *if V is finite-dimensional*.

Proof. (i) If $\hat{0}$ is the zero transformation on V, then by the definition of the adjoint of a linear transformation, we have

$[\alpha, \hat{0}'g]=[\hat{0}\alpha, g]$ for every g in V' and α in V

$=[0, g]$ for every g in V' [$\because$ $\hat{0}(\alpha)=0 \; \forall \; \alpha \in V$]

$=0$ [$\because$ $g(0)=0$]

$=[\alpha, \hat{0}] \; \forall \; \alpha \in V$ [Here $\hat{0} \in V'$ and $\hat{0}(\alpha)=0$]

$=[\alpha, \hat{0}\, g] \; \forall \; g \in V'$ and $\alpha \in V$

[Here $\hat{0}$ is the zero transformation on V']

Thus we have

$[\alpha, \hat{0}'\, g]=[\alpha, \hat{0}\, g]$ for all g in V' and α in V.

$\therefore \quad \hat{0}'=\hat{0}$.

(ii) If I is the identity transformation on V, then by the definition of the adjoint of a linear transformation, we have

$[\alpha, I'\, g]=[I\alpha, g]$ for every g in V' and α in V

$=[\alpha, g]$ for every g in V' and α in V

[$\because$ $I(\alpha)=\alpha \; \forall \; \alpha \in V$]

$=[\alpha, Ig]$ for every g in V' and α 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

$[\alpha, (T_1+T_2)'\, g]=[(T_1+T_2)\, \alpha, g]$ for every g in V' and α in V

$=[T_1\alpha+T_2\alpha, g]$ [by def. of addition of linear transformations]

$=[T_1\, \alpha, g]+[T_2\, \alpha, g]$ [by linearity property of g]

$=[\alpha, T_1'\, g]+[\alpha, T_2'\, g]$ [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

$[\alpha, (T_1+T_2)'\, g]=[\alpha, (T_1'+T_2')\, g]$ for every g in V' and α 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

$[\alpha, (T_1 T_2)'\, g]=[(T_1 T_2)\, \alpha, g]$ for every g in V' and α in V

$=[(T_1)\, T_2\alpha, g]$ [by def. of product of linear transformations]

$=[T_2\alpha, T_1'\, g]$ [by def. of adjoint]

$=[\alpha, T_2'\, T_1'\, g]$ [by def. of adjoint]

Thus we have

$[\alpha, (T_1 T_2)'\, g]=[\alpha, T_2'\, T_1'\, g]$ for every g in V' and α 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

$[\alpha, (aT)'\, g]=[(aT)\, \alpha, g]$ for every g in V' and α 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]

$\therefore \quad (aT)'=aT'$.

(vi) Suppose T is an invertible linear operator on V. If T^{-1} is the inverse of T, we have

$T^{-1}\, T=I=TT^{-1}$

$\Rightarrow (T^{-1}\, T)'=I'=(TT^{-1})'$

$\Rightarrow T'\,(T^{-1})'=I=(T^{-1})'\, T'$

[Using results (ii) and (iv)]

$\therefore \quad T'$ 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

$[g, T'' L_\alpha] = [g, (T')' L_\alpha] = [T' g, L_\alpha]$ for every $g \in V'$ 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

$\Rightarrow T\alpha - T''\alpha = 0$ for every α in V

$\Rightarrow (T - T'')\alpha = 0$ for every α in V

$\Rightarrow T - T'' = \hat{0}$

$\Rightarrow T = T''$.

§ 23. Adjoints of projections.

Theorem 1. *If E is the projection on W_1 along W_2, then E' is the projection on $W_2{}^0$ along $W_1{}^0$.* (Meerut 1975)

Proof. E is a linear operator on V and E' is a linear operator on V'. E is the projection on W_1 along W_2. Since E is a projection, therefore $E^2 = E$. We have

$$(E')^2 = E' E' = (EE)' = (E^2)' = E'.$$

$\therefore$ E' is also a projection.

If W_1, W_2 are subspaces of V, then $W_1{}^0, W_2{}^0$ are subspaces of V'. Also if $V = W_1 \oplus W_2$, then $V' = W_2{}^0 \oplus W_1{}^0$. [See Ex. 6 page 211] Now E' will be the projection on $W_2{}^0$ along $W_1{}^0$ if we prove that $W_2{}^0 = M$ and $W_1{}^0 = N$ where

$$M = \{f \in V' : E'(f) = f\}$$

and

$$N = \{f \in V' : E'(f) = \hat{0}\}.$$

(i) First we shall prove that $W_2{}^0 = M$.

Let $f \in W_2{}^0$. Then to prove that $f \in M$ *i.e.*, $E' f = f$.

For every α in V, we have

$[\alpha, f] = [E\alpha + (I - E)\alpha, f]$

$= [E\alpha, f] + [(I - E)\alpha, f]$ [$\because$ f is linear]

$= [E\alpha, f] + 0$ [Since $f \in W_2{}^0$ and $(I - E)\alpha \in W_2$, $I - E$ being projection on W_2 along W_1]

$= [\alpha, E' f]$ [by def. of adjoint]

Thus $E' f = f$. Therefore $f \in M$.

$\therefore$ $W_2{}^0 \subseteq M$.

Now let $f \in M$ *i.e.*, let $E' f = f$. Then to prove that $f \in W_2{}^0$.

For every α in W_2, we have

$[\alpha, f] = [\alpha, E' f]$

$= [E\alpha, f]$ [by def of adjoint]

$= [0, f]$ [$\because$ $\alpha \in W_2 \Rightarrow E\alpha = 0$, E being projection on W_1 along W_2]

$= 0$.

$\therefore$ $f \in W_2{}^0$.

$\therefore$ $M \subseteq W_2{}^0$.

Hence $M = W_2{}^0$.

(ii) Now we shall prove that $W_1{}^0 = N$.

Let $f \in W_1{}^0$. Then to prove that $f \in N$ *i.e.*, to prove that $E' f = \hat{0}$. For every α in V, we have

$[\alpha, E' f] = [E\alpha, f]$ [by def. of adjoint]

$= 0$ [Since $f \in W_1{}^0$ and $E\alpha \in W_1$, E being projection on W_1 along W_2]

Thus $[\alpha, E' f] = 0$ for every α in V.

$\therefore$ $E' f = \hat{0}$.

$\therefore$ $f \in N$ and thus $W_1{}^0 \subseteq N$.

Now let $f \in N$ *i.e.*, let $E' f = \hat{0}$. Then to prove that $f \in W_1{}^0$. For every α in W_1, we have

$[\alpha, f] = [E\alpha, f]$ [$\because$ $\alpha \in W_1 \Rightarrow E\alpha = \alpha$, E being projection on W_1 along W_2]

$= [\alpha, E' f]$ [by def of adjoint]

$= [\alpha, \hat{0}]$

$= 0$.

$\therefore$ $f \in W_1{}^0$ and thus $N \subseteq W_1{}^0$

Hence $N = W_1{}^0$.

This completes the proof of the theorem.

Theorem 2. *If W_1 is invariant under T, then $W_1{}^0$ is invariant* under T', if T is reduced by (W_1, W_2), then T' is reduced by (W_1^0, W_2^0).

Proof. Since W_1 is invariant under T, therefore
$$\alpha \in W_1 \Rightarrow T\alpha \in W_1.$$
Now let $f \in W_1^0$. Then to prove that $T'f \in W_1^0$. For every α in W_1, we have
$$[\alpha, T'f] = [T\alpha, f]$$
$$=0 \qquad [\because f \in W_1^0 \text{ and } T\alpha \in W_1]$$
$\therefore \quad T'f \in W_1^0.$

Hence W_1^0 is invariant under T'.

Now suppose that T is reduced by the pair (W_1, W_2). Then we have

(i) $V = W_1 \oplus W_2$,

(ii) both W_1 and W_2 are invariant under T.

Now $V' = W_1^0 \oplus W_2^0$ because $V = W_1 \oplus W_2$.

Also W_1^0 and W_2^0 are both invariant under T' as we have just proved

Hence T' is reduced by the pair (W_1^0, W_2^0).

Solved Examples

Example 1. *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
$$(C')^{-1} = (C^{-1})'.$$
$\therefore \quad A' = (C')^{-1} B' C'$

$\Rightarrow C' A' (C')^{-1} = B'$ [Multiplying on right by $(C')^{-1}$ and on left by C']

$\Rightarrow B'$ is similar to A'

$\Rightarrow A'$ and B' are similar.

Example 2. *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 $\dim V' = n$.

Also $\dim L(V, V) = n^2$, $\dim L(V', V') = n^2$.

Let $\psi : L(V, V) \to L(V', V')$ such that
$$\psi(T) = T' \ \forall \ T \in L(V, V).$$
(i) ψ is linear transformation.

Let $a, b \in F$ 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]

$\therefore \quad \psi$ is a linear transformation from $L(V, V)$ into $L(V', V')$.

(ii) ψ is one-one.

Let $T_1, T_2 \in L(V, V)$. Then
$$\psi(T_1) = \psi(T_2)$$
$$\Rightarrow T_1' = T_2'$$
$$\Rightarrow T_1'' = T_2''$$
$$\Rightarrow T_1 = T_2 \qquad [\because V \text{ is finite-dimensional}]$$
$\therefore \quad \psi$ is one-one.

(iii) ψ is onto.

We have $\quad \dim L(V, V) = \dim L(V', V') = n^2$.

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 3 *If A and B are linear transformations on a finite-dimensional vector space V, then prove that*

(i) $\rho(A+B) \leqslant \rho(A) + \rho(B)$

(ii) $\rho(AB) \leqslant \min\{\rho(A), \rho(B)\}$. **(Poona 1970)**

(iii) *If B is invertible, then $\rho(AB) = \rho(BA) = \rho(A)$.*

Solution. If A is a linear transformation on V, let $R(A)$ denote the range of A. We know that $R(A)$, $R(B)$, and $R(A+B)$ are all subspaces of V. First we shall prove that $R(A+B)$ is a subspace of $R(A)+R(B)$.

Let $\alpha \in R(A+B)$. Then
$$\alpha = (A+B)(\beta) \text{ for some } \beta \in V$$
$$= A(\beta) + B(\beta).$$
But $A(\beta) \in R(A)$ and $B(\beta) \in R(B)$.

$\therefore \quad \alpha \in R(A) + R(B)$.

Thus $\quad R(A+B) \subseteq R(A) + R(B)$.

$\therefore \quad R(A+B)$ is a subspace of $R(A)+R(B)$.

$\therefore \quad \dim R(A+B) \leqslant \dim\{R(A)+R(B)\}$

i.e. $\quad \rho(A+B) \leqslant \dim\{R(A)+R(B)\}$. ...(1)

Now if W_1 and W_2 are subspaces of a finite dimensional vector space V, then
$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)$$
$$\Rightarrow \dim(W_1 + W_2) \leqslant \dim W_1 + \dim W_2$$
i.e. the dimension of the sum $\leqslant$ sum of the dimensions.

$\therefore \quad \dim\{R(A)+R(B)\} \leqslant \dim R(A) + \dim R(B)$,

Scanned with CamScanner

i.e. $\dim \{R(A)+R(B)\} \leqslant \rho(A)+\rho(B)$. ...(2)

From (1) and (2), we get

$\rho(A+B) \leqslant \rho(A)+\rho(B)$.

(ii) For every α in V, we have

$(AB)(\alpha)=A[B(\alpha)]$.

$\therefore$ the range of AB is a subset of the range of A *i.e.*,

$R(AB) \subseteq R(A)$.

$\therefore$ $R(AB)$ is a subspace of $R(A)$.

$\therefore$ $\dim R(AB) \leqslant \dim R(A)$

i.e. $\rho(AB) \leqslant \rho(A)$. ...(1)

Applying the result (1) for the transformations B' and A', we get

$\rho(B'A') \leqslant \rho(B')$

$\Rightarrow \rho[(AB)'] \leqslant \rho(B')$ $[\because (AB)'=B'A']$

$\Rightarrow \rho(AB) \leqslant \rho(B)$ $[\because \rho(B)=\rho(B')$ and $\rho[(AB)']=\rho(AB)]$

Thus $\rho(AB) \leqslant \rho(A)$, and $\rho(AB) \leqslant \rho(B)$. ...(2)

$\therefore$ $\rho(AB) \leqslant$ minimum $\{\rho(A), \rho(B)\}$.

(iii) Let B be invertible. Then we can write

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

$\therefore$ $\rho(A)=\rho[(AB)B^{-1}]$

$\leqslant \rho(AB)$ [by result (1) proved in the proof of (ii)]

Similarly, we can write

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

$\therefore$ $\rho(A)=\rho[B^{-1}(BA)]$

$\leqslant \rho(BA)$. [by result (2) proved in the proof of (ii)]

Now $\rho(A) \leqslant \rho(AB)$, and $\rho(AB) \leqslant \rho(A)$ implies that

$\rho(A)=\rho(AB)$.

Similarly $\rho(A) \leqslant \rho(BA)$, and $\rho(BA) \leqslant \rho(A)$ implies that

$\rho(A)=\rho(BA)$.

Ex. 4. *If A and B are linear transformations on an n-dimensional vector space V, then prove that*

(*i*) $\rho(AB) \geqslant \rho(A)+\rho(B)-n$.

(*ii*) $\nu(AB) \leqslant \nu(A)+\nu(B)$. **(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 $\geqslant h-\nu(T)$.

Since V is finite-dimensional, therefore the subspace W_1 will possess complement. Let $V=W_1 \oplus W_2$. Then

$\dim W_2=n-h=k$ (say).

Since $V=W_1+W_2$, therefore

$T(V)=T(W_1)+T(W_2)$, as can be easily seen.

$\therefore$ $\dim T(V)=\dim [T(W_1)+T(W_2)]$

$\leqslant \dim T(W_1)+\dim T(W_2)$ [$\because$ the dimension of a sum is $\leqslant$ the sum of the dimensions]

But $T(V)=$ the range of T.

$\therefore$ $\dim T(V)=\rho(T)$.

Thus $\dim T(W_1)+\dim T(W_2) \geqslant \rho(T)$. ...(1)

Now $T(W_2)$ is a subspace of W_2. Therefore

$\dim W_2 \geqslant \dim T(W_2)$. ...(2)

From (1) and (2), we get

$\dim T(W_1)+\dim W_2 \geqslant \rho(T)$

$\Rightarrow \dim T(W_1) \geqslant \rho(T)-\dim W_2$

$\Rightarrow \dim T(W_1) \geqslant n-\nu(T)-k$ $[\because \rho(T)+\nu(T)=n]$

$\Rightarrow \dim T(W_1) \geqslant n-k-\nu(T)$

$\Rightarrow \dim T(W_1) \geqslant h-\nu(T)$. ...(3)

Now taking $T=A$ and $W_1=B(V)$ in (3), we get

$\dim A[B(V)] \geqslant \dim B(V)-\nu(A)$

$\Rightarrow \dim (AB)(V) \geqslant \rho(V)-\nu(A)$

[$\because$ $B(V)=$ the range of B]

$\Rightarrow \rho(AB) \geqslant \rho(B)-[n-\rho(A)]$

$\Rightarrow \rho(AB) \geqslant \rho(A)+\rho(B)-n$.

(ii) We have $\rho(AB)+\nu(AB)=n$.

$\therefore$ $\rho(AB)=n-\nu(AB)$.

But $\rho(AB) \geqslant \rho(A)+\rho(B)-n$.

$\therefore$ $n-\nu(AB) \geqslant \rho(A)+\rho(B)-n$

$\Rightarrow \nu(AB) \leqslant [n-\rho(A)]+[n-\rho(B)]$

$\Rightarrow \nu(AB) \leqslant \nu(A)+\nu(B)$ $[\because \rho(A)+\nu(A)=n]$

§ 24. Characteristic Values and Characteristic Vectors.

Throughout this discussion T will be regarded as a linear operator on a finite dimensional vector space.

Definition. *Let T be a linear operator on an n-dimensional vector space V over the field F. Then a scalar* $c \in F$ *is called a* **characteristic value** *of T if there is a* **non-zero** *vector* α *in V such that* $T\alpha=c\alpha$. *Also if c is a characteristic value of T, then any non-zero vector* α *in V such that* $T\alpha=c\alpha$ *is called a* **characteristic vector** *of T* **belonging to the characteristic value** *c*.

(Meerut 1976, 79; Nagarjuna 78)

Characteristic values are sometimes also called **proper values, eigen values, or spectral values.** Similarly characteristic vectors are called **proper vectors, eigen vectors, or spectral vectors.**

The set of all characteristic values of T is called the **spectrum** of T.

Theorem 1. *If α is a characteristic vector of T corresponding to the characteristic value c, then $k\alpha$ is also a characteristic vector of T corresponding to the same characteristic value c. Here k is any non zero scalar.*

Proof. Since α is a characteristic vector of T corresponding to the characteristic value c, therefore $\alpha \neq 0$ and

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

If k is any non-zero scalar, then $k\alpha \neq 0$. ...(1)

Also $$T(k\alpha) = kT(\alpha) = k(c\alpha) = (kc)\alpha$$
$$= (ck)\alpha = c(k\alpha).$$

$\therefore$ $k\alpha$ is a characteristic vector of T corresponding to the characteristic value c.

Thus corresponding to a characteristic value c, there may correspond more than one characteristic vectors.

Theorem 2. *If α is a characteristic vector of T, then α cannot correspond to more than one characteristic values of T.*

Proof. Let α be a characteristic vector of T corresponding to two distinct characteristic values c_1 and c_2 of T. Then

$$T\alpha = c_1\alpha$$
and $$T\alpha = c_2\alpha$$
$$\therefore \quad c_1\alpha = c_2\alpha$$
$$\Rightarrow \quad (c_1 - c_2)\alpha = 0$$
$$\Rightarrow \quad c_1 - c_2 = 0 \qquad [\because \alpha \neq 0]$$
$$\Rightarrow \quad c_1 = c_2.$$

Theorem 3. *Let T be a linear operator on a finite dimensional vector space V and let c be a characteristic value of T. Then the set $W_c = \{\alpha \in V : T\alpha = c\alpha\}$ is a subspace of V.* **(Meerut 1992)**

Proof. Let $\alpha, \beta \in W_c$. Then $T\alpha = c\alpha$, and $T\beta = c\beta$.

If $a, b \in F$, then

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

$\therefore$ $a\alpha + b\beta \in W_c$.

Therefore W_c is a subspace of V.

Note. The set W_c is nothing but the set of all characteristic vectors of T corresponding to the characteristic value c provided we include the zero vector in this set. In other words W_c is the null space of the linear operator $T - cI$. The subspace W_c of V is called the **characteristic space of the characteristic value c** of the linear operator T. It is also called the space of characteristic vectors of T associated with the characteristic value c.

Theorem 4. *Distinct characteristic vectors of T corresponding to distinct characteristic values of T are linearly independent.*

(Poona 1970; Meerut 77; Nagarjuna 78)

Proof. Let $c_1, c_2, \ldots, c_m$ be m distinct characteristic values of T and let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the characteristic vectors of T corresponding to these characteristic values respectively. Then

$$T\alpha_i = c_i\alpha_i \text{ where } 1 \leqslant i \leqslant m.$$

Let $$S = \{\alpha_1, \ldots, \alpha_m\}.$$

Then to prove that the set S is linearly independent. We shall prove the theorem by induction on m, the number of vectors in S.

If $m = 1$, then S is linearly independent because S contains only one non-zero vector. Note that a characteristic vector cannot be 0 by our definition.

Now suppose that the set

$$S_1 = \{\alpha_1, \ldots, \alpha_k\}, \text{ where } k < m,$$

is linearly independent.

Consider the set $S_2 = \{\alpha_1, \ldots, \alpha_k, \alpha_{k+1}\}$.

We shall show that S_2 is linearly independent.

Let $a_1, \ldots, a_{k+1} \in F$ and let

$$a_1\alpha_1 + \ldots + a_{k+1}\alpha_{k+1} = 0 \qquad \ldots(1)$$
$$\Rightarrow T(a_1\alpha_1 + \ldots + a_{k+1}\alpha_{k+1}) = T(0)$$
$$\Rightarrow a_1 T(\alpha_1) + \ldots + a_{k+1} T(\alpha_{k+1}) = 0$$
$$\Rightarrow a_1(c_1\alpha_1) + \ldots + a_{k+1}(c_{k+1}\alpha_{k+1}) = 0 \qquad \ldots(2)$$

Multiplying (1) by the scalar c_{k+1} and subtracting from (2), we get

$$a_1(c_1 - c_{k+1})\alpha_1 + \ldots + a_k(c_k - c_{k+1})\alpha_k = 0.$$

$\therefore$ $a_1 = 0, \ldots, a_k = 0$ since $\alpha_1, \ldots, \alpha_k$ are linearly independent according to our assumption and $c_1, \ldots, c_{k+1}$ are all distinct.

Putting each of $a_1, \ldots, a_k$ equal to 0 in (1), we get

$$a_{k+1}\alpha_{k+1} = 0$$
$$\Rightarrow a_{k+1} = 0 \text{ since } \alpha_{k+1} \neq 0.$$

Thus the relation (1) implies that

$$a_1 = 0, \ldots, a_k = 0, a_{k+1} = 0.$$

$\therefore$ the set S_2 is linearly independent.

Now the proof is complete by induction.

Corollary. *If T is a linear operator on an n-dimensional vector*

Scanned with CamScanner

Proof We have
$$[T-cI]_B=[T]_B-c[I]_B=A-cI.$$
If $\alpha \neq 0$, then the coordinate vector X of α is also non-zero.

Now $[(T-cI)(\alpha)]_B=[T-cI]_B\,[\alpha]_B$ [See theorem 2 of § 13]

$=(A-cI)\ X.$

$\therefore\ (T-cI)\ (\alpha)=0$ iff $(A-cI)\ X=O$

or $T\ (\alpha)=c\alpha$ iff $AX=cX$

or α is an eigenvector of T iff X is an eigenvector of A.

Thus with the help of this theorem we see that our definition of characteristic vector of a matrix is sensible Now we shall define the characteristic polynomial of a linear operator. Before doing so we shall prove the following theorem.

Theorem 7. *Similar matrices A and B have the same characteristic polynomial and hence the same eigenvalues. If X is an eigenvector of A corresponding to the eigenvalue c, then* $P^{-1}X$ *is an eigenvector of B corresponding to the eigenvalue c where*
$$B=P^{-1}AP.$$
(Meerut 1976, 83, 92)

Proof. Suppose A and B are similar matrices. Then there exists an invertible matrix P such that
$$B=P^{-1}AP.$$
We have $B-xI=P^{-1}AP-xI$

$=P^{-1}AP-P^{-1}(xI)P$ $\quad [\because\ P^{-1}(xI)P=xP^{-1}IP=xI]$

$=P^{-1}(A-xI)P.$

$\therefore\ \det(B-xI)=\det P^{-1}\det(A-xI)\det P$

$=\det P^{-1}.\det P.\det(A-xI)=\det(P^{-1}P).\det(A-xI)$

$=\det I.\det(A-xI)=1.\det(A-xI)=\det(A-xI).$

Thus the matrices A and B have the same characteristic polynomial and consequently they will have the same characteristic values.

If c is an eigenvalue of A and X is a corresponding eigenvector, then $AX=cX$, and hence
$$B(P^{-1}X)=(P^{-1}AP)P^{-1}X=P^{-1}AX=P^{-1}(cX)=c(P^{-1}X).$$
$\therefore\ P^{-1}X$ is an eigenvector of B corresponding to c. This completes the proof of the theorem.

Now suppose that T is a linear operator on an n-dimensional vector space V. If B_1, B_2 are any two ordered bases for V, then we know that the matrices $[T]_{B_1}$ and $[T]_{B_2}$ are similar. Also similar matrices have the same characteristic polynomial. This enables us to define sensibly the characteristic polynomial of T as follows :

Characteristic polynomial of a linear operator.

Definition. *Let T be a linear operator on an n-dimensional vector space V. The characteristic polynomial of T is the characteristic polynomial of any* $n\times n$ *matrix which represents T in some ordered basis for V.* On account of the above discussion the polynomial of T as defined by us will be unique.

If B is any ordered basis for V and A is the matrix of T with respect to B then

$\det(T-xI)=\det[T-xI]_B=\det([T]_B-x[I]_B)$

$=\det(A-xI)$

$=$ the characteristic polynomial of A and so also that of T.

$\therefore$ the characteristic polynomial of $T=\det(T-xI)$

The equation $\det(T-xI)=0$ is called the characteristic equation of T.

Existence of characteristic values. Let T be a linear operator on an n-dimensional vector space V over the field F. Then c belonging to F will be a characteristic value of T iff c is a root of the characteristic equation of T *i e.* iff c is a root of the equation
$$\det(T-xI)=0. \qquad ...(1)$$
The equation (1) is of degree n in x. If the field F is *algebraically closed i.e.* if every polynomial equation in F possesses a root then T will definitely have at least one characteristic value. If the field F is not algebraically closed, then T may or may not have a characteristic value according as the equation (1) has or has not a root in F. Since the equation (1) is of degree n in x, therefore if T has a characteristic value then it cannot have more than n distinct characteristic values. The field of complex numbers is algebraically closed. By fundamental theorem of algebra we know that every polynomial equation over the field of complex numbers is solvable. Therefore if F is the field of complex numbers then T will definitely have at least one characteristic value. The field of real numbers is not algebraically closed. If F is the field of real numbers, then T may or may not have a characteristic value.

Example. Consider the linear operator T on V_2 (R) which is represented in the standard ordered basis by the matrix
$$A=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}.$$
The characteristic polynomial for T (or for A) is

Scanned with CamScanner

space V, then T cannot have more than n distinct characteristic values.

Proof. Suppose T has more than n distinct characteristic values. Then the corresponding set of distinct characteristic vectors of T will be linearly independent. Thus we shall have a linearly independent subset of V containing more than n vectors which is not possible because V is of dimension n. Hence T cannot have more than n distinct characteristic values.

Theorem 5. *Let T be a linear operator on a finite-dimensional vector space V. Then the following are equivalent.*

(i) *c is a characteristic value of T.*

(ii) *The operator $T-cI$ is singular (not invertible).*

(iii) *$\det(T-cI)=0$.*

Proof. (i) $\Rightarrow$ (ii). **(Meerut 1979)**

c is a characteristic value of T implies that there exists a non-zero vector α in V such that

$$T\alpha=c\alpha$$

or $\quad T\alpha=cI\alpha$ where I is the identity operator on V

or $\quad T\alpha=(cI)\,\alpha$

or $\quad (T-cI)\,\alpha=\mathbf{0}$.

Thus $(T-cI)\,\alpha=\mathbf{0}$ while $\alpha\neq\mathbf{0}$. Therefore the operator $T-cI$ is singular and thus $T-cI$ is not invertible.

(ii) $\Rightarrow$ (iii).

If the operator $T-cI$ is singular, then it is not invertible. Therefore $\det(T-cI)=0$.

(iii) $\Rightarrow$ (i).

If $\det(T-cI)=0$, then $T-cI$ is not invertible.

If $T-cI$ is not invertible, then $T-cI$ is singular because every non-singular operator on a finite-dimensional vector space is invertible. Now $T-cI$ is singular means that there is a non-zero vector α in V such that

$$(T-cI)\,\alpha=\mathbf{0}$$

or $$T\alpha-cI\,\alpha=\mathbf{0}$$

or $$T\alpha=c\alpha.$$

$\therefore$ c is a characteristic value of T.

This completes the proof of the theorem.

Let T be a linear operator on an n-dimensional vector space V. Let B be an ordered basis for V and let A be the matrix of T with respect to B i.e. let $A=[T]_B$. If c is any scalar, we have

$$[T-cI]_B=[T]_B-c\,[I]_B$$
$$=A-cI \text{ where } I \text{ is the unit matrix of order } n.$$

[Note that $[I]_B=I$].

We have $\quad \det(T-cI)=\det[T-cI]_B$
$$=\det(A-cI).$$

Therefore c is a characteristic value of T iff $\det(A-cI)=0$. This enables us to make the following definition.

Characteristic values of a matrix. Definition.

Let $A=[a_{ij}]_{n\times n}$ be a square matrix of order n over the field F. An element c in F is called a characteristic value of A if $\det(A-cI)=0$ where I is the unit matrix of order n.

Now suppose T is a linear operator on an n-dimensional vector space V and A is the matrix of T with respect to any ordered basis B. Then c is a characteristic value of T iff c is a characteristic value of the matrix A. Therefore our definition of characteristic values of a matrix is sensible.

Characteristic equation of a matrix Definition.

Let A be a square matrix of order n over the field F. Consider the matrix $A-xI$. The elements of this matrix are polynomials in x of degree at most 1. If we evaluate $\det(A-xI)$, then it will be a polynomial in x of degree n. The coefficients of x in this polynomial will be $(-1)^n$. Let us denote this polynomial by $f(x)$.

Then $f(x)=\det(A-xI)$ is called the **characteristic polynomial** of the matrix A. The equation $f(x)=0$ is called the **characteristic equation** of the matrix A. Now c is a characteristic value of the matrix A iff $\det(A-cI)=0$ *i.e.* iff $f(c)=0$ *i.e.* iff c is a root of the characteristic equation of A. Thus in order to find the characteristic values of a matrix we should first obtain its characteristic equation and then we should find the roots of this equation.

Characteristic vector of a matrix.

Definition. *If c is a characteristic value of an $n\times n$ matrix A, then a non-zero matrix X of the type $n\times 1$ such that $AX=cX$ is called a characteristic vector of A corresponding to the characteristic value c.*

Theorem 6. *Let T be a linear operator on an n-dimensional vector space V and A be the matrix of T relative to any ordered basis B. Then a vector α in V is an eigenvector of T corresponding to its eigenvalue c if and only if its coordinate vector X relative to the basis B is an eigen-vector of A corresponding to its eigenvalue c.*

Scanned with CamScanner

Proof We have

$$[T-cI]_B=[T]_B-c[I]_B=A-cI.$$

If $\alpha \neq 0$, then the coordinate vector X of α is also non-zero.

Now $$[(T-cI)(\alpha)]_B=[T-cI]_B\,[\alpha]_B$$ [See theorem 2 of § 13]

$$=(A-cI)\,X.$$

$\therefore$ $(T-cI)(\alpha)=0$ iff $(A-cI)\,X=O$

or $T(\alpha)=c\alpha$ iff $AX=cX$

or α is an eigenvector of T iff X is an eigenvector of A.

Thus with the help of this theorem we see that our definition of characteristic vector of a matrix is sensible. Now we shall define the characteristic polynomial of a linear operator. Before doing so we shall prove the following theorem.

Theorem 7. *Similar matrices A and B have the same characteristic polynomial and hence the same eigenvalues. If X is an eigenvector of A corresponding to the eigenvalue c, then $P^{-1}X$ is an eigenvector of B corresponding to the eigenvalue c where*

$$B=P^{-1}AP.$$ **(Meerut 1976, 83, 92)**

Proof. Suppose A and B are similar matrices. Then there exists an invertible matrix P such that

$$B=P^{-1}AP.$$

We have $B-xI=P^{-1}AP-xI$

$=P^{-1}AP-P^{-1}(xI)\,P$ $\quad [\because P^{-1}(xI)\,P=xP^{-1}IP=xI]$

$=P^{-1}(A-xI)\,P.$

$\therefore$ $\det(B-xI)=\det P^{-1}\det(A-xI)\det P$

$=\det P^{-1}.\det P.\det(A-xI)=\det(P^{-1}P).\det(A-xI)$

$=\det I.\det(A-xI)=1.\det(A-xI)=\det(A-xI).$

Thus the matrices A and B have the same characteristic polynomial and consequently they will have the same characteristic values.

If c is an eigenvalue of A and X is a corresponding eigenvector, then $AX=cX$, and hence

$$B(P^{-1}X)=(P^{-1}AP)\,P^{-1}X=P^{-1}AX=P^{-1}(cX)=c(P^{-1}X).$$

$\therefore$ $P^{-1}X$ is an eigenvector of B corresponding to c. This completes the proof of the theorem.

Now suppose that T is a linear operator on an n-dimensional vector space V. If B_1, B_2 are any two ordered bases for V, then we know that the matrices $[T]_{B_1}$ and $[T]_{B_2}$ are similar. Also similar matrices have the same characteristic polynomial. This enables us to define sensibly the characteristic polynomial of T as follows:

Characteristic polynomial of a linear operator.

Definition. *Let T be a linear operator on an n-dimensional vector space V. The characteristic polynomial of T is the characteristic polynomial of any* $n \times n$ *matrix which represents T in some ordered basis for V.* On account of the above discussion the polynomial of T as defined by us will be unique.

If B is any ordered basis for V and A is the matrix of T with respect to B then

$\det(T-xI)=\det[T-xI]_B=\det([T]_B-x[I]_B)$

$=\det(A-xI)$

$=$ the characteristic polynomial of A and so also that of T.

$\therefore$ the characteristic polynomial of $T=\det(T-xI)$

The equation $\det(T-xI)=0$ is called the characteristic equation of T.

Existence of characteristic values. Let T be a linear operator on an n-dimensional vector space V over the field F. Then c belonging to F will be a characteristic value of T iff c is a root of the characteristic equation of T *i.e.* iff c is a root of the equation

$$\det(T-xI)=0. \qquad ...(1)$$

The equation (1) is of degree n in x. If the field F is *algebraically closed i.e.* if every polynomial equation in F possesses a root then T will definitely have at least one characteristic value. If the field F is not algebraically closed, then T may or may not have a characteristic value according as the equation (1) has or has not a root in F. Since the equation (1) is of degree n in x, therefore if T has a characteristic value then it cannot have more than n distinct characteristic values. The field of complex numbers is algebraically closed. By fundamental theorem of algebra we know that every polynomial equation over the field of complex numbers is solvable. Therefore if F is the field of complex numbers then T will definitely have at least one characteristic value. The field of real numbers is not algebraically closed. If F is the field of real numbers, then T may or may not have a characteristic value.

Example. Consider the linear operator T on $V_2(\mathbf{R})$ which is represented in the standard ordered basis by the matrix

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

The characteristic polynomial for T (or for A) is

Scanned with CamScanner

$$\det(A-xI)=\det\begin{bmatrix}0-x & -1\\ 1 & 0-x\end{bmatrix}=\begin{vmatrix}-x & -1\\ 1 & -x\end{vmatrix}=x^2+1.$$

The polynomial equation $x^2+1=0$ has no roots in **R**. Therefore T has no characteristic values.

However if T is a linear operator on $V_2(C)$, then the characteristic equation of T has two distinct roots i and $-i$ in C. In this case T has two characteristic values i and $-i$.

Algebraic and geometric multiplicity of a characteristic value. **Definition.** Let T be a linear operator on an n-dimensional vector space V and let c be a characteristic value of T. By *geometric multiplicity* of c we mean the dimension of the characteristic space W_c of c. By *algebraic multiplicity* of c we mean the multiplicity of c as root of the characteristic equation of T.

Method of finding the characteristic values and the corresponding characteristic vectors of a linear operator T. Let T be a linear operator on an n-dimensional vector space V over the field F. Let B be any ordered basis for V and let $A=[T]_B$. The roots of the equation $\det(A-xI)=0$ will give the characteristic values of A or also of T. Let c be a characteristic value of T. Then $0\neq\alpha$ will be a characteristic vector corresponding to this characteristic value if

$$(T-cI)\,\alpha=\mathbf{0}$$

i.e. if $\quad [T-cI]_B\,[\alpha]_B=[\mathbf{0}]_B$

i.e. if $\quad (A-cI)\,X=O, \qquad \ldots(1)$

where $X=[\alpha]_B=$ a column matrix of the type $n\times 1$ and O is the null matrix of the type $n\times 1$. Thus to find the coordinate matrix of α with respect to B, we should solve the matrix equation (1) for X.

Matric Polynomials. Definition *An expression of the form*

$$f(x)=A_0+A_1x+A_2x^2+\ldots+A_mx^m,$$

where $A_0, A_1, A_2, \ldots, A_m$ are all square matrices of order n, is called a Matric polynomial of degree m provided A_m is not a null matrix. The symbol x is called **indeterminate.**

Equality of Matric Polynomials. Two matric polynomials are equal iff the coefficients of the like powers of x are the same.

Lemma. *Every square matrix over the field F whose elements are ordinary polynomials in x over F, can essentially be expressed as a matric polynomial in x of degree m, where m is the highest power of x occurring in any element of the matrix.*

We shall illustrate this theorem by the following example :

Consider the matrix

$$A=\begin{bmatrix}1+2x+3x^2 & x^2 & 4-6x\\ 1+x^3 & 3+4x^2 & 1-2x+4x^3\\ 2-3x+2x^3 & 5 & 6\end{bmatrix}$$

in which the highest power of x occurring in any element is 3. Rewriting each element as a cubic in x, supplying missing coefficients with zeros, we get

$$A=\begin{bmatrix}1+2.x+3\,x^2+0\,x^3 & 0+0.x+1.x^2+0.x^3 & 4-6.x+0\,x^2+0\,x^3\\ 1+0.x+0.x^2+1.x^3 & 3+0\,x+4\,x^2+0.x^3 & 1-2\,x+0\,x^2+4.x^3\\ 2-3.x+0.x^2+2.x^3 & 5+0\,x+0.x^2+0.x^3 & 6+0.x+0.x^2+0.x^3\end{bmatrix}$$

Obviously A can be written as the matrix polynomial

$$A=\begin{bmatrix}1 & 0 & 4\\ 1 & 3 & 1\\ 1 & 5 & 6\end{bmatrix}+x\begin{bmatrix}2 & 0 & -6\\ 0 & 0 & -2\\ -3 & 0 & 0\end{bmatrix}+x^2\begin{bmatrix}3 & 1 & 0\\ 0 & 4 & 0\\ 0 & 0 & 0\end{bmatrix}+x^3\begin{bmatrix}0 & 0 & 0\\ 1 & 0 & 4\\ 2 & 0 & 0\end{bmatrix}.$$

Theorem 8. The Cayley-Hamilton Theorem.

Let T be a linear operator on an n-dimensional vector space V(F). Then T satisfies its characteristic equation i.e. if f(x) be the characteristic polynomial of T, then $f(T)=\hat{0}$.

Or

Every square matrix satisfies its characteristic equation.

(Meerut 1980, 81, 82, 85, 87, 92; Banaras 73; Poona 70; Raj 65; G.N D.U. Amritsar 90; Andhra 92; S.V.U. Tirupati 90)

Proof. Let T be a linear operator on an n-dimensional vector space V over the field F. Let B be any ordered basis for V and A be the matrix of T relative to B i.e. let $A=[T]_B$. The characteristic polynomial of T is the same as the characteristic polynomial of A. If $A=[a_{ij}]_{n\times n}$, then the characteristic polynomial $f(x)$ of A is given by

$$f(x)=\det(A-xI)=\begin{vmatrix}a_{11}-x & a_{12} & \ldots a_{1n}\\ a_{21} & a_{22}-x & \ldots a_{2n}\\ \ldots\ldots & \ldots\ldots & \ldots\ldots\\ \ldots\ldots & \ldots\ldots & \ldots\ldots\\ a_{n1} & a_{n2} & a_{nn}-x\end{vmatrix} \qquad \ldots(1)$$

$$=a_0+a_1x+a_2x^2+\ldots+a_nx^n \text{ (say)},$$

where the a_i are in F.

The characteristic equation of A is $f(x)=0$

i.e. $\qquad a_0+a_1x+a_2x^2+\ldots+a_nx^n=0.$

Since the elements of the matrix $A-xI$ are polynomials at most of the first degree in x, therefore the elements of the matrix adj $(A-xI)$ are ordinary polynomials in x of degree $n-1$ or less. Note that the elements of the matrix adj $(A-xI)$ are the cofactors of the elements of the matrix $A-xI$. Therefore adj $(A-xI)$ can be written as a matrix polynomial in x in the form

$$\text{adj}\,(A-xI)=B_0+B_1x+B_2x^2+\ldots+B_{n-1}x^{n-1}, \qquad \ldots(2)$$

where the B_i's are square matrices of order n over F with elements independent of x.

Now by the property of adjoints, we know that

$$(A-xI).\ \text{adj.}\ (A-xI)=\{\det.\ (A-xI)\}\ I.$$

$$\therefore\quad (A-xI)\ \{B_0+x\ B_1+x^2B_2+\ldots+x^{n-1}\ B_{n-1}\}$$
$$=\{a_0+a_1x+\ldots+a_nx^n\}\ I \qquad \text{[from (1) and (2)]}$$

Equating the coefficients of like powers of x on both sides, we get

$$AB_0=a_0I$$
$$AB_1-IB_0=a_1I$$
$$AB_2-IB_1=a_2I$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$AB_{n-1}-IB_{n-2}=a_{n-1}\ I$$
$$-IB_{n-1}=a_nI.$$

Premultiplying these equations successively by $I, A, A^2,\ldots,A^n$ and adding, we get

$$a_0I+a_1A+a_2A^2+\ldots+a_nA^n=O,$$

where O is the null matrix of order n.

Thus $\qquad f(A)=O.$

Now $\quad f(T)=a_0I+a_1T+a_2T^2+\ldots+a_nT^n.$

$$\therefore\quad [f(T)]_B=[a_0I+a_1T+a_2T^2+\ldots+a_nT^n]_B$$
$$=a_0[I]_B+a_1\ [T]_B+a_2\ [T^2]_B+\ldots+a_n\ [T^n]_B$$
$$=f(A).$$

$$\therefore\quad f(A)=O$$

$$\Rightarrow\quad [f(T)]_B=O=[\hat{0}]_B$$

$$\Rightarrow\quad f(T)=\hat{0}$$

$$\Rightarrow\quad a_0I+a_1T+a_2T^2+\ldots+a_nT^n=\hat{0}. \qquad \ldots(3)$$

Corollary. We have

$$f(x)=a_0+a_1x+\ldots+a_nx^n=\det.\ (A-xI).$$

$$\therefore\quad f(0)=a_0=\det.\ A=\det\ T.$$

If T is non-singular, then T is invertible and

$$\det.\ T\neq 0\ \text{i.e.},\ a_0\neq 0.$$

Then from (3), we get

$$a_0I=-(a_1T+a_2T^2+\ldots+a_nT^n)$$

$$\Rightarrow\quad I=-\left(\frac{a_1}{a_0}\ I+\frac{a_2}{a_0}\ T+\ldots+\frac{a_n}{a_0}\ T^{n-1}\right)T$$

$$\Rightarrow\quad T^{-1}=-\left(\frac{a_1}{a_0}\ I+\frac{a_2}{a_0}\ T+\ldots+\frac{a_n}{a_0}\ T^{n-1}\right).$$

Diagonalizable Operators. Definition. *Suppose T is a linear operator on the finite-dimensional vector space V. Then T is said to be diagonalizable if there is a basis B for V each vector of which is a characteristic vector of T.*

Matrix of a diagonalizable operator. [Meerut 1976] Let T be a diagonalizable operator on an n-dimensional vector space V. Let $B=\{\alpha_1,\ldots,\alpha_n\}$ be an ordered basis for V such that each α_j is a characteristic vector of T. Let $T\alpha_i=c_i\alpha_i$. Then

$$T\alpha_1=c_1\alpha_1=c_1\alpha_1+0\alpha_2+\ldots+0\alpha_n$$
$$T\alpha_2=c_2\alpha_2=0\alpha_1+c_2\alpha_2+\ldots+0\alpha_n$$
$$\ldots\quad\ldots\quad\ldots\quad\ldots\quad\ldots$$
$$\ldots\quad\ldots\quad\ldots\quad\ldots\quad\ldots$$
$$T\alpha_n=c_n\alpha_n=0\alpha_1+0\alpha_2+\ldots+0\alpha_{n-1}+c_n\alpha_n.$$

Therefore the matrix of T relative to B is

$$[T]_B=\begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & c_n \end{bmatrix}.$$

This matrix is a diagonal matrix. Note that a square matrix of order n is said to be a diagonal matrix if all the elements lying above and below the principal diagonal are equal to zero. The scalars $c_1,\ldots,c_n$ need not all be distinct. If V is n dimensional, then T is diagonalizable iff T has n linearly independent characteristic vectors.

Diagonalizable Matrix. Definition. *A matrix A over a field F is said to be diagonalizable if it is similar to a diagonal matrix over the field F.* 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 9. *A necessary and sufficient condition that an $n\times n$ matrix A over a field F be diagonalizable is that A has n linearly independent characteristic vectors in $V_n(F)$.*

Proof. If A is diagonalizable, then A is similar to a diagonal matrix D. Therefore there exists an invertible matrix P such that

$$P^{-1}AP=D \qquad \ldots(1)$$

or $$AP=PD.$$

If $c_1, c_2,\ldots,c_n$ are the diagonal elements of D, then $c_1, c_2,\ldots,c_n$ are the characteristic values of D as can be easily seen. But similar matrices have the same characteristic values. Therefore $c_1, c_2,\ldots,c_n$ are the characteristic values of A.

Scanned with CamScanner

Now suppose $P_1, P_2, \ldots, P_n$ are the column vectors of the matrix P. Then equating corresponding columns on each side of (1), we get

$$AP_i = c_i P_i \ (i=1, 2, \ldots, n). \qquad \ldots(2)$$

But (2) shows that P_i is a characteristic vector of A corresponding to the characteristic value c_i. Since the matrix P is invertible, therefore its column vectors $P_1, P_2, \ldots, P_n$ are n linearly independent vectors belonging to $V_n(F)$. Thus A has n linearly independent characteristic vectors $P_1, \ldots, P_n$.

Conversely, if $P_1, \ldots, P_n$ are n linearly independent characteristic vectors of A corresponding to characteristic values $c_1, \ldots, c_n$, then equations (2) hold. Therefore equation (1) holds where P is the matrix with columns $P_1, \ldots, P_n$. Since the columns of P are linearly independent, therefore P is invertible and hence (1) implies $P^{-1}AP=D$. Thus A is similar to a diagonal matrix and so is diagonalizable.

This completes the proof of the theorem.

Remark. In the proof of the above theorem we have shown that if A is diagonalizable and P diagonalizes A, then

$$P^{-1}AP = \begin{bmatrix} c_1 & 0 & \ldots & 0 \\ 0 & c_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & c_n \end{bmatrix}$$

if and only if the j^{th} column of P is a characteristic vector of A corresponding to the characteristic value c_j of A, $(j=1, 2, \ldots, n)$.

Theorem 10. *A linear operator T on an n-dimensional vector space $V(F)$ is diagonalizable if and only if its matrix A relative to any ordered basis B of V is diagonalizable.*

Proof. Suppose T is diagonalizable. Then T has n linearly independent characteristic vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ in V. Suppose $X_1, X_2, \ldots, X_n$ are the co-ordinate vectors of $\alpha_1, \alpha_2, \ldots, \alpha_n$ relative to the basis B. Then $X_1, \ldots, X_n$ are also linearly independent since V is isomorphic to $V_n(F)$ by isomorphism which takes a vector in V to its co-ordinate vector in $V_n(F)$. Under an isomorphism a linearly independent set is mapped onto a linearly independent set. Further $X_1, \ldots, X_n$ are the characteristic vectors of the matrix A [see theorem 6]. Therefore the matrix A is diagonalizable. [See theorem 9].

Conversely suppose the matrix A is diagonalizable. Then A has n linearly independent characteristic vectors $X_1, \ldots, X_n$ in $V_n(F)$. If $\alpha_1, \ldots, \alpha_n$ are the vectors in V having $X_1, \ldots, X_n$ as their coordinate vectors, then $\alpha_1, \ldots, \alpha_n$ will be n linearly independent characteristic vectors of T. So T is diagonalizable.

Theorem 11. *Let T be any linear operator on a finite dimensional vector space V, let $c_1, c_2, \ldots, c_k$ be the distinct characteristic values of T, and let W_i be the null space of $(T-c_iI)$. Then the subspaces $W_1, \ldots, W_k$ are independent.* (Meerut 1972, 74, 77)

Further show that if in addition T is diagonalizable, then V is the direct sum of the subspaces $W_1, \ldots, W_k$. (Meerut 1993, 93P)

Proof. By definition of W_i, we have

$$W_i = \{\alpha : \alpha \in V \text{ and } (T-c_iI)\,\alpha = 0 \text{ i.e. } T\alpha = c_i\alpha\}.$$

Now let α_i be in W_i, $i=1, \ldots, k$, and suppose that

$$\alpha_1 + \alpha_2 + \ldots + \alpha_k = 0. \qquad \ldots(1)$$

Let j be any integer between 1 and k and let

$$U_j = \prod_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} (T - c_iI).$$

Note that U_j is the product of the operators $(T-c_iI)$ for $i \neq j$. In other words $U_j = (T-c_1I)(T-c_2I)\ldots(T-c_kI)$ where in the product the factor $T-c_jI$ is missing.

Let us find $U_j\alpha_i$, $i=1, \ldots, k$. By the definition of W_i, we have $(T-c_iI)\,\alpha_i = 0$. Since the operators $(T-c_iI)$ all commute, being polynomials in T, therefore $U_j\alpha_i = 0$ for $i \neq j$. Note that for each $i \neq j$, U_j contains a factor $(T-c_iI)$ and $(T-c_iI)\,\alpha_i = 0$.

Also

$$U_j\alpha_j = [(T-c_1I)\ldots(T-c_kI)]\,\alpha_j$$
$$= [(T-c_1I)\ldots(T-c_{k-1}I)]\,[T\alpha_j - c_kI\alpha_j]$$
$$= [(T-c_1I)\ldots(T-c_{k-1}I)]\,(c_j\alpha_j - c_k\alpha_j) \qquad [\because T\alpha_j = c_j\alpha_j \text{ and } I\alpha_j = \alpha_j]$$
$$= [(T-c_1I)\ldots(T-c_{k-1}I)]\,(c_j - c_k)\,\alpha_j$$
$$= (c_j - c_k)\,[(T-c_1I)\ldots(T-c_{k-1}I)]\,\alpha_j$$
$$= (c_j - c_k)(c_j - c_{k-1})\ldots(c_j - c_1)\,\alpha_j,$$ the factor $c_j - c_j$ will be missing. Thus $\qquad \ldots(2)$

$$U_j\alpha_j = \left[\prod_{\substack{1 \leqslant i \leqslant k \\ i \neq j}} (c_j - c_i)\right]\alpha_j.$$

Scanned with CamScanner

Now applying U_j to both sides of (1), we get

$$U_j\alpha_1+U_j\alpha_2+\dots+U_j\alpha_k=\mathbf{0} \qquad [\because\ U_j\alpha_i=\mathbf{0} \text{ if } i\neq j]$$

$$\Rightarrow U_j\alpha_j=\mathbf{0} \qquad [\text{by } (2)]$$

$$\Rightarrow \left[\prod_{i\neq j}(c_j-c_i)\right]\alpha_j=\mathbf{0}$$

Since the scalars c_i are all distinct, therefore the product

$$\prod_{i\neq j}(c_j-c_i)$$

is a non-zero scalar. Hence $\left[\prod_{i\neq j}(c_j-c_i)\right]\alpha_j=\mathbf{0}$

$\Rightarrow \alpha_j=\mathbf{0}$. Thus $\alpha_j=\mathbf{0}$ for every integer j between 1 and k.

In this way $\alpha_1+\dots+\alpha_k=\mathbf{0}$

$\Rightarrow \alpha_i=\mathbf{0}$ for each i. Hence the subspaces $W_1,\dots,W_k$ are independent.

Second Part. Now suppose that T is diagonalizable. Then we shall show that $V=W_1+\dots+W_k$. Since T is diagonalizable, therefore there exists a basis of V each vector of which is a characteristic vector of T. Thus there exists a basis of V consisting of vectors belonging to the characteristic subspaces $W_1,\dots,W_k$. If $\alpha\in V$, then α can be expressed as a linear combination of these basis vectors. Thus α can be written as $\alpha=\alpha_1+\dots+\alpha_k$ where $\alpha_i\in W_i$, $i=1,\dots,k$. In this way $\alpha\in W_1+\dots+W_k$. Therefore $V=W_1+\dots+W_k$. But in the first part, we have proved that the subspaces $W_1,\dots,W_k$ are independent. Hence

$$V=W_1\oplus\dots\oplus W_k.$$

Theorem 12. *If T is a diagonalizable operator on a finite dimensional vector space V, and $c_1,\dots,c_k$ are the distinct characteristic values of T, then there are linear operators $E_1,\dots,E_k$ on V such that*

(*a*) $T=c_1E_1+\dots+c_kE_k$;

(*b*) $I=E_1+\dots+E_k$;

(*c*) $E_iE_j=\hat{\mathbf{0}}$, $i\neq j$;

(*d*) $E_i^2=E_i$;

(*e*) *the range of E_i is the space of characteristic vectors of T associated with the characteristic value c_i.* **(Meerut 1984P)**

Proof. Let W_i be the null space of the operator $T-c_iI$, for $i=1,\dots,k$. Then $W_1,\dots,W_k$ are the characteristic spaces of the characteristic values $c_1,\dots,c_k$ respectively. By theorem 11, V is

the direct sum of the subspaces $W_1,\dots,W_k$. If $\alpha\in V$, then α can be uniquely written as

$$\alpha=\alpha_1+\dots+\alpha_k \text{ where each } \alpha_i\in W_i,\ 1\leqslant i\leqslant k.$$

Now let E_i be a function from V into V defined by the rule

$$E_i(\alpha)=\alpha_i.$$

Then E_i is a linear transformation on V. For its proof and for the proofs of parts (b), (c), (d), and (e) of this theorem consult theorem 5 on page 223.

Now it remains to prove the part (a) of this theorem. Let $\alpha\in V$ and let

$$\alpha=\alpha_1+\dots+\alpha_k, \text{ where } \alpha_i\in W_i \text{ for } i=1,\dots,k.$$

We have

$$T\alpha=T(\alpha_1+\dots+\alpha_k)$$

$$=T\alpha_1+\dots+T\alpha_k$$

$$=c_1\alpha_1+\dots+c_k\alpha_k \qquad [\because\ \alpha_i\in W_i \text{ and by def. of } W_i, \text{ we have } T\alpha_i=c_i\alpha_i]$$

$$=c_1E_1\alpha+\dots+c_kE_k\alpha \qquad [\text{by def. of } E_i, \text{ we have } E_i\alpha=\alpha_i]$$

$$=(c_1E_1+\dots+c_kE_k)\,\alpha.$$

Thus we have $T\alpha=(c_1E_1+\dots+c_kE_k)\,\alpha\ \ \forall\ \alpha\in V$. Hence $T=c_1E_1+\dots+c_kE_k$.

Solved Examples

Example 1. *Let V be a n-dimensional vector space over F. What is the characteristic polynomial of (i) the identity operator on V, (ii) the zero operator on V.* **(Meerut 1981)**

Solution. Let B be any ordered basis for V.

(*i*) If I is the identity operator on V, then

$$[I]_B=I.$$

The characteristic polynomial of $I=\det(I-xI)$

$$=\begin{vmatrix} 1-x & 0 & \dots & 0 \\ 0 & 1-x & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1-x \end{vmatrix}=(1-x)^n.$$

(ii) If $\hat{\mathbf{0}}$ is the zero operator on V, then $[\hat{\mathbf{0}}]_B=O$ i.e. the null matrix of order n.

Scanned with CamScanner

The characteristic polynomial of $\hat{0}=\det(O-xI)$

$$=\begin{vmatrix} -x & 0 & \dots & 0 \\ 0 & -x & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & -x \end{vmatrix}=(-1)^n x^n.$$

Example 2. *Let T be a linear operator on a finite dimensional vector space V and let c be a characteristic value of T. Show that the characteristic space of c i.e., W_c is invariant under T.*

Solution. We have by definition,

$$W_c=\{\alpha\in V : T\alpha=c\alpha\}.$$

Let $\alpha\in W_c$. Then $T\alpha=c\alpha$.

Since W_c is a subspace, therefore

$$c\in F \text{ and } \alpha\in W_c \Rightarrow c\alpha\in W_c.$$

Thus $\alpha\in W_c \Rightarrow T\alpha\in W_c$.

Hence W_c is invariant under T.

Example 3. *If T be a linear operator on a finite dimensional vector space V and c be a characteristic value of T, then show that the characteristic space of c i.e., W_c is the null space of the operator $T-cI$.*

Solution. Let $\alpha\in W_c$. Then

$$T\alpha=c\alpha$$

$$\Rightarrow (T-cI)\,\alpha=0$$

$$\Rightarrow \alpha\in \text{the null space of } T-cI.$$

Again let $\alpha\in$ the null space of $T-cI$.

Then $(T-cI)\,\alpha=0$

$$\Rightarrow T\alpha=c\alpha$$

$$\Rightarrow \alpha\in W_c.$$

Hence W_c=the null space of $T-cI$.

Example 4. *Show that the characteristic values of a diagonal matrix are precisely the elements in the diagonal. Hence show that if a matrix B is similar to a diagonal matrix D, then the diagonal elements of D are the characteristic values of B.*

Solution. Let

$$A=\begin{bmatrix} a_{11} & 0 & 0 & \dots & 0 \\ 0 & a_{22} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

be a diagonal matrix of order n. The characteristic equation of A is

$$\det(A-xI)=0$$

i.e. $$(a_{11}-x)(a_{22}-x)\dots(a_{nn}-x)=0$$

whose roots are $x=a_{11},\dots,a_{nn}$ i.e., the diagonal elements.

Now we know that similar matrices have the same characteristic values. Hence if B is similar to D, then the eigen values of B are the diagonal elements of D.

Example 5. *Let T be a linear operator on a finite dimensional vector space V. Then show that 0 is a characteristic value of T iff T is not invertible.*

Solution. Suppose 0 is a characteristic value of T. Then there exists a non-zero vector α in V such that

$$T\alpha=0\alpha$$

$$\Rightarrow T\alpha=0.$$

$\therefore$ T is singular and so T is not invertible.

Conversely suppose that T is not invertible. Since T is a linear operator on a finite-dimensional vector space V, therefore T is not invertible means that T is singular. Thus there exists a non-zero vector α in V such that

$$T\alpha=0=0\alpha.$$

$\therefore$ 0 is a characteristic value of T.

Example 6. *If c is a characteristic value of an invertible transformation T, then show that c^{-1} is a characteristic value of T^{-1}.*

Solution Since T is invertible, therefore $c\neq 0$. So c^{-1} exists.

Now c is a characteristic value of T. Therefore there exists a non-zero vector α in V such that

$$T\alpha=c\alpha$$

$$\Rightarrow T^{-1}(T\alpha)=T^{-1}(c\alpha) \Rightarrow (T^{-1}T)\,\alpha=c\,T^{-1}(\alpha)$$

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

$$\Rightarrow T^{-1}(\alpha)=c^{-1}\alpha,\ \alpha\neq 0.$$

$\therefore$ c^{-1} is a characteristic value of T^{-1}.

Example 7. *If $c\in F$ is a characteristic value of a linear operator T on a vector space $V(F)$, then for any polynomial $p(x)$ over F, $p(c)$ is a characteristic value of $p(T)$.*

Solution. Since c is a characteristic value of T, therefore there exists a non-zero vector α in V such that

$$T\alpha=c\alpha$$

$$\Rightarrow T(T\alpha)=T(c\alpha) \Rightarrow T^2\alpha=c\,T\alpha \qquad [\because T\alpha=c\alpha]$$

$$\Rightarrow T^2\alpha=c\,(c\alpha)$$

$$\Rightarrow T^2\alpha=c^2\alpha.$$

Scanned with CamScanner

$\therefore$ c^2 is a characteristic value of T^2.

Repeating this process k times, we get

$$T^k\alpha = c^k\alpha.$$

$\therefore$ c^k is a characteristic value of T^k where k is any positive integer.

Let $p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_mx^m$ where the a_i's $\in F$.

Then $\quad p(T) = a_0I + a_1T + a_2T^2 + \ldots + a_mT^m$.

We have $[p(T)]\,\alpha = (a_0I + a_1T + \ldots + a_mT^m)\,\alpha$

$= a_0I\alpha + a_1T\alpha + \ldots + a_mT^m\,\alpha$

$= a_0\alpha + a_1(c\alpha) + \ldots + a_m(c^m\alpha)$

$= (a_0 + a_1c + \ldots + a_mc^m)\,\alpha.$

$\therefore$ $p(c) = a_0 + a_1c + \ldots + a_mc^m$ is a characteristic value of $p(T)$.

Example 8. *Let A be a square matrix of order n over the field F. If $c \in F$ is a characteristic value of A, then for any polynomial $p(x)$ over F, $p(c)$ is a characteristic value of $p(A)$.*

Solution. Proceed as in Ex. 7. Replace the linear operator T by the matrix A.

Example 9. *If A and B are similar linear operators on a finite dimensional vector space V, then A and B have the same characteristic polynomial.*

Solution. Suppose A and B are similar linear transformations on a finite dimensional vector space V. Then there exists an invertible linear operator C on V such that

$$A = CBC^{-1}.$$

We have $A - xI = CBC^{-1} - xI = CBC^{-1} - C(xI)\,C^{-1}$

$= C(B - xI)\,C^{-1}$.

$\therefore$ $\det(A - xI) = \det\{C(B - xI)\,C^{-1}\}$

$= \det C.\ \det(B - xI).\ \det C^{-1} = \det C.\ \det C^{-1}.\ \det(B - xI)$

$= \det(CC^{-1})\ \det(B - xI) = \det I.\ \det(B - xI) = 1.\ \det(B - xI)$

$= \det(B - xI)$.

Now the characteristic polynomial of $A = \det(A - xI)$.

$\therefore$ the characteristic polynomial of A = the characteristic polynomial of B.

Example 10. *Suppose S and T are two linear operators on a finite dimensional vector space V. If S and T have the same characteristic polynomial, then $\det S = \det T$.*

Solution. Let $\dim V = n$. Let B be any ordered basis for V. If A is the matrix of S relative to B, then the characteristic polynomial $f(x)$ of S is given by

$$f(x) = \det(A - xI).$$

Let $\quad f(x) = a_0 + a_1x + \ldots + a_nx^n = \det(A - xI)$.

The constant term in this polynomial is

$$= a_0 = f(0) = \det A.$$

Similarly the constant term in the characteristic polynomial of $T = \det C$, where C is the matrix of T relative to B.

Since S and T have the same characteristic polynomial, therefore the two constant terms must be equal.

$\therefore$ $\det A = \det C$

$\Rightarrow$ $\det [S]_B = \det [T]_B$

$\Rightarrow$ $\det S = \det T$.

Example 11. *Find all (complex) proper values and proper vectors of the following matrices.*

(a) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ $\quad$ (b) $\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$ $\quad$ (c) $\begin{bmatrix} 1 & 1 \\ 0 & i \end{bmatrix}$.

Solution. (a) $\quad$ Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

We have $\quad A - xI = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} - x\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -x & 1 \\ 0 & -x \end{bmatrix}$.

$\therefore$ characteristic polynomial of $A = \det(A - xI)$

$$= \begin{vmatrix} -x & 1 \\ 0 & -x \end{vmatrix} = x^2.$$

$\therefore$ the characteristic equation of A is

$$\det(A - xI) = 0 \text{ i.e. } x^2 = 0.$$

The only root of this equation is $x = 0$.

$\therefore$ 0 is the only characteristic value of A.

Now let x_1, x_2 be the components of a characteristic vector α corresponding to this characteristic value. Let X be the coordinate matrix of α. $\quad$ Then $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

Now X will be given by a non-zero solution of the equation

$$(A - 0I)\,X = O$$

i.e. $\quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

i.e. $\quad \begin{bmatrix} x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Thus $x_2 = 0$, $x_1 = k$ where k is any non-zero complex number.

$\therefore$ $\quad X = \begin{bmatrix} k \\ 0 \end{bmatrix}$ where k is any non-zero complex number.

Scanned with CamScanner

(b) Let $A=\begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix}$.

We have $A-xI=\begin{bmatrix}1-x & 0\\ 0 & i-x\end{bmatrix}$.

$\therefore$ the characteristic equation of A is

$$\det(A-xI)=0$$

i.e. $$\begin{vmatrix}1-x & 0\\ 0 & i-x\end{vmatrix}=0$$

i.e. $$(1-x)(i-x)=0.$$

The roots of this equation are $x=1$, $x=i$.

$\therefore$ 1 and i are the two characteristic values of A.

Now let x_1, x_2 be the components of a characteristic vector corresponding to the characteristic value 1. Let X be the coordinate matrix of this vector. Then $X=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$.

Now X will be given by a non-zero solution of the equation

$$(A-1I)X=O$$

i.e. $$\begin{bmatrix}1-1 & 0\\ 0 & i-1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}0 & 0\\ 0 & i-1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}0\\ (i-1)x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.$$

Thus $x_2=0$, $x_1=k$ where k is any non-zero complex number.

$\therefore$ $X=\begin{bmatrix}k\\ 0\end{bmatrix}$ where k is any non-zero complex number.

Similarly to find characteristic vectors corresponding to the characteristic value i, we consider the equation

$$(A-iI)X=O$$

i.e. $$\begin{bmatrix}1-i & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}(1-i)x_1\\ 0\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.$$

Thus $x_1=0$, $x_2=c$ where c is any non-zero complex number.

$\therefore$ $X=\begin{bmatrix}0\\ c\end{bmatrix}$ where $c\neq 0$.

(c) Let $A=\begin{bmatrix}1 & 1\\ 0 & i\end{bmatrix}$.

We have $A-xI=\begin{bmatrix}1-x & 1\\ 0 & i-x\end{bmatrix}$.

$\therefore$ the characteristic equation of A is $\det(A-xI)=0$

i.e. $$\begin{vmatrix}1-x & 1\\ 0 & i-x\end{vmatrix}=0$$

i.e. $$(1-x)(i-x)=0.$$

The roots of this equation are $x=1, i$.

$\therefore$ 1 and i are two characteristic values of A.

Let $X=\begin{bmatrix}x_1\\ x_2\end{bmatrix}$ be the coordinate matrix of a characteristic vector corresponding to the characteristic value $x=1$. Then X will be given by a non-zero solution of the equation

$$(A-I)X=O$$

i.e. $$\begin{bmatrix}0 & 1\\ 0 & i-1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}x_2\\ (i-1)x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.$$

Thus $x_2=0$, $x_1=k$ where $k\neq 0$.

$\therefore$ $X=\begin{bmatrix}k\\ 0\end{bmatrix}$ where $k\neq 0$.

To find the characteristic vectors corresponding to the characteristic value i we consider the equation

$$(A-iI)X=O$$

i.e. $$\begin{bmatrix}1-i & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}(1-i)x_1+x_2\\ 0\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

i.e. $$(1-i)x_1+x_2=0.$$

Let $x_1=c$. Then $x_2=(i-1)c$.

$\therefore$ $X=\begin{bmatrix}c\\ (i-1)c\end{bmatrix}$ where $c\neq 0$.

Example 12. *Find all (complex) characteristic values and characteristic vectors of the following matrices* (Meerut 1968)

(a) $\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$, (b) $\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$.

Solution. (a) Let $A=\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$.

We have $$A-xI=\begin{bmatrix}1-x & 1 & 1\\ 1 & 1-x & 1\\ 1 & 1 & 1-x\end{bmatrix}.$$

Scanned with CamScanner

$\therefore$ the characteristic polynomial of A is

$$=\det(A-xI)=\begin{vmatrix}1-x & 1 & 1\\ 1 & 1-x & 1\\ 1 & 1 & 1-x\end{vmatrix}$$

$$=\begin{vmatrix}3-x & 1 & 1\\ 3-x & 1-x & 1\\ 3-x & 1 & 1-x\end{vmatrix} \quad C_1+C_2+C_3$$

$$=(3-x)\begin{vmatrix}1 & 1 & 1\\ 1 & 1-x & 1\\ 1 & 1 & 1-x\end{vmatrix}$$

$$=(3-x)\begin{vmatrix}1 & 1 & 1\\ 0 & -x & 0\\ 0 & 0 & -x\end{vmatrix} \quad R_2-R_1,\ R_3-R_1$$

$$=(3-x)\,x^2.$$

$\therefore$ the characteristic equation of A is

$$(3-x)\,x^2=0.$$

The only roots of this equation are $x=3, 0$.

$\therefore$ 0 and 3 are the only characteristic values of A.

Let $X=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$ be the coordinate matrix of a characteristic vector corresponding to the characteristic value $x=0$. Then X will be given by a non-zero solution of the equation

$$(A-0I)\,X=O$$

i.e. $$\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}x_1+x_2+x_3\\ x_1+x_2+x_3\\ x_1+x_2+x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

i.e. $$x_1+x_2+x_3=0.$$

This equation has two linearly independent solutions i.e.

$$X_1=\begin{bmatrix}1\\ 0\\ -1\end{bmatrix}, \text{ and } X_2=\begin{bmatrix}0\\ 1\\ -1\end{bmatrix}.$$

Every non-zero multiple of these column matrices X_1 and X_2 is a characteristic vector of A corresponding to the characteristic value 0.

The characteristic space of this characteristic value will be the subspace W spanned by these two vectors X_1 and X_2. Any non-zero vector in W will be a characteristic vector corresponding to this characteristic value.

To find the characteristic vectors corresponding to the characteristic value 3 we consider the equation

$$(A-3I)\,X=O$$

i.e. $$\begin{bmatrix}-2 & 1 & 1\\ 1 & -2 & 1\\ 1 & 1 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}-2x_1+x_2+x_3\\ x_1-2x_2+x_3\\ x_1+x_2-2x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

i.e. $$-2x_1+x_2+x_3=0,$$
$$x_1-2x_2+x_3=0,$$
$$x_1+x_2-2x_3=0.$$

Solving these equations, we get

$$x_1=x_2=x_3=k.$$

$\therefore$ $X=\begin{bmatrix}k\\ k\\ k\end{bmatrix}$ where $k\neq 0$.

(b) Let $A=\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$.

The characteristic equation of A is

$$(1-x)^3=0.$$

$\therefore$ $x=1$ is the only characteristic value of A.

Let $X=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$ be the coordinate matrix of a characteristic vector corresponding to the characteristic value 1. Then X will be given by a non-zero solution of the equation

$$(A-I)\,X=O$$

i.e. $$\begin{bmatrix}0 & 1 & 1\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

i.e. $$\begin{bmatrix}x_2+x_3\\ x_2\\ 0\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

i.e. $$x_2+x_3=0,\ x_2=0.$$

$\therefore$ $x_1=k,\ x_2=0,\ x_3=0.$

Thus $X=\begin{bmatrix}k\\ 0\\ 0\end{bmatrix}$ where $k\neq 0$.

Scanned with CamScanner

Example 13. *Let T be a linear operator on the n-dimensional vector space V, and suppose that T has n distinct characteristic values. Prove that T is diagonalizable.* **(Meerut 1989)**

Solution. We know that distinct characteristic vectors of a linear operator T corresponding to distinct characteristic values of T are linearly independent. [See theorem 4 of § 24].

Since T has n distinct characteristic values, therefore T has n linearly independent characteristic vectors. These n linearly independent vectors will form a basis for V because dim $V=n$. Thus we get a basis for V each vector of which is a characteristic vector of T. Hence T is diagonalizable.

Example 14. *If an $n \times n$ matrix A has n distinct eigenvalues, A is diagonalizable.*

Solution. Let $k_1, k_2, \ldots, k_n$ be the n distinct eigenvalues of A and let X_i be an eigenvector of A corresponding to the eigenvalue k_i, $i=1, 2, \ldots, n$.

Then $\quad AX_i = k_i X_i$.

We shall prove that the eigenvectors $X_1, \ldots, X_n$ are linearly independent.

If $X_1, X_2, \ldots, X_n$ are linearly dependent we can choose r so that $1 \leqslant r < n$ and $X_1, X_2, \ldots, X_r$ are linearly independent but $X_1, X_2, \ldots, X_r, X_{r+1}$ are linearly dependent. Hence we can choose scalars $c_1, c_2, \ldots, c_{r+1}$, not all zero, such that

$$c_1X_1+c_2X_2+\ldots+c_{r+1}X_{r+1}=O. \quad \ldots(1)$$

Multiplying (1) on the left by A, we get

$$c_1AX_1+c_2AX_2+\ldots+c_{r+1}AX_{r+1}=O$$

or
$$c_1k_1X_1+c_2k_2X_2+\ldots+c_{r+1}k_{r+1}X_{r+1}=O. \quad \ldots(2)$$

$$[\because \; AX_i=k_iX_i]$$

Now multiplying (1) by the scalar k_{r+1} and subtracting from (2), we get

$$c_1(k_1-k_{r+1})X_1+c_2(k_2-k_{r+1})X_2+\ldots +c_r(k_r-k_{r+1})X_r=O. \quad \ldots(3)$$

But since $X_1, X_2, \ldots, X_r$ are linearly independent according to our assumption and $k_1, k_2, \ldots, k_{r+1}$ are all distinct, therefore from (3), we get

$$c_1=0, c_2=0, \ldots, c_r=0.$$

Putting $c_1=0, \ldots, c_r=0$ in (1), we get

$$c_{r+1}X_{r+1}=O$$

$$\Rightarrow c_{r+1}=0, \text{ since } X_{r+1}\neq O.$$

Thus the relation (1) implies that

$$c_1=0, c_2=0, \ldots, c_r=0, c_{r+1}=0.$$

But this contradicts our assumption that the scalars

$$c_1, \ldots, c_{r+1}$$

are not all zero.

Hence our initial assumption is wrong and the vectors

$$X_1, \ldots, X_n$$

are linearly independent. Since the matrix A has n linearly independent eigenvectors, therefore it is diagonalizable.

Example 15. *Let T be the linear operator on $\mathbf{R}^3$ which is represented in the standard basis by the matrix*

$$\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}.$$

Prove that T is diagonalizable. **(Meerut 1984, 87, 90)**

Solution.

Let $\quad A=\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}.$

The characteristic equation of A is

$$\begin{vmatrix} -9-x & 4 & 4 \\ -8 & 3-x & 4 \\ -16 & 8 & 7-x \end{vmatrix}=0$$

or
$$\begin{vmatrix} -1-x & 4 & 4 \\ -1-x & 3-x & 4 \\ -1-x & 8 & 7-x \end{vmatrix}=0, \text{ applying } C_1+C_2+C_3$$

or
$$-(1+x)\begin{vmatrix} 1 & 4 & 4 \\ 1 & 3-x & 4 \\ 1 & 8 & 7-x \end{vmatrix}=0$$

or
$$(1+x)\begin{vmatrix} 1 & 4 & 4 \\ 0 & -1-x & 0 \\ 0 & 1 & 3-x \end{vmatrix}=0 \text{ applying } R_2-R_1, R_3-R_1$$

or
$$(1+x)(1+x)(3-x)=0.$$

The roots of this equation are $-1, -1, 3$.

$\therefore$ The eigen values of the matrix A are $-1, -1, 3$.

The characteristic vector X of A corresponding to the eigenvalue -1 are given by the equation

$$(A-(-1)\,I)\,X=O$$

or $$(A+I)\,X=O$$

or $$\begin{bmatrix}-8 & 4 & 4\\ -8 & 4 & 4\\ -16 & 8 & 8\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

These equations are equivalent to the equations

$$\begin{bmatrix}-8 & 4 & 4\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix},\ \text{applying } R_2-R_1,\ R_3-2R_1.$$

The matrix of coefficients of these equations has rank 1. Therefore these equations have two linearly independent solutions. We see that these equations reduce to the single equation

$$-2x_1+x_2+x_3=0.$$

Obviously

$$X_1=\begin{bmatrix}1\\ 1\\ 1\end{bmatrix},\quad X_2=\begin{bmatrix}0\\ 1\\ -1\end{bmatrix}$$

are two linearly independent solutions of this equation. Therefore X_1 and X_2 are two linearly independent eigenvectors of A corresponding to the eigenvalue -1.

Now the eigenvectors of A corresponding to the eigenvalue 3 are given by

$$(A-3I)\,X=O$$

i.e. $$\begin{bmatrix}-12 & 4 & 4\\ -8 & 0 & 4\\ -16 & 8 & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

These equations are equivalent to the equations

$$\begin{bmatrix}-12 & 4 & 4\\ 4 & -4 & 0\\ -4 & 4 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix},$$

applying R_2-R_1, R_3-R_1.

The matrix of coefficients of these equations has rank 2. Therefore these equations will have a non-zero solution. Also these equations will have $3-2=1$ linearly independent solution. These equations can be written as

$$-12x_1+4x_2+4x_3=0$$
$$4x_1-4x_2=0$$
$$-4x_1+4x_2=0.$$

From these, we get

$$x_1=x_2=1,\ \text{say}.$$

Then $$x_3=2.$$

$$\therefore\quad X_3=\begin{bmatrix}1\\ 1\\ 2\end{bmatrix}$$

is an eigenvector of A corresponding to the eigenvalue 3.

Now let $$P=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 1\\ 1 & -1 & 2\end{bmatrix}.$$

We have det $P=1\neq 0$. Therefore the matrix P is invertible. Therefore the columns of P are linearly independent vectors belonging to $\mathbf{R}^3$. Since the matrix A has three linearly independent eigenvectors in $\mathbf{R}^3$, therefore it is diagonalizable. Consequently the linear operator T is diagonalizable. Also the diagonal form D of A is given by

$$P^{-1}AP=\begin{bmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 3\end{bmatrix}=D.$$

Example 16. *Is the matrix*

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

similar over the field $\mathbf{R}$ *to a diagonal matrix? Is A similar over the field* $\mathbf{C}$ *to a diagonal matrix?*

Solution. The characteristic equation of A is

$$\begin{vmatrix}1-x & 1\\ -1 & 1-x\end{vmatrix}=0$$

or $$(1-x)^2+1=0$$

or $$x^2-2x+2=0.$$

The roots of this equation are $1+i$, $1-i$. Since the characteristic equation of A has no roots in $\mathbf{R}$, therefore the matrix A has no eigenvalue if we regard it as a matrix over $\mathbf{R}$. Consequently A has no eigenvector in $\mathbf{R}^2$. Therefore the matrix A is not diagonalizable over the field $\mathbf{R}$.

If we regard A as a matrix over $\mathbf{C}$, then it has two eigenvalues

i.e. $$1+i,\ 1-i$$

Since A has two distinct eigen values, therefore it will ha two linearly independent eigenvectors. Consequently A is diagon lizable.

The eigenvectors of A corresponding to the eigenvalues $1+i$, $1-i$ are given by the system of equations

$$\begin{bmatrix}-i & 1\\ -1 & i\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix},\ \begin{bmatrix}i & 1\\ -1 & i\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

respectively.

Scanned with CamScanner

From these, we get

$$X_1=\begin{bmatrix}1\\ i\end{bmatrix},\ X_2=\begin{bmatrix}1\\ -i\end{bmatrix}$$

as the eigenvectors of A corresponding to the eigenvalues $1+i$, $1-i$ respectively.

If $P=\begin{bmatrix}1 & i\\ i & -i\end{bmatrix}$, then

$$P^{-1}AP=\begin{bmatrix}1+i & 0\\ 0 & 1-i\end{bmatrix}$$

gives the diagonal form of A.

Example 17. *Prove that the matrix*

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

is not diagonalizable over the field **C**.

Solution. The characteristic equation of A is

$$\begin{vmatrix}1-x & 2\\ 0 & 1-x\end{vmatrix}=0$$

or $$(1-x)^2=0.$$

The roots of this equation are 1, 1. Therefore the only distinct eigenvalue of A is 1. The eigenvectors of A corresponding to this eigenvalue are given by

$$\begin{bmatrix}0 & 2\\ 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

or $$0x_1+2x_2=0.$$

This equation has only one linearly independent solution. We see that

$$X=\begin{bmatrix}1\\ 0\end{bmatrix}$$

is the only linearly independent eigenvector of A. Since A has not two linearly independent eigenvectors, therefore it is not diagonalizable.

Example 18. *Show that the matrix*

$$A=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$$ *is not diagonalizable.* (Meerut 1992)

§ 25. Minimal polynomial and Minimal equation of a linear operator or of a matrix

Annihilating polynomials. Suppose T is a linear operator on a finite dimensional vector space over the field F and $f(x)$ is a polynomial over F. If $f(T)=\hat{0}$, then we say that the polynomial $f(x)$ *annihilates* the linear operator T. Similarly suppose A is a square matrix of order n over the field F and $f(x)$ is a polynomial over F. If $f(A)=O$, then we say that the polynomial $f(x)$ *annihilates* **the matrix A. We know that every linear operator T on an** n-dimensional vector space $V(F)$ satisfies its characteristic equation. Also the characteristic polynomial of T is a non-zero polynomial i.e. a polynomial in which the coefficients of various terms are not all zero. Note that if A is the matrix of T in some ordered basis, then the characteristic polynomial of T is $|A-xI|$ in which the coefficient of x^n is $(-1)^n$ which is not zero. Thus we see that at least the characteristic polynomial of T is a non-zero polynomial which annihilates T. Therefore the set of those non-zero polynomials which annihilate T is not empty.

Monic polynomial. Definition. A polynomial in x over a field F is called a monic polynomial if the coefficient of the highest power of x in it is unity. Thus $x^3-2x^2+\frac{5}{7}x+5$ is a monic polynomial of degree 3 over the field of rational numbers.

Among these non-zero polynomials which annihilate a linear operator T, the polynomial which is monic and which is of the lowest degree is of special interest. It is called the *minimal polynomial of* the linear operator T.

Minimal polynomial of a linear operator. Definition.
[Meerut 1976, 80, Nagarjuna 74]

Suppose T is a linear operator on an n-dimensional vector space $V(F)$. The monic polynomial of lowest degree over the field F that annihilates T is called the minimal polynomial of T. Also if $f(x)$ is the minimal polynomial of T, the equation $f(x)=0$ is called the minimal equation of the linear operator T.

Similarly we can define the minimal polynomial of a matrix. Suppose A is a square matrix of order n over the field F. The monic polynomial of lowest degree over the field F that annihilates A is called the *minimal polynomial* of A.

Now suppose T is a linear operator on an n-dimensional vector space $V(F)$ and A is the matrix of T in some ordered basis B. If $f(x)$ is any polynomial over F, then $[f(T)]_B=f(A)$. Therefore $f(T)=\hat{0}$ if and only if $f(A)=O$. Thus $f(x)$ annihilates T iff it annihilates A. Therefore if $f(x)$ is the polynomial of lowest degree that annihilates T, then it is also the polynomial of lowest degree that annihilates A and conversely. Hence T and A have the same minimal polynomial. Further the characteristic polynomial of the matrix A is of degree n. Since the characteristic polynomial of A annihilates A, therefore the minimal polynomial of A cannot be of degree greater than n. Its degree must be less than or equal to n.

Scanned with CamScanner

Theorem 1. *The minimal polynomial of a matrix or of a linear operator is unique.*

Proof. Suppose the minimal polynomial of a matrix A is of degree r. Then no non-zero polynomial of degree less than r can annihilate A. Let $f(x)=x^r+a_1x^{r-1}+a_2x^{r-2}+\ldots+a_{r-1}x+a_r$ and $g(x)=x^r+b_1x^{r-1}+b_2x^{r-2}+\ldots+b_{r-1}x+b_r$ be two minimal polynomials of A. Then both $f(x)$ and $g(x)$ annihilate A. Therefore we have $f(A)=O$ and $g(A)=O$. These give

$$A^r+a_1A^{r-1}+\ldots+a_{r-1}A+a_r I=O,$$

and $$A^r+b_1A^{r-1}+\ldots+b_{r-1}A+b_r I=O. \quad \ldots(1)$$

$\ldots(2)$

Subtracting (1) from (2), we get

$$(b_1-a_1)A^{r-1}+\ldots+(b_r-a_r)I=O. \quad \ldots(3)$$

From (3), we see that the polynomial $(b_1-a_1)x^{r-1}+\ldots+(b_r-a_r)$ also annihilates A. Since its degree is less than r, therefore it must be a zero polynomial. This gives $b_1-a_1=0$, $b_2-a_2=0,\ldots,b_r-a_r=0$. Thus $a_1=b_1,\ldots,a_r=b_r$. Therefore $f(x)=g(x)$ and thus the minimal polynomial of A is unique.

Theorem 2 *The minimal polynomial of a matrix (linear operator) is a divisor of every polynomial that annihilates the matrix (linear operator).*

Proof. Suppose $m(x)$ is the minimal polynomial of a matrix A. Let $h(x)$ be any polynomial that annihilates A. Since $m(x)$ and $h(x)$ are two polynomials, therefore by the division algorithm there exist two polynomials $q(x)$ and $r(x)$ such that

$$h(x)=m(x)\,q(x)+r(x), \quad \ldots(1)$$

where either $r(x)$ is a zero polynomial or its degree is less than the degree of $m(x)$. Putting

$x=A$ on both sides of (1), we get

$$h(A)=m(A)\,q(A)+r(A)$$

$\Rightarrow O=O\,q(A)+r(A)$ [$\because$ both $m(x)$ and $h(x)$ annihilate A]

$\Rightarrow r(A)=O.$

Thus $r(x)$ is a polynomial which also annihilates A. If $r(x)\neq 0$, then it is a non-zero polynomial of degree smaller than the degree of the minimal polynomial $m(x)$ and thus we arrive at a contradiction that $m(x)$ is the minimal polynomial of A. Therefore $r(x)$ must be a zero polynomial. Then (1) gives

$$h(x)=m(x)\,q(x) \Rightarrow m(x) \text{ is a divisor of } h(x).$$

Corollary. *The minimal polynomial of a matrix is a divisor of the characteristic polynomial of that matrix.*

Proof. Suppose $f(x)$ is the characteristic polynomial of a matrix A. Then $f(A)=O$ by Cayley-Hamilton theorem. Thus $f(x)$ annihilates A. If $m(x)$ is the minimal polynomial of A, then by the above theorem we see that $m(x)$ must be a divisor of $f(x)$.

Theorem 3. *Let T be a linear operator on an n-dimensional vector space V [or, let A be an $n\times n$ matrix]. The characteristic and minimal polynomials for T [for A] have the same roots, except for multiplicities.* **(Meerut 1984, G.N.D.U. Amritsar 90)**

Proof. Suppose $f(x)$ is the characteristic polynomial of a linear operator T and $m(x)$ is its minimal polynomial. First we shall prove that every root of the equation $m(x)=0$ is also a root of the equation $f(x)=0$. We know that the minimal polynomial is a divisor of the characteristic polynomial. Therefore $m(x)$ is a divisor of $f(x)$. Then there exists a polynomial $q(x)$ such that

$$f(x)=m(x)\,q(x). \quad \ldots(1)$$

Suppose c is a root of the equation $m(x)=0$. Then $m(c)=0$. Putting $x=c$ on both sides of (1) we get $f(c)=m(c)\,q(c)=0\,q(c)=0$. Therefore c is also a root of $f(x)=0$. Thus c is also a characteristic root of the linear operator T.

Conversely suppose that c is a characteristic value of T. Then there exists a non-zero vector α such that $T\alpha=c\alpha$. Since $m(x)$ is a polynomial, therefore we have

$$[m(T)]\,(\alpha)=m(c)\,\alpha. \quad \text{[See Ex. 7 page 263]}$$

But $m(x)$ is the minimal polynomial for T. So $m(x)$ annihilates T i.e., $m(T)=\hat{0}$.

$\therefore\ \hat{0}\,(\alpha)=m(c)\,\alpha$

$\Rightarrow 0=m(c)\,\alpha$ [$\because \hat{0}(\alpha)=0$]

$\Rightarrow m(c)=0.$ [$\because \alpha\neq 0$]

Thus c is a root of the minimal equation of T.

Hence every root of the minimal equation of T is also a root of its characteristic equation and every root of the characteristic equation of T is also a root of its minimal equation.

Theorem 4. *Let T be a diagonalizable linear operator and let $c_1,\ldots,c_k$ be the distinct characteristic values of T. Then the minimal polynomial for T is the polynomial*

$$p(x)=(x-c_1)\ldots(x-c_k). \quad \textbf{(Meerut 1980, 84P)}$$

Proof. We know that each characteristic value of T is a root of the minimal polynomial for T. Therefore each of the scalars

Scanned with CamScanner

$c_1, \ldots, c_k$ is a root of the minimal polynomial for T and so each of the polynomials $x-c_1, \ldots, x-c_k$ is a factor of minimal polynomial for T. Therefore the polynomial $p(x)=(x-c_1)\ldots(x-c_k)$ will be the minimal polynomial for T provided it annihilates T *i.e.* provided $p(T)=\hat{0}$.

Let α be a characteristic vector of T. Then one of the operators $T-c_1I, \ldots, T-c_kI$ sends α into $\mathbf{0}$. Therefore

$$(T-c_1I)\ldots(T-c_kI)\,\alpha=\mathbf{0} \text{ i.e. } p(T)\,\alpha=\mathbf{0}$$

for every characteristic vector α.

Now T is a diagonalizable operator. Let V be the underlying vector space. Then there exists a basis B for V which consists of characteristic vectors. If β is any vector in V, then β can be expressed as a linear combination of the vectors in the basis B. But we have just shown that $p(T)\,\alpha=\mathbf{0}$ for every characteristic vector. Therefore we have

$$p(T)\,\beta=\mathbf{0}, \ \forall\ \beta \in V$$

$$\Rightarrow \quad p(T)=\hat{0}.$$

$\therefore$ $p(x)$ annihilates T and so $p(x)$ is the minimal polynomial for T.

Thus we have proved that *if T is a diagonalizable linear operator, the minimal polynomial for T is a product of distinct linear factors.*

Corollary. *If the roots of the characteristic equation of a linear operator T are all distinct say $c_1, c_2, \ldots, c_n$, then the minimal polynomial for T is the polynomial*

$$(x-c_1)\ldots(x-c_n).$$

Proof. Since the roots of the characteristic equation of T are all distinct, therefore T is diagonalizable. Hence by the above theorem, the minimal polynomial for T is the polynomial

$$(x-c_1)(x-c_2)\ldots(x-c_n).$$

Solved Examples

Example 1. *Let V be a finite-dimensional vector space. What is the minimal polynomial for the identity operator on V? What is the minimal polynomial for the zero operator?* **(Meerut 1981)**

Solution. We have $I-1\,I=I-I=\hat{0}$. Therefore the monic polynomial $x-1$ annihilates the identity operator I and it is the polynomial of lowest degree that annihilates I. Hence $x-1$ is the minimal polynomial for I.

Again we see that the monic polynomial x annihilates the zero operator $\hat{0}$ and it is the polynomial of lowest degree that annihilates $\hat{0}$. Hence x is the minimal polynomial for $\hat{0}$.

Example 2. *Let V be an n-dimensional vector space and let T be a linear operator on V. Suppose that there exists some positive integer k so that $T^k=\hat{0}$. Prove that $T^n=\hat{0}$.*

Solution. Since $T^k=\hat{0}$, therefore the polynomial x^k annihilates T. So the minimal polynomial for T is a divisor of x^k. Let x^r be the minimal polynomial for T where $r \leqslant n$. Then $T^r=\hat{0}$.

Now $T^n=T^{n-r}\,T^r=T^{n-r}\,\hat{0}=\hat{0}$.

Example 3. *Find the minimal polynomial for the real matrix*

$$A=\begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}.$$

Solution. We have

$$|A-xI|=\begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ -4 & -4 & 4-x \end{vmatrix}$$

$$=\begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ 0 & 3-x & 3-x \end{vmatrix}, \text{ by } R_3+R_2$$

$$=(3-x)\begin{vmatrix} 7-x & 4 & -1 \\ 4 & 7-x & -1 \\ 0 & 1 & 1 \end{vmatrix}$$

$$=(3-x)\begin{vmatrix} 7-x & 4 & -5 \\ 4 & 7-x & x-8 \\ 0 & 1 & 0 \end{vmatrix}, \text{ by } C_3-C_2$$

$$=-(3-x)\begin{vmatrix} 7-x & -5 \\ 4 & x-8 \end{vmatrix}, \text{ expanding along third row}$$

$$=-(3-x)\begin{vmatrix} 3-x & 3-x \\ 4 & x-8 \end{vmatrix}, \text{ by } R_1-R_2$$

$$=-(3-x)^2\begin{vmatrix} 1 & 1 \\ 4 & x-8 \end{vmatrix}=-(3-x)^2(x-12).$$

Therefore the roots of the equation $|A-xI|=0$ are $x=3, 3, 12$. These are the characteristic roots of A.

Scanned with CamScanner

Let us now find the minimal polynomial of A. We know that each characteristic root of A is also a root of its minimal polynomial. So if $m(x)$ is the minimal polynomial for A, then both $x-3$ and $x-12$ are factors of $m(x)$. Let us try whether the polynomial $h(x)=(x-3)(x-12)=x^2-15x+36$ annihilates A or not.

We have $A^2=\begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix}$.

$$\therefore \quad A^2-15A+36I=\begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix}$$

$$-15\begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}+\begin{bmatrix} 36 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 36 \end{bmatrix}$$

$$=\begin{bmatrix} 105 & 60 & -15 \\ 60 & 105 & -15 \\ -60 & -60 & 60 \end{bmatrix}-\begin{bmatrix} 105 & 60 & -15 \\ 60 & 105 & -15 \\ -60 & -60 & 60 \end{bmatrix}=O.$$

$\therefore$ $h(x)$ annihilates A. Thus $h(x)$ is the monic polynomial of lowest degree which annihilates A. Hence $h(x)$ is the minimal polynomial for A.

Note. In order to find the minimal polynomial of a matrix A, we should not forget that each characteristic root of A must also be a root of the minimal polynomial. We should try to find the monic polynomial of lowest degree which annihilates A and which has also the characteristic roots of A as its roots.

Example 4. *Show that similar matrices have the same minimal polynomial.* (Meerut 1976)

Solution: Suppose A and B are two similar matrices. Then there exists a non-singular matrix P such that

$$B=P^{-1}AP.$$

Now we are to show that the matrices A and $P^{-1}AP$ have the same monic polynomial. First we shall show that a monic polynomial $f(x)$ annihilates A if and only if it annihilates $P^{-1}AP$. We have

$$(P^{-1}AP)^2=P^{-1}AP\,P^{-1}AP=P^{-1}A^2P.$$

Proceeding in this way we can show that $(P^{-1}AP)^k=P^{-1}A^kP$, where k is any positive integer.

Let $f(x)=x^r+a_1x^{r-1}+\dots+a_{r-1}x+a_r$.

Then $f(A)=A^r+a_1A^{r-1}+\dots+a_{r-1}A+a_rI$.

Also $f(P^{-1}AP)=(P^{-1}AP)^r+\dots+a_{r-1}(P^{-1}AP)+a_rI$

$=P^{-1}A^rP+\dots+a_{r-1}(P^{-1}AP)+a_rP^{-1}P$

$=P^{-1}(A^r+\dots+a_{r-1}A+a_rI)P$

$=P^{-1}f(A)P.$

Since P is non-singular, therefore

$P^{-1}f(A)P=O$ if and only if $f(A)=O$.

Thus $f(x)$ annihilates A if and only if it annihilates $P^{-1}AP$. Therefore if $f(x)$ is the polynomial of lowest degree that annihilates $P^{-1}AP$ and conversely. Hence A and $P^{-1}AP$ have the same minimal polynomial.

Exercises

1. Find the characteristic roots of the matrix
$$\begin{bmatrix} 5 & 6 & 8 \\ 0 & 7 & 2 \\ 0 & 0 & 4 \end{bmatrix}.$$ (Meerut 1976)

2. Write the characteristic polynomial and the minimal polynomial of the matrix
$$\begin{bmatrix} 4 & 3 & 0 \\ 2 & 1 & 0 \\ 5 & 7 & 9 \end{bmatrix}.$$

3. State whether the following statement is true or false:
The matrices
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$
have the same characteristic roots. (Meerut 1977)

4. Show that the characteristic equation of the complex matrix
$$A=\begin{bmatrix} 0 & 0 & c \\ 1 & 0 & b \\ 0 & 1 & a \end{bmatrix}$$
is $x^3-ax^2-bx-c=0$. (Meerut 1980)

5. Show that the minimal polynomial of the real matrix
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$
is x^2+1.

6. Show that the minimal polynomial of the real matrix
$$\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$
is $(x-1)(x-2)$.

Scanned with CamScanner

7. Find all (complex) eigenvalues and eigenvectors of the following matrices.

(a) $\begin{bmatrix} 2 & 4 \\ 3 & 13 \end{bmatrix}$. (b) $\begin{bmatrix} 2 & -3 \\ -3 & 1 \end{bmatrix}$. (c) $\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}$.

(d) $\begin{bmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{bmatrix}$.

8. What are the eigenvalues and eigenvectors of the identity matrix ?

9. For each of the following matrices over the field **C**, find the diagonal form and a diagonalizing matrix P.

(a) $\begin{bmatrix} 20 & 18 \\ -27 & -25 \end{bmatrix}$. (b) $\begin{bmatrix} 3 & 4 \\ -4 & 3 \end{bmatrix}$.

(c) $\begin{bmatrix} 4 & 2 & -2 \\ -5 & 3 & 2 \\ -2 & 4 & 1 \end{bmatrix}$. (d) $\begin{bmatrix} -17 & 18 & -6 \\ -18 & 19 & -6 \\ -9 & 9 & 2 \end{bmatrix}$.

10. Show that distinct eigenvectors of a matrix A corresponding to distinct eigenvalues of A are linearly independent.

11. Let T be the linear operator on $\mathbf{R}^3$ which is represented in the standard ordered basis by the matrix

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

Find the characteristic values of A and prove that T is diagonalizable. (Meerut 1985)

12. Is the matrix

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

similar over the field **R** to a diagonal matrix ? Is A similar over the field **C** to a diagonal matrix ? (Meerut 1983)

13. Is the matrix

$$A=\begin{bmatrix} 6 & -3 & -2 \\ 4 & -1 & -2 \\ 10 & -5 & -3 \end{bmatrix}$$

similar over the field **R** to a diagonal matrix ? Is A similar over the field **C** to a diagonal matrix ? (G.N.D.U. Amritsar 1990)

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**. (Meerut 1988)

Answers

1. 5, 7, 4.
2. Characteristic polynomial is $(9-x)(x^2-5x-2)$, minimal polynomial is $(x-9)(x^2-5x-2)$.
3. True.
7. (a) 14, 1 ; [1, 3], [4, -1].
   (b) $\frac{3}{2}\pm\frac{1}{2}\sqrt{37}$; [6, $1-\sqrt{37}$], [6, $1+\sqrt{37}$].
   (c) $2\pm\sqrt{3}i$; [2, $1-\sqrt{3}i$], [2, $1+\sqrt{3}i$].
   (d) 8, -1, -1; linearly independent eigenvectors are [2, 1, 2], [0, 2, -1], [1, 0, -1].
8. All eigenvalues are 1. Every non-zero vector is an eigenvector.
9. (a) $D=\begin{bmatrix} 2 & 0 \\ 0 & -7 \end{bmatrix}$, $P=\begin{bmatrix} 1 & -3 \\ -1 & 2 \end{bmatrix}$.
   (b) $D=\begin{bmatrix} 3+4i & 0 \\ 0 & 3-4i \end{bmatrix}$, $P=\begin{bmatrix} 1 & -1 \\ i & i \end{bmatrix}$.
   (c) $D=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$, $P=\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$.
   (d) $D=\begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $P=\begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$.
11. 1, 2, 2.
12. Roots of the characteristic equation of A are 1, 2, 2. A is not similar over the field **R** to a diagonal matrix. Here A has only two linearly independent eigenvectors belonging to $\mathbf{R}^3$. A is also not similar over the field **C** to a diagonal matrix.
13. Roots of the characteristic equation of A are 2, i, $-i$. A is not diagonalizable over **R**. But A is diagonalizable over **C** because in this case A has 3 distinct eigenvalues.

Scanned with CamScanner

3
Inner Product Spaces

Throughout this chapter we shall deal only with *real* or *complex* vector spaces. Thus if V is the vector space over the field F, then F will not be an arbitrary field. In this chapter F will be either the field **R** of real numbers or the field **C** of complex numbers.

Before defining inner product and inner product spaces, we shall just give some important properties of complex numbers.

Let $z \in$ **C** *i.e.*, let z be a complex number. Then $z=x+iy$ where $x, y \in$ **R** and $i=\sqrt{-1}$. Here x is called the real part of z and y is called the imaginary part of z. We write $x=\text{Re } z$, and $y=\text{Im } z$. The *modulus* of the complex number $z=x+iy$ is the non-negative real number $\sqrt{(x^2+y^2)}$ and is denoted by $|z|$. Also if $z=x+iy$ is a complex number, then the complex number $\bar{z}=x-iy$ is called the *conjugate complex of* z. If $z=\bar{z}$, then $x+iy=x-iy$ and therefore $y=0$. Thus $z=\bar{z}$ implies that z is real. Obviously we have

(i) $z+\bar{z}=2x=2 \text{ Re } z$, (ii) $z-\bar{z}=2iy=2i \text{ Im } z$,

(iii) $z\bar{z}=x^2+y^2=|z|^2$, (iv) $|z|=0 \Leftrightarrow x=0, y=0$

i.e., $|z|=0 \Leftrightarrow z=0$, (v) $\overline{(\bar{z})}=z$, (vi) $|\bar{z}|=|z|$, and

(vii) $|z|=\sqrt{(x^2+y^2)} \geqslant x$ *i.e.*, $|z| \geqslant \text{Re } z$.

If z_1 and z_2 are two complex numbers, then

(i) $|z_1+z_2| \leqslant |z_1|+|z_2|$ (ii) $\overline{z_1+z_2}=\bar{z}_1+\bar{z}_2$

(iii) $\overline{z_1 z_2}=\bar{z}_1 \bar{z}_2$, and (iv) $\overline{z_1-z_2}=\bar{z}_1-\bar{z}_2$.

§ 1. **Inner Product Spaces. Definition.** **(Meerut 1982, 83; Allahabad 75; Madras 81; Nagarjuna 78; Andhra 81, 85)**

Let $V(F)$ be a vector space where F is either the field of real numbers or the field of complex numbers. An **inner product** *on V is a function from $V \times V$ into F which assigns to each ordered pair of vectors α, β in V a scalar (α, β) in such a way that*

(*i*) $(\alpha, \beta)=\overline{(\beta, \alpha)}$ [*Here $\overline{(\beta, \alpha)}$ denotes the conjugate complex of the number (β, α)*].

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

(3) $(\alpha, \alpha) \geqslant 0$ *and* $(\alpha, \alpha)=0 \Rightarrow \alpha=0$

for any $\alpha, \beta, \gamma \in V$ *and* $a, b \in F$.

Also the vector space V is then said to be an **inner product space** *with respect to the specified inner product defined on it.*

It should be noted that in the above definition (α, β) does not denote the ordered pair of the vectors α and β. But it denotes the inner product of the vectors α and β. It is an element of V which has been assigned by the function (named as inner product) to the vectors α and β. Sometimes the inner product of the ordered pair of vectors α, β is also written as $(\alpha \mid \beta)$. If $F=$**R**, then (α, β) is a real number and if $F=$**C**, then (α, β) is a complex number.

If F is the field of real numbers, then the complex conjugate appearing in (1) is superfluous and (1) should be read as $(\alpha, \beta)=(\beta, \alpha)$. If F is the field of complex numbers, then from (1), we have $(\alpha, \alpha)=\overline{(\alpha, \alpha)}$ and therefore (α, α) is real. Thus (α, α) is always real whether $F=$**R** or $F=$**C**. Therefore the inequality given in (3) makes sense.

If $V(F)$ is an inner product space, then it is called a **Euclidean space** *if F is the field of real numbers. Also it is called a* **Unitary space** *if F is the field of complex numbers.*

Note 1. The property (3) in the definition of inner product is called *non-negativity*. The property (2) is called the *linearity property*. If $F=$**R**, then the property (1) is called *symmetry* and if $F=$**C**, then it is called *conjugate symmetry*.

Note 2. If in an inner product space $V(F)$, the vector α is **0**, then $(\alpha, \alpha)=0$.

We have $(\mathbf{0}, \mathbf{0})=(0\,\mathbf{0}, \mathbf{0})$ [$\because$ $0\,\mathbf{0}=\mathbf{0}$ in V]

$=0\,(\mathbf{0}, \mathbf{0})$ [by linearity property of inner product]

$=0$ [$\because$ $(\mathbf{0}, \mathbf{0}) \in F$ and therefore $0\,(\mathbf{0}, \mathbf{0})=0$]

EXAMPLES OF INNER PRODUCT SPACES

Example 1. On V_n (**C**) there is an inner product which we call the **standard inner product**.

If $\alpha=(a_1, a_2, \ldots, a_n), \beta=(b_1, b_2, \ldots, b_n) \in V_n$ (**C**), then we define

$$(\alpha, \beta)=a_1\bar{b}_1+a_2\bar{b}_2+\ldots+a_n\bar{b}_n=\sum_{i=1}^{n} a_i\bar{b}_i \qquad \ldots(1)$$

Scanned with CamScanner

Let us see that all the postulates of an inner product hold in (1).

(i) Conjugate symmetry. From the definition of product given in (1), we have $(\beta, \alpha)=b_1\bar{a}_1+\ldots+b_n\bar{a}_n$.

$$\therefore \quad \overline{(\beta, \alpha)}=\overline{(b_1\bar{a}_1+\ldots+b_n\bar{a}_n)}=\overline{(b_1\bar{a}_1)}+\ldots+\overline{(b_n\bar{a}_n)}$$

$$=\bar{b}_1\overline{(\bar{a}_1)}+\ldots+\bar{b}_n\overline{(\bar{a}_n)}=\bar{b}_1a_1+\ldots+\bar{b}_na_n$$

$$=a_1\bar{b}_1+\ldots+a_n\bar{b}_n \qquad [\because \text{ multiplication in C is commutative}]$$

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

Thus $\qquad (\alpha, \beta)=\overline{(\beta, \alpha)}$.

(ii) Linearity. Let $\gamma=(c_1,\ldots, c_n) \in V_n(\mathbf{C})$ and let $a, b \in \mathbf{C}$. We have $a\alpha+b\beta=a\,(a_1,\ldots, a_n)+b\,(b_1,\ldots, b_n)$

$$=(aa_1+bb_1,\ldots, aa_n+bb_n).$$

$$\therefore \quad (a\alpha+b\beta, \gamma)=(aa_1+bb_1)\,\bar{c}_1+\ldots+(aa_n+bb_n)\bar{c}_n \qquad [\text{by } (1)]$$

$$=(aa_1\bar{c}_1+\ldots+aa_n\bar{c}_n)+(bb_1\bar{c}_1+\ldots+bb_n\bar{c}_n)$$

$$=a\,(a_1\bar{c}_1+\ldots+a_n\bar{c}_n)+b\,(b_1\bar{c}_1+\ldots+b_n\bar{c}_n)$$

$$=a\,(\alpha, \gamma)+b\,(\beta, \gamma) \qquad [\text{by } (1)]$$

(iii) Non-negativity.

$$(\alpha, \alpha)=a_1\bar{a}_1+\ldots+a_n\bar{a}_n$$

$$=|a_1|^2+\ldots+|a_n|^2. \qquad [\text{by } (1)] \qquad \ldots(2)$$

Now a_i is a complex number. Therefore $|a_i|^2 \geqslant 0$. Thus (2) is a sum of n non-negative real numbers and therefore it is $\geqslant 0$. Thus $(\alpha, \alpha) \geqslant 0$. Also $(\alpha, \alpha)=0$

$\Rightarrow \ |a_1|^2+\ldots+|a_n|^2=0$

$\Rightarrow$ each $|a_i|^2=0$ and so each $a_i=0$

$\Rightarrow \ \alpha=0$.

Hence the product defined in (1) is an inner product on $V_n(\mathbf{C})$ and with respect to this inner product $V_n(\mathbf{C})$ is an inner product space.

If α, β are two vectors in $V_n(\mathbf{C})$, then the standard inner product of α and β is also called the **dot product** of α and β and is denoted by $\alpha \cdot \beta$. Thus if $\alpha=(a_1,\ldots, a_n)$, $\beta=(b_1,\ldots, b_n)$ then

$$\alpha \cdot \beta=a_1\bar{b}_1+a_2\bar{b}_2+\ldots+a_n\bar{b}_n.$$

Example 2. On $V_n(\mathbf{R})$ there is an inner product which we call the **standard inner product**.

If $\quad \alpha=(a_1,\ldots, a_n)$, $\beta=(b_1,\ldots, b_n) \in V_n(\mathbf{R})$, then we define

$$(\alpha, \beta)=a_1b_1+\ldots+a_nb_n.$$

As shown in example 1, we can see that this definition satisfies all the postulates of an inner product.

If α, β are two vectors in $V_n(\mathbf{R})$, then the standard inner product of α and β is also called the dot product of α and β and is denoted by $\alpha \cdot \beta$. Thus if $\alpha=(a_1,\ldots, a_n)$, $\beta=(b_1,\ldots, b_n)$, then

$$\alpha \cdot \beta=a_1b_1+a_2b_2+\ldots+a_nb_n.$$

Example 3. If $\alpha=(a_1, a_2)$, $\beta=(b_1, b_2) \in V_2(\mathbf{R})$, let us define

$$(\alpha, \beta)=a_1b_1-a_2b_1-a_1b_2+4a_2b_2. \qquad \ldots(1)$$

We shall show that all the postulates of an inner product hold good in (1).

(i) Symmetry. We have

$$(\beta, \alpha)=b_1a_1-b_2a_1-b_1a_2-4b_2a_2 \qquad [\text{from } (1)]$$

$$=a_1b_1-a_2b_1-a_1b_2+4a_2b_2 \qquad [\because a_1, a_2, b_1, b_2 \in \mathbf{R}]$$

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

(ii) Linearity. If $a, b \in \mathbf{R}$, we have

$$a\alpha+b\beta=a\,(a_1, a_2)+b\,(b_1, b_2)=(aa_1+bb_1, aa_2+bb_2).$$

Let $\gamma=(c_1, c_2) \in V_2(\mathbf{R})$. Then

$$(a\alpha+b\beta, \gamma)=(aa_1+bb_1)\,c_1-(aa_2+bb_2)\,c_1-(aa_1+bb_1)\,c_2$$

$$+4\,(aa_2+bb_2)\,c_2 \qquad [\text{from } (1)]$$

$$=(aa_1c_1-aa_2c_1-aa_1c_2+4aa_2c_2)+(bb_1c_1-bb_2c_1-bb_1c_2+4bb_2c_2)$$

$$=a\,(a_1c_1-a_2c_1-a_1c_2+4a_2c_2)+b\,(b_1c_1-b_2c_1-b_1c_2+4b_2c_2) \qquad [\text{from } (1)]$$

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

(iii) Non-negativity. We have

$$(\alpha, \alpha)=a_1a_1-a_2a_1-a_1a_2+4a_2a_2=a_1^2-2a_1a_2+4a_2^2$$

$$=(a_1-a_2)^2+3a_2^2. \qquad \ldots(2)$$

Now (2) is a sum of two non-negative real numbers. Therefore it is $\geqslant 0$. Thus $(\alpha, \alpha) \geqslant 0$.

Also $\qquad (\alpha, \alpha)=0$

$\Rightarrow (a_1-a_2)^2+3a_2^2=0$

$\Rightarrow (a_1-a_2)^2=0,\ 3a_2^2=0$

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

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

Hence the product defined in (1) is an inner product on $V_2(\mathbf{R})$. Also with respect to this inner product $V_2(\mathbf{R})$ is an inner product space.

Scanned with CamScanner

Note. There can be defined more than one inner products on a vector space. For example, we have the standard inner product on $V_2(\mathbf{R})$ and the inner product defined in example 3.

Example 4. Let $V(\mathbf{C})$ be the vector space of all continuous complex-valued functions on the unit interval, $0 \leqslant t \leqslant 1$. If $f(t), g(t) \in V$, let us define

$$(f(t), g(t)) = \int_0^1 f(t)\, \overline{g(t)}\, dt. \qquad \ldots(1)$$

We shall show that all the postulates of an inner product hold in (1).

(i) **Conjugate Symmetry.** We have

$$(g(t), f(t)) = \int_0^1 g(t)\, \overline{f(t)}\, dt. \qquad \text{[from (1)]}$$

$$\therefore \overline{(g(t), f(t))} = \overline{\left[\int_0^1 g(t) \overline{f(t)}\, dt\right]} = \int_0^1 \overline{\left[g(t)\, \overline{f(t)}\right]}\, dt$$

$$= \int_0^1 \overline{g(t)}\, f(t)\, dt = \int_0^1 f(t)\, \overline{g(t)}\, dt = (f(t), g(t)).$$

(ii) **Linearity.** Let $a, b \in \mathbf{C}$ and $h(t) \in V$. Then

$$(af(t) + bg(t), h(t)) = \int_0^1 [af(t) + bg(t)]\, \overline{h(t)}\, dt$$

$$= a \int_0^1 f(t)\, \overline{h(t)}\, dt + b \int_0^1 g(t)\, \overline{h(t)}\, dt$$

$$= a\, (f(t), h(t)) + b\, (g(t), h(t)).$$

(iii) **Non-negativity.** We have

$$(f(t), f(t)) = \int_0^1 f(t)\, \overline{f(t)}\, dt$$

$$= \int_0^1 |f(t)|^2\, dt. \qquad \ldots(2)$$

Since $|f(t)|^2 \geqslant 0$ for every t lying in the closed interval $[0, 1]$, therefore (2) $\geqslant 0$. Thus $(f(t), f(t)) \geqslant 0$.

Also $\qquad (f(t), f(t)) = 0$

$\Rightarrow \int_0^1 |f(t)|^2\, dt = 0$

$\Rightarrow |f(t)|^2 = 0$ for every t lying in $[0, 1]$

$\Rightarrow f(t) = 0$ for every t lying in $[0, 1]$

$\Rightarrow f(t) = 0$.

Hence the product defined in (1) is an inner product on $V(\mathbf{C})$.

Example 5. Let $V(\mathbf{C})$ be the vector space of all polynomials in t with coefficients in $\mathbf{C}$. If $f(t), g(t) \in V$, let us define

$$(f(t), g(t)) = \int_0^1 f(t)\, \overline{g(t)}\, dt. \qquad \ldots(1)$$

As in example 4, we can show that all the postulates of an inner product are satisfied by (1). Since V is not a finite dimensional vector space, therefore this example gives us an inner product on a vector space which is not finite-dimensional.

Important note. In some books the inner product space is defined as follows :

The vector space V over F is said to be an inner product space if there is defined for any two vectors $\alpha, \beta \in V$ an element $(\alpha \mid \beta) \in F$ such that

(1) $(\alpha \mid \beta) = \overline{(\beta \mid \alpha)}$

(2) $(a\alpha + b\beta \mid \gamma) = a\, (\alpha \mid \gamma) + b\, (\beta \mid \gamma)$.

(3) $(\alpha \mid \alpha) > 0$ *if* $\alpha \neq 0$

for any $\alpha, \beta, \gamma \in V$ and $a, b \in F$.

There is a slight difference in the postulate (3). It can be easily seen that both the definitions are equivalent.

§ 2. Norm or length of a vector in an inner product space. Consider the vector space $V_3(\mathbf{R})$ with standard inner product defined on it. If $\alpha = (a_1, a_2, a_3) \in V_3(\mathbf{R})$, we have

$$(\alpha, \alpha) = a_1^2 + a_2^2 + a_3^2.$$

Now we know that in the three dimensional Euclidean space $\sqrt{(a_1^2 + a_2^2 + a_3^2)}$ is the length of the vector $\alpha = (a_1, a_2, a_3)$. Taking motivation from this fact, we make the following definition.

Definition. *Let V be an inner product space. If $\alpha \in V$, then the norm or the length of the vector α, written as $\|\alpha\|$, is defined as the positive square root of (α, α) i.e.,*

$$\|\alpha\| = \sqrt{(\alpha, \alpha)}. \qquad \textbf{(Nagarjuna 1978)}$$

Unit vector. Definition. *Let V be an inner product space. If $\alpha \in V$ is such that $\|\alpha\| = 1$, then α is called a unit vector. Thus in an inner product space a vector is called a unit vector if its length is 1.*

Theorem 1. *In an inner product space $V(F)$, prove that*

(*i*) $(a\alpha - b\beta, \gamma) = a\, (\alpha, \gamma) - b\, (\beta, \gamma)$

(*ii*) $(\alpha, a\beta + b\gamma) = \bar{a}\, (\alpha, \beta) + \bar{b}\, (\alpha, \gamma)$.

Proof. (i) We have

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

Scanned with CamScanner

$$=a\,(\alpha,\gamma)+(-b)\,(\beta,\gamma) \quad \text{[by linearity property]}$$
$$=a\,(\alpha,\gamma)-b\,(\beta,\gamma).$$

(ii) $(\alpha, a\beta+b\gamma)=\overline{(a\beta+b\gamma, \alpha)}$ [by conjugate symmetry]
$$=\overline{a(\beta,\alpha)+b\,(\gamma,\alpha)} \quad \text{[by linearity property]}$$
$$=\overline{a\,(\beta,\alpha)}+\overline{b\,(\gamma,\alpha)}=\bar{a}\,\overline{(\beta,\alpha)}+\bar{b}\,\overline{(\gamma,\alpha)}$$
$$=\bar{a}\,(\alpha,\beta)+\bar{b}\,(\alpha,\gamma).$$

Note 1. If $F=\mathbf{R}$, then the result (ii) can be simply read as
$$(\alpha, a\beta+b\gamma)=a\,(\alpha,\beta)+b\,(\alpha,\gamma).$$

Note 2. Similarly it can be proved that
$$(\alpha, a\beta-b\gamma)=\bar{a}\,(\alpha,\beta)-\bar{b}\,(\alpha,\gamma).$$

Also $(\alpha, \beta+\gamma)=(\alpha, 1\beta+1\gamma)=\bar{1}\,(\alpha,\beta)+\bar{1}\,(\alpha,\gamma)$
$$=(\alpha,\beta)+(\alpha,\gamma).$$

Theorem 2. *In an inner product space $V(F)$, prove that*

(i) $\|\alpha\| \geqslant 0$; *and* $\|\alpha\|=0$ *if and only if* $\alpha=\mathbf{0}$.

(ii) $\|a\alpha\|=|a|.\|\alpha\|$.

Proof. (i) We have

$\|\alpha\|=\sqrt{(\alpha,\alpha)}$ [by def. of norm]

$\Rightarrow \|\alpha\|^2=(\alpha,\alpha)$

$\Rightarrow \|\alpha\|^2 \geqslant 0$ [$\because (\alpha,\alpha) \geqslant 0$]

$\Rightarrow \|\alpha\| \geqslant 0.$

Also $(\alpha,\alpha)=0$ iff $\alpha=\mathbf{0}$.

$\therefore \|\alpha\|^2=0$ iff $\alpha=\mathbf{0}$ *i.e.* $\|\alpha\|=0$ iff $\alpha=\mathbf{0}$.

Thus in an inner product space, $\|\alpha\|>0$ iff $\alpha\neq\mathbf{0}$.

(ii) We have $\|a\alpha\|^2=(a\alpha, a\alpha)$ [by def. of norm]

$=a\,(\alpha, a\alpha)$ [by linearity property]

$=a\bar{a}\,(\alpha,\alpha)$ [by theorem 1]

$=|a|^2.\|\alpha\|^2.$

Thus $\|a\alpha\|^2=|a|^2.\|\alpha\|^2$.

Taking square root, we get $\|a\alpha\|=|a|.\|\alpha\|$.

Note. If α is any non-zero vector of an inner product space V, then $\dfrac{1}{\|\alpha\|}\alpha$ is a unit vector in V. We have $\|\alpha\|\neq 0$ because $\alpha\neq\mathbf{0}$.

Therefore $\dfrac{1}{\|\alpha\|}\alpha\in V$.

Now $\left(\dfrac{1}{\|\alpha\|}\alpha, \dfrac{1}{\|\alpha\|}\alpha\right)=\dfrac{1}{\|\alpha\|}\left(\alpha, \dfrac{1}{\|\alpha\|}\alpha\right)$
$$=\frac{1}{\|\alpha\|^2}(\alpha,\alpha)=\frac{1}{\|\alpha\|^2}\|\alpha\|^2=1.$$

Therefore $\left\|\dfrac{\alpha}{\|\alpha\|}\right\|=1$ and thus $\dfrac{\alpha}{\|\alpha\|}$ is a unit vector.

For example if $\alpha=(2,1,2)$ is a vector in $V_3(\mathbf{R})$ with standard inner product, then $\|\alpha\|=\sqrt{(\alpha,\alpha)}=\sqrt{(4+1+4)}=3$.

Therefore $\frac{1}{3}(2,1,2)$ *i.e.*, $\left(\frac{2}{3},\frac{1}{3},\frac{2}{3}\right)$ is a unit vector.

Theorem 3. Schwarz's Inequality. *In an inner product space $V(F)$, prove that* $|(\alpha,\beta)| \leqslant \|\alpha\|.\|\beta\|$.

(Meerut 1981, 83, 89, 92, 93P; Madras 83; Andhra 92; S.V.U. Tirupati 90, 93; Nagarjuna 78)

Proof. If $\alpha=\mathbf{0}$, then $\|\alpha\|=0$. Also in that case
$$(\alpha,\beta)=(\mathbf{0},\beta)=(0\mathbf{0},\beta)=0\,(\mathbf{0},\beta)=0.$$
$\therefore |(\alpha,\beta)|=0.$

Thus if $\alpha=\mathbf{0}$, then $|(\alpha,\beta)|=0$ and $\|\alpha\|\,\|\beta\|=0$.

$\therefore$ the inequality $|(\alpha,\beta)| \leqslant \|\alpha\|.\|\beta\|$ is valid.

Now let $\alpha\neq\mathbf{0}$. Then $\|\alpha\|>0$. Therefore $\dfrac{1}{\|\alpha\|^2}$ is a positive real number. Consider the vector
$$\gamma=\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha. \text{ We have}$$
$$(\gamma,\gamma)=\left(\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha, \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)$$
$$=\left(\beta, \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)-\frac{(\beta,\alpha)}{\|\alpha\|^2}\left(\alpha, \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)$$
[by linearity property]
$$=(\beta,\beta)-\overline{\frac{(\beta,\alpha)}{\|\alpha\|^2}}(\beta,\alpha)-\frac{(\beta,\alpha)}{\|\alpha\|^2}(\alpha,\beta)+\frac{(\beta,\alpha)\overline{(\beta,\alpha)}}{\|\alpha\|^2.\|\alpha\|^2}(\alpha,\alpha)$$
$$=\|\beta\|^2-\frac{\overline{(\beta,\alpha)}\,(\beta,\alpha)}{\|\alpha\|^2}-\frac{\overline{(\alpha,\beta)}\,(\alpha,\beta)}{\|\alpha\|^2}+\frac{\overline{(\beta,\alpha)}\,(\beta,\alpha)}{\|\alpha\|^2\,\|\alpha\|^2}.\|\alpha\|^2$$
$$=\|\beta\|^2-\frac{(\alpha,\beta)\overline{(\alpha,\beta)}}{\|\alpha\|^2}, \text{ the second and the fourth terms cancel}$$
$$=\|\beta\|^2-\frac{|(\alpha,\beta)|^2}{\|\alpha\|^2}. \quad [\because z\bar{z}=|z|^2 \text{ if } z\in\mathbf{C}]$$

But $(\gamma,\gamma)=\|\gamma\|^2 \geqslant 0$.
$$\therefore \|\beta\|^2-\frac{|(\alpha,\beta)|^2}{\|\alpha\|^2} \geqslant 0$$
or $\|\beta\|^2.\|\alpha\|^2 \geqslant |(\alpha,\beta)|^2$

or $|(\alpha,\beta)| \leqslant \|\alpha\|.\|\beta\|$, taking square root of both sides.

Schwarz's inequality has very important applications in mathematics.

Scanned with CamScanner

Theorem 4. Triangle inequality. *If α, β are vectors in an inner product space V, prove that*

$$\| \alpha+\beta \| \leqslant \| \alpha \| + \| \beta \|.$$

(Meerut 1974, 76, 77, 81, 90, 92, 93)

Proof. We have

$\| \alpha+\beta \|^2=(\alpha+\beta, \alpha+\beta)$ [by def. of norm]

$=(\alpha, \alpha+\beta)+(\beta, \alpha+\beta)$ [by linearity property]

$=(\alpha, \alpha)+(\alpha, \beta)+(\beta, \alpha)+(\beta, \beta)$ [by theorem 1]

$=\| \alpha \|^2+(\alpha, \beta)+\overline{(\alpha, \beta)}+\| \beta \|^2$ [$\because$ $(\beta, \alpha)=\overline{(\alpha, \beta)}$]

$=\| \alpha \|^2+2\text{Re }(\alpha, \beta)+\| \beta \|^2$ [$\because$ $z+\bar{z}=2\text{Re } z$]

$\leqslant \| \alpha \|^2+2 \mid (\alpha, \beta) \mid+\| \beta \|^2$ [$\because$ $\text{Re } z \leqslant \mid z \mid$]

$\leqslant \| \alpha \|^2+2 \| \alpha \| . \| \beta \|+\| \beta \|^2$

[$\because$ by Schwarz inequality $\mid (\alpha, \beta) \mid \leqslant \| \alpha \| . \| \beta \|$]

$=(\| \alpha \|+\| \beta \|)^2.$

Thus $\quad \| \alpha+\beta \|^2 \leqslant (\| \alpha \|+\| \beta \|)^2.$

Taking square root of both sides, we get

$$\| \alpha+\beta \| \leqslant \| \alpha \|+\| \beta \|.$$

Geometrical interpretation. Let α, β be the vectors in the inner product space $V_3(\mathbf{R})$ with standard inner product defined on it. Suppose the vectors α, β represent the sides AB and BC respectively of a triangle ABC in the three dimensional Euclidean space. Then $\| \alpha \|=AB$, $\| \beta \|=BC$. Also the vector $\alpha+\beta$ represents the side AC of the triangle ABC and $\| \alpha+\beta \|=AC$. Then from the above inequality, we have $AC \leqslant AB+BC$.

Special cases of Schwarz's inequality.

Case I. Consider the vector space $V_n(\mathbf{C})$ with standard inner product defined on it.

Let $\alpha=(a_1, a_2, \ldots, a_n)$ and $\beta=(b_1, b_2, \ldots, b_n) \in V_n(\mathbf{C})$.

Then $a_1, \ldots, a_n, b_1, \ldots, b_n$ are all complex numbers.

We have $(\alpha, \beta)=a_1\bar{b}_1+a_2\bar{b}_2+\ldots+a_n\bar{b}_n$.

$\therefore \quad \mid (\alpha, \beta) \mid^2=\mid a_1\bar{b}_1+a_2\bar{b}_2+\ldots+a_n\bar{b}_n \mid^2.$

Also $\| \alpha \|^2=(\alpha, \alpha)=a_1\bar{a}_1+\ldots+a_n\bar{a}_n$

$=\mid a_1 \mid^2+\ldots+\mid a_n \mid^2.$

Similarly $\| \beta \|^2=\mid b_1 \mid^2+\ldots+\mid b_n \mid^2.$

By Schwarz's inequality, we have

$$\mid (\alpha, \beta) \mid^2 \leqslant \| \alpha \|^2 . \| \beta \|^2.$$

$\therefore$ If $a_1, \ldots, a_n, b_1, \ldots, b_n$ are complex numbers, then

$$\mid a_1\bar{b}_1+a_2\bar{b}_2+\ldots+a_n\bar{b}_n \mid^2 \leqslant (\mid a_1 \mid^2+\ldots+\mid a_n \mid^2)(\mid b_1 \mid^2+\ldots+\mid b_n \mid^2).$$

This inequality is known as **Cauchy's Inequality.**

If $a_1, \ldots, a_n, b_1, \ldots, b_n$ are all real numbers, then this inequality gives that

$$(a_1b_1+a_2b_2+\ldots+a_nb_n)^2 \leqslant (a_1^2+\ldots+a_n^2)(b_1^2+\ldots+b_n^2).$$

Case II. Consider the vector space $V(\mathbf{C})$ of all continuous complex-valued functions on the unit interval $0 \leqslant t \leqslant 1$, with inner product defined by $(f(t), g(t))=\int_0^1 f(t)\,\overline{g(t)}\,dt$.

We have $\| f(t) \|^2=(f(t), f(t))=\int_0^1 f(t)\,\overline{f(t)}\,dt$

$$=\int_0^1 \mid f(t) \mid^2 dt.$$

Similarly $\| g(t) \|^2=\int_0^1 \mid g(t) \mid^2 dt.$

Also $\mid (f(t), g(t)) \mid^2=\left| \int_0^1 f(t)\,\overline{g(t)}\,dt \right|^2.$

By Schwarz's inequality, we have

$$\mid (f(t), g(t)) \mid^2 \leqslant \| f(t) \|^2 . \| g(t) \|^2.$$

Therefore if $f(t), g(t)$ are continuous complex valued functions on the unit interval $[0, 1]$, then

$$\left| \int_0^1 f(t)\,\overline{g(t)}\,dt \right|^2 \leqslant \left(\int_0^1 \mid f(t) \mid^2 dt \right)\left(\int_0^1 \mid g(t) \mid^2 dt \right).$$

Case III. Consider the vector space $V_3(\mathbf{R})$ with standard inner product defined on it *i.e.*, if

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

then $\quad (\alpha, \beta)=a_1b_1+a_2b_2+a_3b_3. \quad \ldots(1)$

We see that (1) is nothing but the dot product of two vectors α and β in three dimensional Euclidean space. If θ is the angle between the non-zero vectors α and β, then we know that

$$\cos^2 \theta=\frac{(a_1b_1+a_2b_2+a_3b_3)^2}{(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)}=\frac{\{(\alpha, \beta)\}^2}{(\alpha, \alpha)(\beta, \beta)}$$

$$=\frac{\mid (\alpha, \beta) \mid^2}{\| \alpha \|^2 \| \beta \|^2}.$$

[$\because$ if (α, β) is real then $\{(\alpha, \beta)\}^2=\mid (\alpha, \beta) \mid^2$]

But by Schwarz's inequality, we have

$$\mid (\alpha, \beta) \mid^2 \leqslant \| \alpha \|^2 . \| \beta \|^2.$$

$\therefore \quad \cos^2 \theta \leqslant \dfrac{\| \alpha \|^2 \| \beta \|^2}{\| \alpha \|^2 \| \beta \|^2}$ *i.e.*, $\cos^2 \theta \leqslant 1.$

Thus the absolute value of the cosine of a real angle cannot be greater than 1.

Scanned with CamScanner

Normed vector space. Definition. *Let $V(F)$ be a vector space where F is either the field of real numbers or the field of complex numbers. Then V is said to be a normed vector space if to each vector α there corresponds a real number denoted by $\| \alpha \|$ called the norm of α in such a manner that*

(1) $\| \alpha \| \geqslant 0$ *and* $\| \alpha \| = 0 \Rightarrow \alpha = 0$.

(2) $\| a\alpha \| = | a | . \| \alpha \|, \ \forall \ a \in F$.

(3) $\| \alpha + \beta \| \leqslant \| \alpha \| + \| \beta \|, \ \forall \ \alpha, \beta \in V$.

We have shown in theorems 2 and 4 that the norm of an inner product space satisfies all the three conditions of the norm of a normed vector space. Hence *every inner product space is a normed vector space.*

Distance in an inner product space.

Definition. Let $V(F)$ be an inner product space. Then we define the distance (α, β) between two vectors α and β by

$$d(\alpha, \beta) = \| \alpha - \beta \| = \sqrt{[(\alpha - \beta, \alpha - \beta)]}.$$

Theorem 5. *In an inner product space $V(F)$ we define the distance $d(\alpha, \beta)$ from α to β by $d(\alpha, \beta) = \| \alpha - \beta \|$. Prove that*

(1) $d(\alpha, \beta) \geqslant 0$ *and* $d(\alpha, \beta) = 0$ *iff* $\alpha = \beta$.

(2) $d(\alpha, \beta) = d(\beta, \alpha)$.

(3) $d(\alpha, \beta) \leqslant d(\alpha, \gamma) + d(\gamma, \beta)$. **[triangle inequality]**

(4) $d(\alpha, \beta) = d(\alpha + \gamma, \beta + \gamma)$. **(Meerut 1984)**

Proof. (1) We have $d(\alpha, \beta) = \| \alpha - \beta \|$. [by definition]

Now $\| \alpha - \beta \| \geqslant 0$ and $\| \alpha - \beta \| = 0$ if and only if $\alpha - \beta = 0$.

$\therefore$ $d(\alpha, \beta) \geqslant 0$ and $d(\alpha, \beta) = 0$ if and only if $\alpha = \beta$.

(2) We have $d(\alpha, \beta) = \| \alpha - \beta \|$ [by def.]

$= \| (-1)(\beta - \alpha) \|$

$= | -1 | \, \| \beta - \alpha \|$ $[\because \ \| a\alpha \| = | a | \, \| \alpha \|]$

$= \| \beta - \alpha \| = d(\beta, \alpha)$.

(3) We have $d(\alpha, \beta) = \| \alpha - \beta \| = \| (\alpha - \gamma) + (\gamma - \beta) \|$

$\leqslant \| \alpha - \gamma \| + \| \gamma - \beta \|$ [by theorem 4]

$= d(\alpha, \gamma) + d(\gamma, \beta)$.

$\therefore$ $d(\alpha, \beta) \leqslant d(\alpha, \gamma) + d(\gamma, \beta)$.

(4) We have $d(\alpha, \beta) = \| (\alpha - \beta) \| = \| (\alpha + \gamma) - (\beta + \gamma) \|$

$= d(\alpha + \gamma, \beta + \gamma)$.

Matrix of an inner product.

Before defining the matrix of an inner product, we shall define some special types of matrices over the complex field **C**.

Conjugate transpose of a Matrix. Definition. *Let $A = [a_{ij}]_{n \times n}$ be a square matrix of order n over the field C of complex numbers. Then the matrix $[\bar{a}_{ji}]_{n \times n}$ is called the conjugate transpose of A and we shall denote it by A^*.*

Thus in order to obtain A^* from A, we should first replace each element of A by its conjugate complex and then we should write the transpose of the new matrix.

If in place of the field **C** we take the field **R** of the real numbers, then $\bar{a}_{ji} = a_{ji}$. So in this case A^* will simply be the transpose of the matrix A.

If $A = A^*$, then the matrix A is said to be a **self-adjoint** matrix or a **Hermitian matrix**.

Symmetric matrix. Definition. *A square matrix A over a field F is said to be a symmetric matrix if it is equal to its transpose i.e., if* $A = A^T$.

Obviously a Hermitian matrix over the field of real numbers is a symmetric matrix.

Theorem. *If $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an ordered basis of a finite dimensional vector space V, then an inner product on V is completely determined by the values which it takes on pairs of vectors in B.*

(Meerut 1976)

Proof. Suppose we are given a particular inner product on V. We shall show that this inner product on V is completely determined by the values

$$g_{ij} = (\alpha_j, \alpha_i) \text{ where } i = 1, 2, \ldots\ldots, n; \ j = 1, 2, \ldots, n.$$

Let $\alpha = \sum_{j=1}^{n} x_j \alpha_j$ and $\beta = \sum_{i=1}^{n} y_i \alpha_i$ be any two vectors in V. Then

$(\alpha, \beta) = \left(\sum_{j=1}^{n} x_j \alpha_j, \beta \right) = \sum_{j=1}^{n} x_j (\alpha_j, \beta)$, by linearity property of the inner product

$= \sum_{j=1}^{n} x_j \left(\alpha_j, \sum_{i=1}^{n} y_i \alpha_i \right) = \sum_{j=1}^{n} x_j \sum_{i=1}^{n} \bar{y}_i (\alpha_j, \alpha_i)$ [See theorem 1, part (ii) on page 289]

$= \sum_{j=1}^{n} \sum_{i=1}^{n} \bar{y}_i \, g_{ij} \, x_j$ ...(1)

$= Y^* G X$,

Scanned with CamScanner

where X, Y are the coordinate matrices of α, β in the ordered basis B and G is the matrix $[g_{ij}]_{n\times n}$.

From (1) we observe that the inner product (α, β) is completely determined by the matrix G *i.e.*, by the scalars g_{ij}. Hence the result of the theorem.

Definition. *Let $B=\{\alpha_1, \ldots, \alpha_n\}$ be an ordered basis for an n-dimensional inner product space V. The matrix $G=[g_{ij}]_{n\times n}$, where $g_{ij}=(\alpha_j, \alpha_i)$, is called the matrix of the underlined inner product in the ordered basis B.*

We observe that $\overline{g_{ij}}=\overline{(\alpha_j, \alpha_i)}=(\alpha_i, \alpha_j)=g_{ji}$. Therefore the matrix G is such that $G^*=G$. Thus G is a Hermitian matrix.

Further we know that in an inner product space, $(\alpha, \alpha)>0$ if $\alpha\neq 0$. Therefore from the relation (1) given above we observe that the matrix G is such that

$$X^*GX>0, \quad X\neq 0. \qquad \ldots(2)$$

From the relation (2) we conclude that the matrix G is invertible. For if G is not invertible then there exists an $X\neq 0$ such that $GX=0$. For any such X the relation (2) is impossible. Hence G must be invertible.

In the last, from the relation (2) we observe that if $x_1, x_2, \ldots, x_n$ are scalars not all of which are zero, then

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{x}_i\, g_{ij}\, x_j>0. \qquad \ldots(3)$$

Now suppose that out of the n scalars $x_1, \ldots, x_n$ we take $x_i=1$ and each of the remaining $n-1$ scalars is taken as 0. Then from (3) we conclude that $g_{ii}>0$. Thus $g_{ii}>0$ for each $i=1, \ldots, n$. Hence each entry along the principal diagonal of the matrix G is positive.

Solved Examples

Example 1. *Show that we can always define an inner product on a finite dimensional vector space real or complex.*

Solution. Let V be a finite dimensional vector space over the field F real or complex.

Let $\quad B=\{\alpha_1, \ldots, \alpha_n\}$ be a basis for V.

Let $\alpha, \beta\in V$. Then we can write $\alpha=a_1\alpha_1+\ldots+a_n\alpha_n$,

and $\quad \beta=b_1\alpha_1+\ldots+b_n\alpha_n$

where $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are uniquely determined elements of F. Let us define

$$(\alpha, \beta)=a_1\bar{b}_1+\ldots+a_n\bar{b}_n. \qquad \ldots(1)$$

We shall show that (1) satisfies all the conditions for an inner product.

(i) **Conjugate Symmetry.** We have

$$(\beta, \alpha)=b_1\bar{a}_1+\ldots+b_n\bar{a}_n.$$

$$\therefore \quad \overline{(\beta, \alpha)}=\overline{(b_1\bar{a}_1+\ldots+b_n\bar{a}_n)}=\bar{b}_1a_1+\ldots+\bar{b}_na_n$$
$$=a_1\bar{b}_1+\ldots+a_n\bar{b}_n=(\alpha, \beta).$$

(ii) **Linearity.** Let $\gamma=c_1\alpha_1+\ldots+c_n\alpha_n\in V$ and $a, b\in F$. We have

$$a\alpha+b\beta=a(a_1\alpha_1+\ldots+a_n\alpha_n)+b(b_1\alpha_1+\ldots+b_n\alpha_n)$$
$$=(aa_1+bb_1)\alpha_1+\ldots+(aa_n+bb_n)\alpha_n.$$

$$\therefore \quad (a\alpha+b\beta, \gamma)=(aa_1+bb_1)\bar{c}_1+\ldots+(aa_n+bb_n)\bar{c}_n$$
$$=a(a_1\bar{c}_1+\ldots+a_n\bar{c}_n)+b(b_1\bar{c}_1+\ldots+b_n\bar{c}_n)$$
$$=a(\alpha, \gamma)+b(\beta, \gamma).$$

(iii) **Non-negativity.** We have

$$(\alpha, \alpha)=a_1\bar{a}_1+\ldots+a_n\bar{a}_n=|a_1|^2+\ldots+|a_n|^2\geqslant 0.$$

Also $\quad (\alpha, \alpha)=0$

$$\Rightarrow |a_1|^2+\ldots+|a_n|^2=0$$
$$\Rightarrow |a_1|^2=0, \ldots, |a_n|^2=0$$
$$\Rightarrow a_1=0, \ldots, a_n=0$$
$$\Rightarrow \alpha=0.$$

Hence (1) is an inner product on V.

Example 2. *In $V_2(F)$ define for $\alpha=(a_1, a_2)$ and $\beta=(b_1, b_2)$,*

$$(\alpha, \beta)=2a_1\bar{b}_1+a_1\bar{b}_2+a_2\bar{b}_1+a_2\bar{b}_2.$$

Show that this defines an inner product on $V_2(F)$.

Solution. (1) Conjugate Symmetry. We have

$$(\beta, \alpha)=2b_1\bar{a}_1+b_1\bar{a}_2+b_2\bar{a}_1+b_2\bar{a}_2.$$

$$\therefore \quad \overline{(\beta, \alpha)}=\overline{(2b_1\bar{a}_1+b_1\bar{a}_2+b_2\bar{a}_1+b_2\bar{a}_2)}$$
$$=2\bar{b}_1a_1+\bar{b}_1a_2+\bar{b}_2a_1+\bar{b}_2a_2$$
$$=2a_1\bar{b}_1+a_1\bar{b}_2+a_2\bar{b}_1+a_2\bar{b}_2=(\alpha, \beta).$$

(2) **Linearity.** Let $a, b\in F$ and $\gamma=(c_1, c_2)\in V_2(F)$. Then

$a\alpha+b\beta=a(a_1, a_2)+b(b_1, b_2)=(aa_1+bb_1, aa_2+bb_2)$.

$$\therefore \quad (a\alpha+b\beta, \gamma)=2(aa_1+ab_1)\bar{c}_1+(aa_1+bb_1)\bar{c}_2$$
$$+(aa_2+bb_2)\bar{c}_1+(aa_2+bb_2)\bar{c}_2$$
$$=a(2a_1\bar{c}_1+a_1\bar{c}_2+a_2\bar{c}_1+a_2\bar{c}_2)+b(2b_1\bar{c}_1+b_1\bar{c}_2+b_2\bar{c}_1+b_2\bar{c}_2)$$
$$=a(\alpha, \gamma)+b(\beta, \gamma).$$

(3) **Non-negativity.** We have

$$(\alpha, \alpha)=2a_1\bar{a}_1+a_1\bar{a}_2+a_2\bar{a}_1+a_2\bar{a}_2$$
$$=a_1\bar{a}_1+(a_1+a_2)(\bar{a}_1+\bar{a}_2)=|a_1|^2+(a_1+a_2)\overline{(a_1+a_2)}$$
$$=|a_1|^2+|a_1+a_2|^2\geqslant 0.$$

Scanned with CamScanner

Also $(\alpha, \alpha)=0 \Rightarrow |a_1|^2+|a_1+a_2|^2=0$
$\Rightarrow a_1=0,\ a_1+a_2=0$
$\Rightarrow a_1=0,\ a_2=0$
$\Rightarrow \alpha=0.$

Hence the result

Example 3. *Let V be a vector space over F. Show that the sum of two inner products on V is an inner product on V. Is the difference of two inner products an inner product? Show that a positive multiple of an inner product is an inner product.* **(Meerut 1979)**

Solution. Let p and q be two inner products on $V(F)$. Then p and q are both mappings from $V \times V$ into F and are such that $\forall\ \alpha, \beta, \gamma \in V$ and $\forall\ a, b \in F$, we have

$$p(\alpha, \beta)=\overline{p(\beta, \alpha)},$$
$$p(a\alpha+b\beta, \gamma)=a\, p(\alpha, \gamma)+b\, p(\beta, \gamma),$$
$$p(\alpha, \alpha)>0 \text{ if } \alpha \neq 0;$$

and similar results for the inner product q. Note that here $p(\alpha, \beta)$ is an element of F which is the image of the ordered pair (α, β) under the inner product p.

Now let us define the sum $p+q$ of the inner products p and q by the relation

$$(p+q)(\alpha, \beta)=p(\alpha, \beta)+q(\alpha, \beta) \ \forall\ (\alpha, \beta) \in V \times V.$$

We shall show that $p+q$ is also an inner product on V *i.e.* all the postulates of an inner product hold for $p+q$.

(i) Conjugate symmetry. We have

$(p+q)(\beta, \alpha)=p(\beta, \alpha)+q(\beta, \alpha)$. [by def. of $p+q$]

$\therefore\ \overline{(p+q)(\beta, \alpha)}=\overline{p(\beta, \alpha)+q(\beta, \alpha)}$

$=\overline{p(\beta, \alpha)}+\overline{q(\beta, \alpha)}$ [$\because\ \overline{z_1+z_2}=\bar{z}_1+\bar{z}_2$]

$=p(\alpha, \beta)+q(\alpha, \beta)$ [$\because$ both p and q are inner products]

$=(p+q)(\alpha, \beta)$ [by def. of $p+q$]

(ii) Linearity. We have

$(p+q)(a\alpha+b\beta, \gamma)=p(a\alpha+b\beta, \gamma)+q(a\alpha+b\beta, \gamma)$ [by def. of $p+q$]

$=ap(\alpha, \gamma)+bp(\beta, \gamma)+aq(\alpha, \gamma)+bq(\beta, \gamma)$ [by linearity of p and q]

$=a\{p(\alpha, \gamma)+q(\alpha, \gamma)\}+b\{p(\beta, \gamma)+q(\beta, \gamma)\}$

$=a\{(p+q)(\alpha, \gamma)\}+b\{(p+q)(\beta, \gamma)\}.$

(iii) Non-negativity. Suppose $\alpha \neq 0$. We have

$$(p+q)(\alpha, \alpha)=p(\alpha, \alpha)+q(\alpha, \alpha).$$

Since $p(\alpha, \alpha)>0$, and $q(\alpha, \alpha)>0$, therefore

$$(p+q)(\alpha, \alpha)>0.$$

Hence $p+q$ is an inner product on V.

If we define $p-q$ by the relation

$$(p-q)(\alpha, \beta)=p(\alpha, \beta)-q(\alpha, \beta),$$

then $p-q$ is not necessarily an inner product on V.

In this case the postulate of non-negativity might not be satisfied. Note that even if $p(\alpha, \alpha)>0$ and $q(\alpha, \alpha)>0$, yet $(p-q)(\alpha, \alpha)=p(\alpha, \alpha)-q(\alpha, \alpha)$ may come out to be negative.

Again suppose we define $np=p+p+p+\ldots$upto n terms, where n is a +ive integer. Then np is an inner product on V. As we have proved above, $p+p=2p$ is an inner product on V. Then $2p+p=3p$ is also an inner product, $3p+p=4p$ is also an inner product and so on. By induction we can show that np is an inner product on V.

Example 4. *Let V be an inner product space.*

(a) *Show that* $(0, \beta)=0$ *for all* β *in* V.

(b) *Show that if* $(\alpha, \beta)=0$ *for all* β *in* V, *then* $\alpha=0$.

Solution. (a) We have $(0, \beta)=(00, \beta)\ \forall\ \beta \in V$
$=0\ (0, \beta)=0.$

(b) Let $(\alpha, \beta)=0$ for all β in V. Then taking $\beta=\alpha$, we get
$(\alpha, \alpha)=0$
$\Rightarrow \alpha=0.$

Example 5. *Let V be an inner product space, and* α, β *be vectors in V. Show that* $\alpha=\beta$ *if and only if* $(\alpha, \gamma)=(\beta, \gamma)$ *for every* γ *in V.* **(Meerut 1979; S.V.U. Tirupati 93)**

Solution. Let $\alpha=\beta$. Then for every γ in V we have
$(\alpha, \gamma)=(\beta, \gamma).$

Conversely, let $(\alpha, \gamma)=(\beta, \gamma)\ \forall\ \gamma \in V$.

Then $(\alpha, \gamma)-(\beta, \gamma)=0\ \forall\ \gamma \in V$
$\Rightarrow (\alpha-\beta, \gamma)=0\ \forall\ \gamma \in V$
$\Rightarrow (\alpha-\beta, \alpha-\beta)=0$ [taking $\gamma=\alpha-\beta$]
$\Rightarrow \alpha-\beta=0$ [$\because\ (\alpha, \alpha)=0 \Rightarrow \alpha=0$]
$\Rightarrow \alpha=\beta.$

Example 7. *If* α *and* β *are vectors in an inner product space then show that* $\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2\|\alpha\|^2+2\|\beta\|^2$.
(Parallelogram law) (Meerut 1969, 74; Madras 81, 83; Nagarjuna 78; Andhra 81; S.V.U. Tirupati 93)

Solution. We have

$$\|\alpha+\beta\|^2=(\alpha+\beta,\alpha+\beta) \qquad \text{[by def. of norm]}$$
$$=(\alpha,\alpha+\beta)+(\beta,\alpha+\beta) \qquad \text{[by linearity property]}$$
$$=(\alpha,\alpha)+(\alpha,\beta)+(\beta,\alpha)+(\beta,\beta)$$
$$=\|\alpha\|^2+(\alpha,\beta)+(\beta,\alpha)+\|\beta\|^2 \qquad \ldots(1)$$

Also $\|\alpha-\beta\|^2=(\alpha-\beta,\alpha-\beta)=(\alpha,\alpha-\beta)-(\beta,\alpha-\beta)$
$$=(\alpha,\alpha)-(\alpha,\beta)-(\beta,\alpha)+(\beta,\beta)$$
$$=\|\alpha\|^2-(\alpha,\beta)-(\beta,\alpha)+\|\beta\|^2. \qquad \ldots(2)$$

Adding (1) and (2), we get
$$\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2\|\alpha\|^2+2\|\beta\|^2.$$

Geometrical interpretation. Let α and β be vectors in the vector space $V_2(\mathbf{R})$ with standard inner product defined on it. Suppose the vector α is represented by the side AB and the vector β by the side BC of a parallelogram $ABCD$. Then the vectors $\alpha+\beta$ and $\alpha-\beta$ represent the diagonals AC and DB of the parallelogram.

$$\therefore \quad AC^2+DB^2=2AB^2+2BC^2,$$

i.e., the sum of the squares of the sides of a parallelogram is equal to the sum of the squares of its diagonals.

Example 7. *If α, β are vectors in an inner product space $V(F)$ and a, $b \in F$, then prove that*

(i) $\|a\alpha+b\beta\|^2=|a|^2\|\alpha\|^2+a\bar{b}(\alpha,\beta)+\bar{a}b(\beta,\alpha)+|b|^2\|\beta\|^2$,

(ii) $Re(\alpha,\beta)=\frac{1}{4}\|\alpha+\beta\|^2-\frac{1}{4}\|\alpha-\beta\|^2$.

Solution. (i) We have

$\|a\alpha+b\beta\|^2=(a\alpha+b\beta,a\alpha+b\beta)$
$$=a(\alpha,a\alpha+b\beta)+b(\beta,a\alpha+b\beta)$$
$$=a\{\bar{a}(\alpha,\alpha)+\bar{b}(\alpha,\beta)\}+b\{\bar{a}(\beta,\alpha)+\bar{b}(\beta,\beta)\}$$
$$=a\bar{a}(\alpha,\alpha)+a\bar{b}(\alpha,\beta)+b\bar{a}(\beta,\alpha)+b\bar{b}(\beta,\beta)$$
$$=|a|^2\|\alpha\|^2+a\bar{b}(\alpha,\beta)+\bar{a}b(\beta,\alpha)+|b|^2\|\beta\|^2.$$

(ii) We have

$$\|\alpha+\beta\|^2=(\alpha+\beta,\alpha+\beta)=(\alpha,\alpha+\beta)+(\beta,\alpha+\beta)$$
$$=(\alpha,\alpha)+(\alpha,\beta)+(\beta,\alpha)+(\beta,\beta)$$
$$=\|\alpha\|^2+(\alpha,\beta)+\overline{(\alpha,\beta)}+\|\beta\|^2$$
$$=\|\alpha\|^2+2\text{Re}(\alpha,\beta)+\|\beta\|^2. \qquad \ldots(1)$$

Also $\|\alpha-\beta\|^2=(\alpha-\beta,\alpha-\beta)=(\alpha,\alpha-\beta)-(\beta,\alpha-\beta)$
$$=(\alpha,\alpha)-(\alpha,\beta)-(\beta,\alpha)+(\beta,\beta)$$
$$=\|\alpha\|^2-\{(\alpha,\beta)+(\beta,\alpha)\}+\|\beta\|^2$$
$$=\|\alpha\|^2-\{(\alpha,\beta)+\overline{(\alpha,\beta)}\}+\|\beta\|^2$$
$$=\|\alpha\|^2-2\text{ Re}(\alpha,\beta)+\|\beta\|^2 \qquad \ldots(2)$$

Subtracting (2) from (1), we get
$$\|\alpha+\beta\|^2-\|\alpha-\beta\|^2=4\text{ Re}(\alpha,\beta).$$
$$\therefore \quad \text{Re}(\alpha,\beta)=\tfrac{1}{4}\|\alpha+\beta\|^2-\tfrac{1}{4}\|\alpha-\beta\|^2.$$

Note. If $F=\mathbf{R}$, then $\text{Re}(\alpha,\beta)=(\alpha,\beta)$.

Example 8. *Prove that if α and β are vectors in a unitary space, then*

(i) $4(\alpha,\beta)=\|\alpha+\beta\|^2-\|\alpha-\beta\|^2+i\|\alpha+i\beta\|^2-i\|\alpha-i\beta\|^2$. **(Meerut 1987, 88, 91)**

(ii) $(\alpha,\beta)=Re(\alpha,\beta)+i\,Re(\alpha,i\beta)$. **(Meerut 1985, 88)**

Solution. (i) We have

$$\|\alpha+\beta\|^2=\|\alpha\|^2+(\alpha,\beta)+(\beta,\alpha)+\|\beta\|^2 \qquad \text{[See Ex. 7]}$$

Also $\|\alpha-\beta\|^2=\|\alpha\|^2-(\alpha,\beta)-(\beta,\alpha)+\|\beta\|^2$ [See Ex. 7]

$$\therefore \quad \|\alpha+\beta\|^2-\|\alpha-\beta\|^2=2(\alpha,\beta)+2(\beta,\alpha) \qquad \ldots(1)$$

Also $\|\alpha+i\beta\|^2=(\alpha+i\beta,\alpha+i\beta)=(\alpha,\alpha+i\beta)+i(\beta,\alpha+i\beta)$
$$=(\alpha,\alpha)+\bar{i}(\alpha,\beta)+i\{(\beta,\alpha)+\bar{i}(\beta,\beta)\}$$
$$=(\alpha,\alpha)-i(\alpha,\beta)+i\{(\beta,\alpha)-i(\beta,\beta)\}$$
$$[\because \ \bar{i}=-i]$$
$$=\|\alpha\|^2-i(\alpha,\beta)+i(\beta,\alpha)+\|\beta\|^2$$
$$\therefore \quad i\|\alpha+i\beta\|^2=i\|\alpha\|^2+(\alpha,\beta)-(\beta,\alpha)+i\|\beta\|^2. \qquad \ldots(2)$$

Also $\|\alpha-i\beta\|^2=(\alpha-i\beta,\alpha-i\beta)$
$$=(\alpha,\alpha-i\beta)-i(\beta,\alpha-i\beta)$$
$$=(\alpha,\alpha)-\bar{i}(\alpha,\beta)-i\{(\beta,\alpha)-\bar{i}(\beta,\beta)\}$$
$$=\|\alpha\|^2+i(\alpha,\beta)-i\{(\beta,\alpha)+i(\beta,\beta)\}$$
$$=\|\alpha\|^2+i(\alpha,\beta)-i(\beta,\alpha)+\|\beta\|^2.$$
$$\therefore \quad -i\|\alpha-i\beta\|^2=-i\|\alpha\|^2+(\alpha,\beta)-(\beta,\alpha)-i\|\beta\|^2. \qquad \ldots(3)$$

Adding (2) and (3), we get
$$i\|\alpha+i\beta\|^2-i\|\alpha-i\beta\|^2=2(\alpha,\beta)-2(\beta,\alpha). \qquad \ldots(4)$$

Adding (1) and (4), we get
$$4(\alpha,\beta)=\|\alpha+\beta\|^2-\|\alpha-\beta\|^2+i\|\alpha+i\beta\|^2-i\|\alpha-i\beta\|^2.$$

(ii) We have $(\alpha,\beta)=\text{Re}(\alpha,\beta)+i\text{ Im}(\alpha,\beta)$

If $z=x+iy$, then
$$\therefore \quad y=\text{Im } z=\text{Re}\{-i(x+iy)\}=\text{Re}(-iz)$$
$$\therefore \quad \text{Im}(\alpha,\beta)=\text{Re}\{-i(\alpha,\beta)\}$$
$$=\text{Re}(\alpha,i\beta) \qquad [\because \ (\alpha,i\beta)=-i(\alpha,\beta)]$$
$$\therefore \quad (\alpha,\beta)=\text{Re}(\alpha,\beta)+i\text{ Re}(\alpha,i\beta).$$

Example 9. *Suppose that α and β are vectors in an inner product space V. If $|(\alpha,\beta)|=\|\alpha\|\,\|\beta\|$ (that is, if the Schwarz inequality reduces to an equality), then α and β are linearly dependent.* **(Meerut 1969)**

Scanned with CamScanner

It is given that $|(\alpha, \beta)|=\|\alpha\|\cdot\|\beta\|$. ...(1)

Solution. If $\alpha=0$, then (1) is satisfied. Therefore if α and β satisfy (1), then α can be $\mathbf{0}$ also. If $\alpha=\mathbf{0}$, the vectors α and β are linearly dependent because any set of vectors containing zero vector is linearly dependent.

Let us now suppose that α and β satisfy (1) and $\alpha\neq\mathbf{0}$.

If $\alpha\neq 0$, then $\|\alpha\|>0$. Consider the vector

$$\gamma=\beta-\frac{(\beta, \alpha)}{\|\alpha\|^2}\alpha.$$

We have $(\gamma, \gamma)=\left(\beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha, \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)$

$$=\left(\beta, \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)-\frac{(\beta,\alpha)}{\|\alpha\|^2}\left(\alpha, \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha\right)$$

$$=(\beta,\beta)-\frac{\overline{(\beta,\alpha)}}{\|\alpha\|^2}(\beta,\alpha)-\frac{(\beta,\alpha)}{\|\alpha\|^2}(\alpha,\beta)+\frac{(\beta,\alpha)\overline{(\beta,\alpha)}}{\|\alpha\|^2\|\alpha\|^2}(\alpha,\alpha)$$

$$=\|\beta\|^2-\frac{\overline{(\beta,\alpha)}(\beta,\alpha)}{\|\alpha\|^2}-\frac{\overline{(\alpha,\beta)}(\alpha,\beta)}{\|\alpha\|^2}+\frac{(\beta,\alpha)\overline{(\beta,\alpha)}}{\|\alpha\|^2\|\alpha\|^2}\|\alpha\|^2$$

$$=\|\beta\|^2-\frac{|(\alpha,\beta)|^2}{\|\alpha\|^2}=\frac{\|\beta\|^2\|\alpha\|^2-|(\alpha,\beta)|^2}{\|\alpha\|^2}$$

$=0$. [from (1)]

Now $(\gamma,\gamma)=0 \Rightarrow \gamma=\mathbf{0}$

$$\Rightarrow \beta-\frac{(\beta,\alpha)}{\|\alpha\|^2}\alpha=\mathbf{0}$$

$\Rightarrow \alpha$ and β are linearly dependent.

Example 10. *If in an inner product space the vectors α and β are linearly dependent, then*

$$|(\alpha,\beta)|=\|\alpha\|\cdot\|\beta\|.$$

Solution. If $\alpha=0$, then $|(\mathbf{0},\beta)|=0$ and $\|\alpha\|=0$. Therefore the given result is true.

Also if $\beta=\mathbf{0}$, then $(\alpha,\mathbf{0})=\overline{(\mathbf{0},\alpha)}=\overline{0}=0$ and $\|\beta\|=0$.

So let us suppose that both α and β are non-zero vectors. Since they are linearly dependent, therefore $\alpha=c\beta$ where c is some scalar. We have

$$(\alpha,\beta)=(c\beta,\beta)=c(\beta,\beta)=c\|\beta\|^2.$$

$\therefore\quad |(\alpha,\beta)|=|c|\,\|\beta\|^2.$

Also $\quad \|\alpha\|=\|c\beta\|=|c|.\|\beta\|.$

$\therefore\quad \|\alpha\|\,\|\beta\|=|c|.\|\beta\|^2.$

Hence $\quad |(\alpha,\beta)|=\|\alpha\|\,\|\beta\|.$

Example 11. *If in an inner product space*

$$\|\alpha+\beta\|=\|\alpha\|+\|\beta\|,$$

then prove that the vectors α and β are linearly dependent. Give an example to show that the converse of this statement is false.

Solution. We have

$$\|\alpha+\beta\|=\|\alpha\|+\|\beta\|$$

$\Rightarrow \|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2+2\|\alpha\|.\|\beta\|$

$\Rightarrow \|\alpha\|^2+2\text{Re}(\alpha,\beta)+\|\beta\|^2=\|\alpha\|^2+\|\beta\|^2+2\|\alpha\|.\|\beta\|$ [See Ex. 7]

$\Rightarrow \text{Re}(\alpha,\beta)=\|\alpha\|.\|\beta\|$

$\Rightarrow \|\alpha\|.\|\beta\|\leqslant|(\alpha,\beta)|$ ...(1) [$\because$ Re $z\leqslant|z|$]

But by Schwarz inequality, we have

$$\|\alpha\|.\|\beta\|\geqslant|(\alpha,\beta)|. \qquad ...(2)$$

$\therefore$ from (1) and (2), we get

$$\|\alpha\|.\|\beta\|=|(\alpha,\beta)|.$$

Thus for the vectors α and β, the Schwarz inequality reduces to an equality. Hence the vectors α and β are linearly dependent.

Note that the proof we have given is applicable whether the inner product is complex or real. In real case,

$$\text{Re}(\alpha,\beta)=(\alpha,\beta).$$

The converse is not true as is obvious from the following example.

Take the inner product space $V_3(\mathbf{R})$ with standard inner product defined on it.

Let $\alpha=(-1, 0, 1)$, $\beta=(2, 0, -2)$ be two vectors in $V_3(\mathbf{R})$. Then $\beta=-2\alpha$, therefore α and β are linearly dependent. We have

$$\|\alpha\|=\sqrt{(1+0+1)}=\sqrt{2},\ \|\beta\|=\sqrt{8}.$$

$\therefore\quad \|\alpha\|+\|\beta\|=\sqrt{2}+\sqrt{8}.$

Also $\quad \alpha+\beta=(1, 0, -1).$

$\therefore\quad \|\alpha+\beta\|=\sqrt{2}.$

Obviously $\|\alpha+\beta\|\neq\|\alpha\|+\|\beta\|$.

Exercises

1. If $\alpha=(a_1, a_2, ..., a_n)$, $\beta=(b_1, b_2, ..., b_n)\in V_n(\mathbf{R})$, then prove that
$$(\alpha,\beta)=a_1b_1+a_2b_2+...+a_nb_n$$
defines an inner product on $V_n(\mathbf{R})$.
2. Which of the following define inner products in $V_2(\mathbf{R})$? Give reasons. [Assume $\alpha=(x_1, x_2)$, $\beta=(y_1, y_2)$].

Scanned with CamScanner

$(\alpha, \beta)=x_1y_1+2x_1y_2+2x_2y_1+5x_2y_2$.

(a) $(\alpha, \beta)=x_1^2-2x_1y_2-2x_2y_1+y_1^2$.

(b) $(\alpha, \beta)=2x_1y_1+5x_2y_2$.

(c) $(\alpha, \beta)=x_1y_1-2x_1y_2-2x_2y_1+4x_2y_2$.

(d) Ans. (a) and (c) are inner products, (b) and (d) are not.

3. Show that for the vectors $\alpha=(x_1, x_2)$ and $\beta=(y_1, y_2)$ from $\mathbf{R}^2$ the following defines an inner product on $\mathbf{R}^2$: **(Meerut 1977)**

$$(\alpha, \beta)=x_1y_1-x_2y_1-x_1y_2+2x_2y_2.$$

4. Let $\alpha=(a_1, a_2)$ and $\beta=(b_1, b_2)$ be any two vectors $\in V_2(\mathbf{C})$. Prove that $(\alpha, \beta)=a_1\bar{b}_1+(a_1+a_2)(\bar{b}_1+\bar{b}_2)$ defines an inner product in $V_2(\mathbf{C})$. Show that the norm of the vector (3, 4) in this inner product space is $\sqrt{(58)}$.

5. If α, β be vectors in a real inner product space such that

$$\| \alpha \|=\| \beta \|,$$

then prove that $(\alpha+\beta, \alpha-\beta)=0$.

6. Show that any two vectors α, β of an inner product space are linearly dependent if and only if $|(\alpha, \beta)|=\| \alpha \| . \| \beta \|$.

7. Let V be a finite-dimensional vector space and let $B=\{\alpha_1, ..., \alpha_n\}$ be a basis for V. Let (|) or (,) be an inner product on V. If $c_1, ..., c_n$ are any n scalars, show that there is exactly one vector α in V such that $(\alpha \mid \alpha_j)$ or $(\alpha, \alpha_j)=c_j, j=1, 2, ..., n$.

§ 3. **Orthogonality. Definition.** *Let α and β be vectors in an inner product space V. Then α is said to be* **orthogonal** *to β if*

$$(\alpha, \beta)=0.$$

The relation of orthogonality in an inner product space is **symmetric.** We have

α is orthogonal to $\beta \Rightarrow (\alpha, \beta)=0 \Rightarrow \overline{(\alpha, \beta)}=\bar{0}$

$\Rightarrow (\beta, \alpha)=0 \Rightarrow \beta$ is orthogonal to α.

So we can say that *two vectors α and β in an inner product space are* **orthogonal** *if* $(\alpha, \beta)=0$.

Note 1. *If α is orthogonal to β, then every scalar multiple of α is orthogonal to β.* Let k be any scalar. Then

$$(k\alpha, \beta)=k(\alpha, \beta)=k0=0 \qquad [\because (\alpha, \beta)=0]$$

Therefore $k\alpha$ is orthogonal to β.

Note 2. *The zero vector is orthogonal to every vector.* For every vector α in V, we have $(0, \alpha)=0$.

Note 3. *The zero vector is the only vector which is orthogonal to itself.*

We have

α is orthogonal to $\alpha \Rightarrow (\alpha, \alpha)=0$

$\Rightarrow \alpha=\mathbf{0}$, by def. of an inner product space.

Definition. *A vector α is said to be orthogonal to a set S if it is orthogonal to each vector in S. Similarly two subspaces are called orthogonal if every vector in each is orthogonal to every vector in the other.*

Orthogonal set. Definition. *Let S be a set of vectors in an inner product space V. Then S is said to be an* **orthogonal set** *provided that any two distinct vectors in S are orthogonal.*

Theorem 1. *Let $S=\{\alpha_1, ..., \alpha_m\}$ be an orthogonal set of non-zero vectors in an inner product space V. If a vector β in V is in the linear span of S, then*

$$\beta=\sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\| \alpha_k \|^2} \alpha_k.$$

Proof. Since $\beta \in L(S)$, therefore β can be expressed as a linear combination of the vectors in S. Let

$$\beta=c_1\alpha_1+...+c_m\alpha_m=\sum_{j=1}^{m} c_j\alpha_j.$$

We have for each k where $1 \leqslant k \leqslant m$,

$$(\beta, \alpha_k)=\left(\sum_{j=1}^{m} c_j\alpha_j, \alpha_k\right)$$

$$=\sum_{j=1}^{m} c_j(\alpha_j, \alpha_k) \qquad \text{[by linearity property of inner product]}$$

$$=c_k(\alpha_k, \alpha_k)$$

[On summing with respect to j. Note that S is an orthogonal set of non-zero vectors and so $(\alpha_j, \alpha_k)=0$ if $j \neq k$]

Now $\alpha_k \neq \mathbf{0}$. Therefore $(\alpha_k, \alpha_k) \neq 0$. Thus

$$\| \alpha_k \|^2 \neq 0.$$

$$\therefore \quad c_k=\frac{(\beta, \alpha_k)}{\| \alpha_k \|^2}, \; 1 \leqslant k \leqslant m.$$

Putting these values of $c_1, ..., c_m$ in (1), we get

$$\beta=\sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\| \alpha_k \|^2} \alpha_k.$$

Theorem 2. *Any orthogonal set of non-zero vectors in an inner product space V is linearly independent.* **[Meerut 1984(P)]**

Proof. Let S be an orthogonal set of non-zero vectors in an inner product space V. Let $S_1=\{\alpha_1, ..., \alpha_m\}$ be a finite subset of S containing m distinct vectors. Let

Scanned with CamScanner

$$\sum_{j=1}^{m} c_j\alpha_j = c_1\alpha_1 + \dots + c_m\alpha_m = 0. \quad \dots(1)$$

We have, for each k where $1 \leqslant k \leqslant m$,

$$\left(\sum_{j=1}^{m} c_j\alpha_j, \alpha_k\right) = \sum_{j=1}^{m} c_j (\alpha_j, \alpha_k)$$
$$= c_k (\alpha_k, \alpha_k) \quad [\because (\alpha_j, \alpha_k) = 0 \text{ if } j \neq k]$$
$$= c_k \| \alpha_k \|^2.$$

But from (1), $\sum_{j=1}^{m} c_j\alpha_j = 0$. Therefore $\left(\sum_{j=1}^{m} c_j\alpha_j, \alpha_k\right) = (0, \alpha) = 0$.

$\therefore$ (1) implies that $c_k \| \alpha_k \|^2 = 0$, $1 \leqslant k \leqslant m$

$\Rightarrow c_k = 0 \quad [\because \alpha_k \neq 0 \Rightarrow \| \alpha_k \|^2 \neq 0]$

$\therefore$ the set S_1 is linearly independent. Thus every finite subset of S is linearly independent. Therefore S is linearly independent.

Orthonormal set. Definition. *Let S be a set of vectors in an inner product space V. Then S is said to be an* **orthonormal set** *if*

(i) $\alpha \in S \Rightarrow \| \alpha \| = 1$ *i.e.* $(\alpha, \alpha) = 1$,

and (ii) $\alpha, \beta \in S$ *and* $\alpha \neq \beta \Rightarrow (\alpha, \beta) = 0$. **(Nagarjuna 1980)**

Thus an orthonormal set is an orthogonal set with the additional property that each vector in it is of length 1. *In other words a set S consisting of mutually orthogonal unit vectors is called an orthonormal set.* Obviously an orthonormal set cannot contain zero vector because $\| 0 \| = 0$.

A finite set $S = \{\alpha_1, \dots, \alpha_m\}$ is orthonormal if

$(\alpha_i, \alpha_j) = \delta_{ij}$ where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. **(Andhra 1981)**

Existence of an orthonormal set. *Every inner product space V which is not equal to zero space possesses an orthonormal set.*

Let $0 \neq \alpha \in V$. Then $\| \alpha \| \neq 0$. The set $\left\{\frac{\alpha}{\| \alpha \|}\right\}$ containing only one vector is necessarily an orthonormal set.

We have $\left(\frac{\alpha}{\| \alpha \|}, \frac{\alpha}{\| \alpha \|}\right) = \frac{1}{\| \alpha \|}\left(\alpha, \frac{\alpha}{\| \alpha \|}\right) = \frac{1}{\| \alpha \|} \cdot \frac{1}{\| \alpha \|}(\alpha, \alpha)$

$$= \frac{1}{\| \alpha \|^2} \| \alpha \|^2 = 1.$$

Theorem 2. *Let $S = \{\alpha_1, \dots, \alpha_m\}$ be an orthonormal set of vectors in an inner product space V. If a vector β is in the linear span of S, then $\beta = \sum_{k=1}^{m} (\beta, \alpha_k)\alpha_k$.*

Proof. Since $\beta \in L(S)$, therefore β can be expressed as a linear combination of the vectors in S. Let

$$\beta = c_1\alpha_1 + \dots + c_m\alpha_m = \sum_{j=1}^{m} c_j\alpha_j. \quad \dots(1)$$

We have for each k where $1 \leqslant k \leqslant m$,

$$(\beta, \alpha_k) = \left(\sum_{j=1}^{m} c_j\alpha_j, \alpha_k\right)$$
$$= \sum_{j=1}^{m} c_j (\alpha_j, \alpha_k) \quad \text{[by linearity of inner product]}$$
$$= \sum_{j=1}^{m} c_j \delta_{jk} \quad [\because S \text{ is an orthonormal set}]$$
$$= c_k$$

[On summing with respect to j and remembering that $\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ if $j \neq k$]

Putting the values of $c_1, \dots, c_m$ in (1), we get

$$\beta = \sum_{k=1}^{m} (\beta, \alpha_k)\alpha_k.$$

Theorem 4. *If $S = \{\alpha_1, \dots, \alpha_m\}$ is an orthonormal set in V and if $\beta \in V$ then $\gamma = \beta - \sum_{i=1}^{m} (\beta, \alpha_i)\alpha_i$ is orthogonal to each of $\alpha_1, \dots, \alpha_m$ and, consequently, to the subspace spanned by S.*

Proof. We have for each k where $1 \leqslant k \leqslant m$,

$$(\gamma, \alpha_k) = \left(\beta - \sum_{i=1}^{m} (\beta, \alpha_i)\alpha_i, \alpha_k\right)$$
$$= (\beta, \alpha_k) - \left(\sum_{i=1}^{m} (\beta, \alpha_i)\alpha_i, \alpha_k\right) \quad \text{[by linearity of inner product]}$$
$$= (\beta, \alpha_k) - \sum_{i=1}^{m} (\beta, \alpha_i)(\alpha_i, \alpha_k) \quad \text{[by linearity of inner product]}$$
$$= (\beta, \alpha_k) - \sum_{i=1}^{m} (\beta, \alpha_i)\delta_{ik} \quad [\because \alpha_i, \alpha_k \text{ belong to an orthonormal set}]$$
$$= (\beta, \alpha_k) - (\beta, \alpha_k) \quad [\because \delta_{ik} = 1 \text{ if } i = k \text{ and } \delta_{ik} = 0 \text{ if } i \neq k]$$
$$= 0.$$

Scanned with CamScanner

Hence the first part of the theorem.

Now let δ be any vector in the subspace spanned by S *i.e.*, let $\delta \in L(S)$. Then $\delta = \sum_{i=1}^{m} a_i \alpha_i$ where each a_i is some scalar.

We have $(\gamma, \delta) = \left(\gamma, \sum_{i=1}^{m} a_i \alpha_i\right) = \sum_{i=1}^{m} \bar{a}_i (\gamma, \alpha_i) = \sum_{i=1}^{m} \bar{a}_i \, 0 = 0.$

Thus γ is orthogonal to every vector δ in $L(S)$. Therefore γ is orthogonal to $L(S)$.

Theorem 5. *Any orthonormal set of vectors in an inner product space is linearly independent.*

Proof. Let S be any orthonormal set of vectors in an inner product space V. Let $S_1 = \{\alpha_1, \ldots, \alpha_m\}$ be a finite subset of S containing m distinct vectors. Let

$$\sum_{j=1}^{m} c_j \alpha_j = c_1 \alpha_1 + \ldots + c_m \alpha_m = \mathbf{0}. \qquad \ldots(1)$$

We have, for each k where $1 \leqslant k \leqslant m$,

$$\left(\sum_{j=1}^{m} c_j \alpha_j, \alpha_k\right) = \sum_{j=1}^{m} c_j (\alpha_j, \alpha_k) \quad \text{[by linearity of inner product]}$$

$$= \sum_{j=1}^{m} c_j \delta_{jk} \quad [\because \; (\alpha_j, \alpha_k) = \delta_{jk}]$$

$$= c_k. \quad \text{[On summing with respect to } j]$$

But from (1), $\sum_{j=1}^{m} c_j \alpha_j = \mathbf{0}$. Therefore $\left(\sum_{j=1}^{m} c_j \alpha_j, \alpha_k\right) = (\mathbf{0}, \alpha_k) = 0.$

$\therefore$ (1) implies that $c_k = 0$ for each $1 \leqslant k \leqslant m$.

$\therefore$ the set S_1 is linearly independent. Thus every finite subset of S is linearly independent. Therefore S is linearly independent.

Complete orthonormal set. Definition. *An orthonormal set is said to be complete if it is not contained in any larger orthonormal set.*

Orthonormal dimension of a finite-dimensional vector space. Definition. *Let V be a finite-dimensional inner product space of dimension n. If S is any orthonormal set in V then S is linearly independent. Therefore S cannot contain more than n distinct vectors because in an n-dimensional vector space a linearly independent set cannot contain more than n vectors.*

The **orthogonal dimension** *of V is defined as the largest number of vectors an orthonormal set in V can contain.*

Obviously the orthogonal dimension of V will be $\leqslant n$ where n is the linear dimension of V.

The following theorem gives us a characterization of completeness *i.e.* it gives us equivalent definitions of completeness.

Theorem 6. *If $S = \{\alpha_1, \ldots, \alpha_n\}$ is any finite orthonormal set in an inner product space V, then the following six conditions on S are equivalent:*

(*i*) *The orthonormal set S is complete.*

(*ii*) *If $(\beta, \alpha_i) = 0$ for $i = 1, \ldots, n$, then $\beta = \mathbf{0}$.*

(*iii*) *The linear span of S is equal to V i.e. $L(S) = V$.*

(*iv*) *If $\beta \in V$, then $\beta = \sum_{i=1}^{n} (\beta, \alpha_i) \alpha_i$.*

(*v*) *If β and γ are in V, then $(\beta, \gamma) = \sum_{i=1}^{n} (\beta, \alpha_i)(\alpha_i, \gamma)$.*

(*vi*) *If β is in V, then $\sum_{i=1}^{n} |(\beta, \alpha_i)|^2 = \|\beta\|^2$.*

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

It is given that S is a complete orthonormal set. Let $\beta \in V$ and $(\beta, \alpha_i) = 0$ for each $i = 1, \ldots, n$.

Then β is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_n$.

If $\beta \neq \mathbf{0}$, then adjoining the vector $\frac{\beta}{\|\beta\|}$ to the set S we obtain an orthonormal set larger than S. This contradicts the given statement that S is a complete orthonormal set. Hence $\beta = \mathbf{0}$.

(ii) $\Rightarrow$ (iii).

It is given that if $(\beta, \alpha_i) = 0$ for $i = 1, \ldots, n$, then $\beta = \mathbf{0}$. To prove that $L(S) = V$.

Let γ be any vector in V. Consider the vector $\delta = \gamma - \sum_{i=1}^{n} (\gamma, \alpha_i) \alpha_i$.

We know that δ is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_n$ i.e. $(\delta, \alpha_i) = 0$ for each $i = 1, \ldots, n$. Therefore according to the given statement $\delta = \mathbf{0}$. This gives $\gamma = \sum_{i=1}^{n} (\gamma, \alpha_i) \alpha_i$. Thus every vector

Scanned with CamScanner

γ in V can be expressed as a linear combination of $\alpha_1, \ldots\ldots, \alpha_n$. Therefore $L(S)=V$.

(iii) $\Rightarrow$ (iv).

It is given that $L(S)=V$. Therefore if $\beta \in V$, then β can be expressed as a linear combination of $\alpha_1, \ldots, \alpha_n$. From theorem 3, we know that this expression for β will be

$$\beta=\sum_{i=1}^{n}(\beta, \alpha_i)\, \alpha_i.$$

(iv) $\Rightarrow$ (v).

It is given that if β is in V, then $\beta=\sum_{i=1}^{n}(\beta, \alpha_i)\, \alpha_i$. If γ is another vector in V, then $\gamma=\sum_{i=1}^{n}(\gamma, \alpha_i)\, \alpha_i$.

We have $(\beta, \gamma)=\left(\sum_{i=1}^{n}(\beta, \alpha_i)\, \alpha_i, \sum_{j=1}^{n}(\gamma, \alpha_j)\, \alpha_j\right)$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}(\beta, \alpha_i)\, \overline{(\gamma, \alpha_j)}\, (\alpha_i, \alpha_j)=\sum_{i=1}^{n}(\beta, \alpha_i)\, \overline{(\gamma, \alpha_i)}$$

[On summing with respect to j]

$$=\sum_{i=1}^{n}(\beta, \alpha_i)\, (\alpha_i, \gamma).$$

(v) $\Rightarrow$ (vi).

It is given that if β and γ are in V, then $(\beta, \gamma)=\sum_{i=1}^{n}(\beta, \alpha_i)\, (\alpha_i, \gamma)$. If β is in V then taking $\gamma=\beta$ in the given result, we get

$$(\beta, \beta)=\sum_{i=1}^{n}(\beta, \alpha_i)\, (\alpha_i, \beta)=\sum_{i=1}^{n}(\beta, \alpha_i)\, \overline{(\beta, \alpha_i)}$$

$$\Rightarrow \|\beta\|^2=\sum_{i=1}^{n}|(\beta, \alpha_i)|^2.$$

(vi) $\Rightarrow$ (i).

It is given that if β is in V, then $\|\beta\|^2=\sum_{i=1}^{n}|(\beta, \alpha_i)|^2$. To prove that S is a complete orthonormal set.

Let S be not a complete orthonormal set *i.e.* let S be contained in a larger orthonormal set S_1.

Then there exists a vector α_0 in S_1 such that $\|\alpha_0\|=1$ and α_0 is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_n$. Since α_0 is in V, therefore from the given condition, we have

$$\|\alpha_0\|^2=\sum_{i=1}^{n}|(\alpha_0, \alpha_i)|^2=0.$$

This contradicts the fact that $\|\alpha_0\|=1$. Hence S must be complete.

Corollary. *Every complete orthonormal set in a finite-dimensional inner product space V forms a basis for V.*

Proof. Let S be a complete orthonormal set in a finite-dimensional inner product space V. Then S is linearly independent. Also by the above theorem $L(S)=V$. Hence S must be a basis for V.

In the next theorem we shall prove that the orthogonal dimension of a finite dimensional inner product space is equal to its linear dimension.

Theorem 7. *If V is an n-dimensional inner product space, then there exist complete orthonormal sets in V, and every complete orthonormal set in V contains exactly n elements. The orthogonal dimension of V is the same as its linear dimension.*

(Meerut 1969, 70, 73, 76, 77, 80)

Proof. Let $0 \neq \alpha \in V$. Then $\left\{\frac{\alpha}{\|\alpha\|}\right\}$ is an orthonormal set in V. If it is not complete, then we can enlarge it by adding one more vector to it so that the resulting set is also an orthonormal set. If this resulting orthonormal set is still not complete, then we enlarge it again. Thus we proceed by induction. Ultimately we must reach a complete orthonormal set because an orthonormal set is linearly independent and so it can contain at most n elements. Thus there exist complete orthonormal sets in V.

Now suppose $S=\{\alpha_1, \ldots, \alpha_m\}$ is a complete orthonormal set in V. Then S is linearly independent. Also the linear span of S is V. Therefore S is a basis for V. Hence the number of vectors in S must be n. Thus we must have $m=n$.

Thus we have proved that there exist complete orthonormal sets in V and each of them will have n elements. Thus n is the

Scanned with CamScanner

largest number of vectors that an orthonormal set in V will contain. Therefore the orthogonal dimension of V is equal to n which is also the linear dimension of V.

Now we shall give an alternative proof of theorem 7. This proof will be a constructive proof *i.e.*, it will also give us a process to construct an orthonormal basis for a finite dimensional inner product space.

Orthonormal basis. Definition. *A basis of an inner product space that consists of mutually orthogonal unit vectors is called an orthonormal basis.*

Gram-Schmidt orthogonalization process. Theorem 8. *Every finite-dimensional inner product space has an orthonormal basis.*

(Meerut 1972, 73, 74, 79, 80, 82, 83P, 92, 93P; Madras 83; Andhra 92; Tirupati 90; Nagarjuna 77)

Proof. Let V be an n-dimensional inner product space and let $B=\{\beta_1, \beta_2, \ldots, \beta_n\}$ be a basis for V. From this set we shall construct an orthonormal set $B_1=\{\alpha_1, \ldots, \alpha_n\}$ of n distinct vectors by means of a construction known as Gram-Schmidt orthogonalization process. The main idea behind this construction is that each α_j, $1 \leqslant j \leqslant n$ will be in the linear span of $\beta_1, \ldots, \beta_j$.

We have $\beta_1 \neq 0$ because the set B is linearly independent. Let $\alpha_1=\dfrac{\beta_1}{\| \beta_1 \|}$. We have $(\alpha_1, \alpha_1)=\left(\dfrac{\beta_1}{\| \beta_1 \|}, \dfrac{\beta_1}{\| \beta_1 \|}\right)=\dfrac{1}{\| \beta_1 \|^2}(\beta_1, \beta_1)$

$$=\frac{1}{\| \beta_1 \|^2} \| \beta_1 \|^2=1.$$

Thus we have constructed an orthonormal set $\{\alpha_1\}$ containing one vector. Also α_1 is in the linear span of β_1.

Now let $\gamma_2=\beta_2-(\beta_2, \alpha_1)\, \alpha_1$. By theorem 4, γ_2 is orthogonal to α_1. Also $\gamma_2 \neq 0$ because if $\gamma_2=0$, then β_2 is a scalar multiple of α_1 and therefore of β_1. But this is not possible because the vectors β_1 and β_2 are linearly independent. Hence $\gamma_2 \neq 0$. Let us now put $\alpha_2=\dfrac{\gamma_2}{\| \gamma_2 \|}$. Then $\| \alpha_2 \|=1$. Also α_2 is orthogonal to α_1 because α_2 is simply a scalar multiple of γ_2 which is orthogonal to α_1. Further $\alpha_2 \neq \alpha_1$. For otherwise β_2 will become a scalar multiple of β_1. Thus $\{\alpha_1, \alpha_2\}$ is an orthonormal set containing two distinct vectors such that α_1 is in the linear span of β_1 and α_2 is in the linear span of β_1, β_2.

The way ahead is now clear. Suppose that we have constructed an orthonormal set $\{\alpha_1, \ldots, \alpha_k\}$ of k (where $k < n$) distinct

vectors such that each α_j $(j=1, \ldots, k)$ is a linear combination of $\beta_1, \ldots, \beta_j$. Consider the vector

$$\gamma_{k+1}=\beta_{k+1}-(\beta_{k+1}, \alpha_1)\, \alpha_1-(\beta_{k+1}, \alpha_2)\, \alpha_2-\ldots-(\beta_{k+1}, \alpha_k)\, \alpha_k. \quad \ldots(1)$$

By theorem 4, γ_{k+1} is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_k$. Suppose $\gamma_{k+1}=\mathbf{0}$. Then β_{k+1} is a linear combination of $\alpha_1, \ldots, \alpha_k$. But according to our assumption each α_j $(j=1, \ldots, k)$ is a linear combination of $\beta_1, \beta_2, \ldots, \beta_j$. Therefore β_{k+1} is a linear combination of $\beta_1, \ldots, \beta_k$. This is not possible because $\beta_1, \ldots, \beta_k, \beta_{k+1}$ are linearly independent.

Therefore we must have $\gamma_{k+1} \neq \mathbf{0}$.

Let us now put $\alpha_{k+1}=\dfrac{\gamma_{k+1}}{\| \gamma_{k+1} \|}$. $\quad \ldots(2)$

We have $\| \alpha_{k+1} \|=1$. Also α_{k+1} is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_k$, because α_{k+1} is simply a scalar multiple of γ_{k+1} which is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_k$. Further obviously $\alpha_{k+1} \neq \alpha_j$, $j=1, \ldots, k$. For otherwise from (1) and (2), we see that β_{k+1} will become a linear combination of $\beta_1, \ldots, \beta_k$. Also from (1) and (2), we see that α_{k+1} is in the linear span of $\beta_1, \ldots, \beta_{k+1}$.

Thus we have been able to construct an orthonormal set

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

containing $k+1$ distinct vectors such that α_j $(j=1, 2, \ldots, k+1)$ is in the linear span of $\beta_1, \ldots, \beta_j$. Our aim is now complete by induction. Thus continuing in this way we shall ultimately obtain an orthonormal set $B_1=\{\alpha_1, \ldots, \alpha_n\}$ containing n distinct vectors. The set B_1 is linearly independent because it is an orthonormal set. Therefore B_1 is a basis for V because the number of vectors in B_1 is equal to the dimension of V. Also the set B_1 is a complete orthonormal set because the maximum number of vectors in an orthonormal set in V can be n. Thus there exist complete orthonormal sets in V. Also the orthogonal dimension of V is equal to n *i e.*, equal to the linear dimension of V.

Note. In the above construction the vector α_2 will be $\dfrac{\gamma_2}{\| \gamma_2 \|}$ where $\gamma_2=\beta_2-(\beta_2, \alpha_1)\, \alpha_1$. Similarly the vector α_3 will be $\dfrac{\gamma_3}{\| \gamma_3 \|}$ where $\gamma_3=\beta_3-(\beta_3, \alpha_1)\, \alpha_1-(\beta_3, \alpha_2)\, \alpha_2$. Similarly the other vectors can be found.

How to apply Gram-Schmidt orthogonalization process to numerical problems ?

Suppose $B=\{\beta_1, \beta_2, \ldots, \beta_n\}$ is a given basis of a finite dimen-

Scanned with CamScanner

sional inner product space V. Let
$$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$$
be an orthonormal basis for V which we are required to construct from the basis B. The vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$ will be obtained in the following way.

Take $\alpha_1 = \dfrac{\beta_1}{\| \beta_1 \|}$,

$$\alpha_2 = \frac{\gamma_2}{\| \gamma_2 \|} \text{ where } \gamma_2 = \beta_2 - (\beta_2, \alpha_1)\,\alpha_1,$$

$$\alpha_3 = \frac{\gamma_3}{\| \gamma_3 \|} \text{ where } \gamma_3 = \beta_3 - (\beta_3, \alpha_1)\,\alpha_1 - (\beta_3, \alpha_2)\,\alpha_2,$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\alpha_n = \frac{\gamma_n}{\| \gamma_n \|} \text{ where } \gamma_n = \beta_n - (\beta_n, \alpha_1)\,\alpha_1 - (\beta_n, \alpha_2)\,\alpha_2 - \ldots - (\beta_n, \alpha_{n-1})\,\alpha_{n-1}.$$

Now we shall give an example to illustrate the Gram-Schmidt process.

Example. *Apply the Gram-Schmidt process to the vectors* $\beta_1 = (1, 0, 1)$, $\beta_2 = (1, 0, -1)$, $\beta_3 = (0, 3, 4)$, *to obtain an orthonormal basis for* $V_3(\mathbf{R})$ *with the standard inner product.*

(Meerut 1980, 81, 83, 88, 93; Nagarjuna 80; S.V.U. Tirupati 90)

Solution. We have $\| \beta_1 \|^2 = (\beta_1, \beta_1) = 1.1 + 0.0 + 1.1$
$$= (1)^2 + (0)^2 + (1)^2 = 2.$$

Let $\alpha_1 = \dfrac{\beta_1}{\| \beta_1 \|} = \dfrac{1}{\sqrt{2}}(1, 0, 1) = \left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right)$.

Now let $\gamma_2 = \beta_2 - (\beta_2, \alpha_1)\,\alpha_1$.

We have $(\beta_2, \alpha_1) = 1.\dfrac{1}{\sqrt{2}} + 0.0 + (-1).\dfrac{1}{\sqrt{2}} = 0$.

$\therefore \quad \gamma_2 = (1, 0, -1) - 0\left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right) = (1, 0, -1)$.

Now $\| \gamma_2 \|^2 = (\gamma_2, \gamma_2) = (1)^2 + (0)^2 + (-1)^2 = 2$.

Let $\alpha_2 = \dfrac{\gamma_2}{\| \gamma_2 \|} = \dfrac{1}{\sqrt{2}}(1, 0, -1) = \left(\dfrac{1}{\sqrt{2}}, 0, -\dfrac{1}{\sqrt{2}}\right)$.

Now let $\gamma_3 = \beta_3 - (\beta_3, \alpha_1)\,\alpha_1 - (\beta_3, \alpha_2)\,\alpha_2$.

We have $(\beta_3, \alpha_1) = \left((0, 3, 4), \left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right)\right)$
$$= 0.\frac{1}{\sqrt{2}} + 3.0 + 4.\frac{1}{\sqrt{2}} = 2\sqrt{2}.$$

Also $(\beta_3, \alpha_2) = \left((0, 3, 4), \left(\dfrac{1}{\sqrt{2}}, 0, -\dfrac{1}{\sqrt{2}}\right)\right)$
$$= 0.\frac{1}{\sqrt{2}} + 3.0 - 4.\frac{1}{\sqrt{2}} = -2\sqrt{2}.$$

$$\therefore \quad \gamma_3 = (0, 3, 4) - 2\sqrt{2}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) + 2\sqrt{2}\left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right)$$
$$= (0, 3, 4) - (2, 0, 2) + (2, 0, -2) = (0, 3, 0).$$

Now $\| \gamma_3 \|^2 = (\gamma_3, \gamma_3) = (0)^2 + (3)^2 + (0)^2 = 9$.

Put $\alpha_3 = \dfrac{\gamma_3}{\| \gamma_3 \|} = \frac{1}{3}(0, 3, 0) = (0, 1, 0)$.

Now $\{\alpha_1, \alpha_2, \alpha_3\}$ *i.e.* $\left\{\left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}}\right), \left(\dfrac{1}{\sqrt{2}}, 0, -\dfrac{1}{\sqrt{2}}\right), (0, 1, 0)\right\}$

is the required orthonormal basis for $V_3(\mathbf{R})$.

Theorem 9. Bessel's inequality.

If $B = \{\alpha_1, \ldots, \alpha_m\}$ is any finite orthonormal set in an inner product space V and if β is any vector in V, then

$$\sum_{i=1}^{m} |(\beta, \alpha_i)|^2 \leqslant \| \beta \|^2.$$ **(Meerut 1978, 79, 83P, 89, 91; Nagarjuna 91)**

Furthermore, equality holds if and only if β is in the subspace spanned by $\alpha_1, \ldots, \alpha_m$.

Proof. Consider the vector $\gamma = \beta - \sum_{i=1}^{m} (\beta, \alpha_i)\,\alpha_i$.

We have $\| \gamma \|^2 = (\gamma, \gamma) = \left(\beta - \sum_{i=1}^{m} (\beta, \alpha_i)\,\alpha_i,\ \beta - \sum_{j=1}^{m} (\beta, \alpha_j)\,\alpha_j\right)$

$$= (\beta, \beta) - \sum_{i=1}^{m} (\beta, \alpha_i)(\alpha_i, \beta) - \sum_{j=1}^{m} \overline{(\beta, \alpha_j)}\,(\beta, \alpha_j)$$
$$+ \sum_{i=1}^{m} \sum_{j=1}^{m} (\beta, \alpha_i)\,\overline{(\beta, \alpha_j)}\,(\alpha_i, \alpha_j)$$

$$= (\beta, \beta) - \sum_{i=1}^{m} (\beta, \alpha_i)\,\overline{(\beta, \alpha_i)} - \sum_{j=1}^{m} \overline{(\beta, \alpha_j)}\,(\beta, \alpha_j) + \sum_{i=1}^{m} (\beta, \alpha_i)\,\overline{(\beta, \alpha_i)}$$

[On summing with respect to j and remembering that $(\alpha_i, \alpha_j) = 0$ when $j \neq i$]
$(\alpha_i, \alpha_j) = 1$ when $j = i$

$$= \| \beta \|^2 - \sum_{i=1}^{m} |(\beta, \alpha_i)|^2 - \sum_{i=1}^{m} |(\beta, \alpha_i)|^2 + \sum_{i=1}^{m} |(\beta, \alpha_i)|^2.$$

Scanned with CamScanner

$$\therefore \quad \| \gamma \|^2 = \| \beta \|^2 - \sum_{i=1}^{m} |(\beta, \alpha_i)|^2 \qquad \ldots(1)$$

Now $\qquad \| \gamma \|^2 \geqslant 0.$

$$\therefore \quad \| \beta \|^2 - \sum_{i=1}^{m} |(\beta, \alpha_i)|^2 \geqslant 0$$

or $$\sum_{i=1}^{m} |(\beta, \alpha_i)|^2 \leqslant \| \beta \|^2$$

If the equality holds *i.e.* if $\sum_{i=1}^{m} |(\beta, \alpha_i)|^2 = \| \beta \|^2$, then from (1) we have $\| \gamma \|^2 = 0$. This implies that $\gamma = 0$ *i.e.* $\beta = \sum_{i=1}^{m} (\beta, \alpha_i) \alpha_i$.

Thus if the equality holds, then β is a linear combination of $\alpha_1, \ldots, \alpha_m$.

If β is a linear combination of $\alpha_1, \ldots, \alpha_m$, then from theorem 3, we know that $\beta = \sum_{i=1}^{m} (\beta, \alpha_i) \alpha_i$. This implies that $\gamma = 0$ which in itself implies that $\| \gamma \|^2 = 0$. Then from (1), we get

$$\sum_{i=1}^{m} |(\beta, \alpha_i)|^2 = \| \beta \|^2$$

and thus the equality holds.

Note. Another statement of Bessel's inequality.

Let $\{\alpha_1, \ldots, \alpha_m\}$ be an orthogonal set of non-zero vectors in an inner product space V. If β is any vector in V, then

$$\sum_{i=1}^{m} \frac{|(\beta, \alpha_i)|^2}{\| \alpha_i \|^2} \leqslant \| \beta \|^2. \qquad \textbf{(Meerut 1971, 74)}$$

Proof. Let $B = \{\delta_1, \ldots, \delta_m\}$ where $\delta_i = \dfrac{\alpha_i}{\| \alpha_i \|}$, $1 \leqslant i \leqslant m$.

Then $\| \delta_i \| = 1$. Thus the set B is an orthonormal set. Now proceeding as in the previous theorem, we get

$$\sum_{i=1}^{m} |(\beta, \delta_i)|^2 \leqslant \| \beta \|^2. \qquad \ldots(1)$$

Also $(\beta, \delta_i) = \left(\beta, \dfrac{\alpha_i}{\| \alpha_i \|}\right) = \dfrac{1}{\| \alpha_i \|} (\beta, \alpha_i).$

$$\therefore \quad |(\beta, \delta_i)|^2 = \frac{|(\beta, \alpha_i)|^2}{\| \alpha_i \|^2}. \qquad \ldots(2)$$

From (1) and (2), we get the required result.

Corollary. *If V is finite dimensional and if $\{\alpha_1, \ldots, \alpha_m\}$ is an orthonormal set in V such that $\sum_{i=1}^{m} |(\beta, \alpha_i)|^2 = \| \beta \|^2$ for every $\beta \in V$, prove that $\{\alpha_1, \ldots, \alpha_m\}$ must be a basis of V.*

Proof. Let β be any vector in V. Consider the vector

$$\gamma = \beta - \sum_{i=1}^{m} (\beta, \alpha_i) \alpha_i. \qquad \ldots(1)$$

As in the proof of Bessel's inequality, we have

$$\| \gamma \|^2 = (\gamma, \gamma) = \| \beta \|^2 - \sum_{i=1}^{m} |(\beta, \alpha_i)|^2 \qquad \text{[prove it here]}$$

$$= 0 \text{ by the given condition.}$$

$$\therefore \quad \gamma = 0 \text{ i.e., } \beta = \sum_{i=1}^{m} (\beta, \alpha_i) \alpha_i. \qquad \text{[from (1)]}$$

Thus every vector β in V can be expressed as a linear combination of the vectors in the set $S = \{\alpha_1, \ldots, \alpha_m\}$ *i.e.* $L(S) = V$. Also S is linearly independent because it is an orthonormal set. Hence S must be a basis for V.

Orthogonal complement. Definition *Let V be an inner product space, and let S be any set of vectors in V. The orthogonal complement of S, written as $S^\perp$ and read as S perpendicular, is defined by*

$$S^\perp = \{\alpha \in V : (\alpha, \beta) = 0 \;\forall\; \beta \in S\}$$

Thus $S^\perp$ is the set of all those vectors in V which are orthogonal to every vector in S.

Theorem 10. *Let S be any set of vectors in an inner product space V. Then $S^\perp$ is a subspace of V.* **(Madras 1981; Andhra 81)**

Proof. We have, by definition

$$S^\perp = \{\alpha \in V : (\alpha, \beta) = 0 \;\forall\; \beta \in S\}.$$

Since $(0, \beta) = 0 \;\forall\; \beta \in S$, therefore at least $0 \in S^\perp$ and thus $S^\perp$ is not empty.

Let $a, b \in F$ and $\gamma, \delta \in S^\perp$. Then $(\gamma, \beta) = 0 \;\forall\; \beta \in S$ and $(\delta, \beta) = 0 \;\forall\; \beta \in S$.

Scanned with CamScanner

For every $\beta \in S$, we have

$(a\gamma+b\delta, \beta)=a(\gamma, \beta)+b(\delta, \beta)=a0+b0=0$.

Therefore $a\gamma+b\delta \in S^{\perp}$. Hence $S^{\perp}$ is a subspace of V

Note. The orthogonal complement of V is the zero subspace and the orthogonal complement of the zero subspace is V itself.

Orthogonal complement of an orthogonal complement.

Definition. *Let S be any subset of an inner product space V. Then $S^{\perp}$ is a subset of V. We define $(S^{\perp})^{\perp}$, written as $S^{\perp\perp}$, by* $S^{\perp\perp}=\{\alpha \in V : (\alpha, \beta)=0 \ \forall\ \beta \in S^{\perp}\}$.

Obviously $S^{\perp\perp}$ is a subspace of V. Also it can be easily seen that $S \subseteq S^{\perp\perp}$.

Let $\alpha \in S$. Then $(\alpha, \beta)=0 \ \forall\ \beta \in S^{\perp}$. Therefore by definition of $(S^{\perp})^{\perp}$, $\alpha \in (S^{\perp})^{\perp}$. Thus $\alpha \in S \Rightarrow \alpha \in S^{\perp\perp}$. Therefore $S \subseteq S^{\perp\perp}$.

The following theorem known as the **projection theorem** is very important.

Theorem 11. *Let W be any subspace of a finite dimensional inner product space V. Then*

(*i*) $V=W \oplus W^{\perp}$, *and* (*ii*) $W^{\perp\perp}=W$.

(Meerut 1974, 76, 84, 85, 90, 93 P; Andhra 92)

Proof (i) First we shall prove that $V=W+W^{\perp}$.

Since W is a subspace of a finite dimensional vector space V therefore W itself is also finite-dimensional. Let $\dim V=n$ and $\dim W=m$.

Now every finite-dimensional vector space possesses an orthonormal basis. Let $B_1=\{\alpha_1, \ldots, \alpha_m\}$ be an orthonormal basis for W.

Let β be any vector in V. Consider the vector

$$\gamma=\beta-\sum_{i=1}^{m}(\beta, \alpha_i)\,\alpha_i. \qquad \ldots(1)$$

By theorem 4, the vector γ is orthogonal to each of the vectors $\alpha_1, \ldots, \alpha_m$ and consequently γ is orthogonal to the subspace W spanned by these vectors. Thus γ is orthogonal to every vector in W. Therefore $\gamma \in W^{\perp}$. Also the vector $\sum_{i=1}^{m}(\beta, \alpha_i)\,\alpha_i$ is in W because it is a linear combination of vectors belonging to a basis for W.

Now from (1), we have

$\beta=\left[\sum_{i=1}^{m}(\beta, \alpha_i)\alpha_i\right]+\gamma$ where $\sum_{i=1}^{m}(\beta, \alpha_i)\,\alpha_i$ is in W and γ is in $W^{\perp}$. Therefore $V=W+W^{\perp}$.

Now we shall prove that the subspaces W and $W^{\perp}$ are disjoint. Let $\alpha \in W \cap W^{\perp}$.

Then $\alpha \in W$ and $\alpha \in W^{\perp}$. Since $\alpha \in W^{\perp}$, therefore α is orthogonal to every vector in W. In particular α is orthogonal to α because $\alpha \in W$. Now $(\alpha, \alpha)=0 \Rightarrow \alpha=0$. Thus 0 is the only vector which belongs to both W and $W^{\perp}$. Hence W and $W^{\perp}$ are disjoint

$$\therefore \quad V=W \oplus W^{\perp}.$$

(ii) We have $V=W \oplus W^{\perp}$. ...(2)

Now $W^{\perp}$ is also a subspace of V. Therefore taking $W^{\perp}$ in place of W and using the result (2), we get

$$V=W^{\perp} \oplus W^{\perp\perp}. \qquad \ldots(3)$$

Since V is the direct sum of W and $W^{\perp}$ and V is finite-dimensional, therefore

$$\dim V=\dim W+\dim W^{\perp}. \qquad \ldots(4)$$

Similarly from (3), we get

$$\dim V=\dim W^{\perp}+\dim W^{\perp\perp}. \qquad \ldots(5)$$

From (4) and (5), we get

$$\dim W=\dim W^{\perp\perp} \qquad \ldots(6)$$

Now we shall prove that $W \subseteq W^{\perp\perp}$.

Let $\alpha \in W$. Then $(\alpha, \beta)=0 \ \forall\ \beta \in W^{\perp}$ Therefore by definition of $(W^{\perp})^{\perp}$, $\alpha \in (W^{\perp})^{\perp}$. Thus $\alpha \in W \Rightarrow \alpha \in W^{\perp\perp}$. Therefore $W \subseteq W^{\perp\perp}$.

Since $W \subseteq W^{\perp\perp}$, therefore W is a subspace of $W^{\perp\perp}$. Also $\dim W=\dim W^{\perp\perp}$ Hence $W=W^{\perp\perp}$.

Corollary. *Let W be any subspace of a finite-dimensional inner product space V. Then*

$$\dim W^{\perp}=\dim V-\dim W$$

Proof. Since V is finite dimensional and

$$V=W \oplus W^{\perp},$$

therefore,

$$\dim V=\dim W+\dim W^{\perp}$$

$$\Rightarrow \dim W^{\perp}=\dim V-\dim W.$$

Definition. *If W is a subspace of a finite dimensional inner product space V, then $V=W \oplus W^{\perp}$. Therefore every vector α in V can be uniquely expressed as $\alpha=\alpha_1+\alpha_2$ where $\alpha_1 \in W$ and $\alpha_2 \in W^{\perp}$.*

Scanned with CamScanner

The vectors α_1 and α_2 are then called the **orthogonal projections** *of α on the subspaces W and $W^{\perp}$.*

Solved Examples

Example 1. *State whether the following statement is true or false. Give reasons to support your answer.*

α is an element of an n-dimensional unitary space V and α is perpendicular to n linearly independent vectors from V, then $\alpha=0$.

(Meerut 1977)

Solution. True. Suppose α is perpendicular to n linearly independent vectors $\alpha_1, ..., \alpha_n$.

Since V is of dimension n, therefore the n linearly independent vectors $\alpha_1, ..., \alpha_n$ constitute a basis for V. So we can write $\alpha=a_1\alpha_1+...+a_n\alpha_n$. Now

$$(\alpha, \alpha)=(a_1\alpha_1+...+a_n\alpha_n, \alpha)=a_1(\alpha_1, \alpha)+...+a_n(\alpha_n, \alpha)$$
$$=a_1\times 0+...+a_n\times 0 \quad [\because \alpha \text{ is } \perp \text{ to each of the vectors } \alpha_1, ..., \alpha_n]$$
$$=0.$$

$\therefore \alpha=0.$

Example 2. *If α and β are orthogonal unit vectors (that is, $\{\alpha, \beta\}$ is an orthonormal set), what is the distance between α and β?*

Solution. If $d(\alpha, \beta)$ denotes the distance between α and β, then $d(\alpha, \beta)=\|\alpha-\beta\|$.

We have
$$\|\alpha-\beta\|^2=(\alpha-\beta, \alpha-\beta)=(\alpha, \alpha-\beta)-(\beta, \alpha-\beta)$$
$$=(\alpha, \alpha)-(\alpha, \beta)-(\beta, \alpha)+(\beta, \beta)$$
$$=\|\alpha\|^2-0-0+\|\beta\|^2 \quad [\because \alpha \text{ is orthogonal to } \beta]$$
$$=1+1 \quad [\because \alpha \text{ and } \beta \text{ are unit vectors}]$$
$$=2.$$

$\therefore d(\alpha, \beta)=\|\alpha-\beta\|=\sqrt{2}.$

Example 3. *Prove that two vectors α and β in a real inner product space are orthogonal if and only if $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$.*

(Nagarjuna 1991; Madurai 85)

Solution. Let α, β be two vectors in a real inner product space V. We have

$$\|\alpha+\beta\|^2=(\alpha+\beta, \alpha+\beta)$$
$$=(\alpha, \alpha)+(\alpha, \beta)+(\beta, \alpha)+(\beta, \beta)$$
$$=\|\alpha\|^2+2(\alpha, \beta)+\|\beta\|^2 \quad [\because (\beta, \alpha)=(\alpha, \beta)]$$

Thus in a real inner product space V, we have

$$\|\alpha+\beta\|^2=\|\alpha\|^2+2(\alpha, \beta)+\|\beta\|^2. \quad ...(1)$$

If α and β are orthogonal, $(\alpha, \beta)=0$.

Therefore from (1), we get $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$.

Conversely, suppose that $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$.

Then from (1), we get $2(\alpha, \beta)=0$ *i.e.*, $(\alpha, \beta)=0$.

Therefore α and β are orthogonal.

Note 1. The above result is known as the **Pythogorean theorem.** Its geometrical interpretation is that if ABC is a triangle in three dimensional Euclidean space, then the angle B is a right angle if and only if $AB^2+BC^2=AC^2$.

Note 2. If V is a complex inner product space, then the above result becomes false.

In this case

$$\|\alpha+\beta\|^2=\|\alpha\|^2+(\alpha, \beta)+\overline{(\alpha, \beta)}+\|\beta\|^2$$
$$=\|\alpha\|^2+2\,\text{Re}\,(\alpha, \beta)+\|\beta\|^2.$$

If α and β are orthogonal, then $(\alpha, \beta)=0$. So Re $(\alpha, \beta)=0$ and we get $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$.

But if $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$, then we get 2 Re $(\alpha, \beta)=0$. This implies that Re $(\alpha, \beta)=0$.

This does not necessarily imply that $(\alpha, \beta)=0$ *i.e.*, α and β are orthogonal. Thus in a complex inner product space if α and β are orthogonal, then we have $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$. But if we have $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$, then it is not necessary that α and β are orthogonal.

Example 4. *If α and β are vectors in a real inner product space, and if $\|\alpha\|=\|\beta\|$, then $\alpha-\beta$ and $\alpha+\beta$ are orthogonal. Interpret the result geometrically.*

Solution. Let α and β be vectors in a real inner product space V. Also let $\|\alpha\|=\|\beta\|$. We have

$$(\alpha-\beta, \alpha+\beta)=(\alpha, \alpha+\beta)-(\beta, \alpha+\beta)$$
$$=(\alpha, \alpha)+(\alpha, \beta)-(\beta, \alpha)-(\beta, \beta)$$
$$=\|\alpha\|^2+(\alpha, \beta)-(\alpha, \beta)-\|\beta\|^2$$
$$=0 \quad [\because \|\alpha\|^2=\|\beta\|^2]$$

$\therefore \alpha-\beta$ and $\alpha+\beta$ are orthogonal.

Geometrical Interpretation. Let V be the three dimensional Euclidean space *i.e.*, let V be the inner product space V_3 (R) with standard inner product defined on it. Let vectors α and β represent the sides AB and BC of a parallelogram $ABCD$. Since the length of α is equal to the length of β, therefore $ABCD$ is a rhombus. The vectors $\alpha+\beta$ and $\alpha-\beta$ are along the diagonals AC and DB of

Scanned with CamScanner

the rhombus. Therefore diagonals of a rhombus intersect at right angles.

Example 5. *If α and β are vectors in a real inner product space, and if $\alpha+\beta$ is orthogonal to $\alpha-\beta$, then prove that $\|\alpha\|=\|\beta\|$. Interpret the result geometrically.*

Solution. We have $\alpha+\beta$ is orthogonal to $\alpha-\beta$

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

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

$\Rightarrow \|\alpha\|^2+(\alpha, \beta)-(\alpha, \beta)-\|\beta\|^2=0$

$\Rightarrow \|\alpha\|^2=\|\beta\|^2 \Rightarrow \|\alpha\|=\|\beta\|$.

Geometrical Interpretation. Let V be the three dimensional Euclidean space i.e., let V be the inner product space $V_3(\mathbf{R})$ with standard inner product defined on it. Let vectors α and β represent the sides AB and BC of a parallelogram $ABCD$. Then the vectors $\alpha+\beta$ and $\alpha-\beta$ are along the diagonals AC and DB of the parallelogram. If these diagonals are at right angles, then the length of α is equal to the length of β. So $AB=BC$ and the parallelogram is a rhombus.

Example 6. *Two vectors α and β in a complex inner product space are orthogonal if and only if $\|a\alpha+b\beta\|^2=\|a\alpha\|^2+\|b\beta\|^2$ for all pairs of scalars a and b.* **(Meerut 1975)**

Solution. Let α and β be any two vectors in a complex inner product space V. Also let a, b be any two scalars. We have

$$\|a\alpha+b\beta\|^2=(a\alpha+b\beta, a\alpha+b\beta)$$
$$=(a\alpha, a\alpha+b\beta)+(b\beta, a\alpha+b\beta)$$
$$=(a\alpha, a\alpha)+(a\alpha, b\beta)+(b\beta, a\alpha)+(b\beta, b\beta)$$
$$=\|a\alpha\|^2+\|b\beta\|^2+a\bar{b}(\alpha, \beta)+b\bar{a}(\beta, \alpha)$$
$$=\|a\alpha\|^2+\|b\beta\|^2+a\bar{b}(\alpha, \beta)+b\bar{a}\,\overline{(\alpha, \beta)}. \quad \ldots(1)$$

If α and β are orthogonal, then $(\alpha, \beta)=0$.

Therefore from (1), we get $\|a\alpha+b\beta\|^2=\|a\alpha\|^2+\|b\beta\|^2$ for all pairs of scalars a and b.

Conversely, suppose that for all pairs of scalars a and b we have $\|a\alpha+b\beta\|^2=\|a\alpha\|^2+\|b\beta\|^2$. Then from (1), for all pairs of scalars a and b, we get

$$a\bar{b}(\alpha, \beta)+b\bar{a}\,\overline{(\alpha, \beta)}=0. \quad \ldots(2)$$

Take $a=1, b=1$. Then (2) gives

$$(\alpha, \beta)+\overline{(\alpha, \beta)}=0$$

$\Rightarrow 2\text{Re}\,(\alpha, \beta)=0$

$\Rightarrow \text{Re}\,(\alpha, \beta)=0.$

Again take $a=i, b=1$. Then (2) gives

$$i(\alpha, \beta)-i\,\overline{(\alpha, \beta)}=0$$

$\Rightarrow (\alpha, \beta)-\overline{(\alpha, \beta)}=0$ $\quad [\because i\neq 0]$

$\Rightarrow 2i\,\text{Im}\,(\alpha, \beta)=0$ $\quad [\because z-\bar{z}=2i\,\text{Im}\,z]$

$\Rightarrow \text{Im}\,(\alpha, \beta)=0.$

Thus we have $\text{Re}\,(\alpha, \beta)=0$, $\text{Im}\,(\alpha, \beta)=0$. Therefore $(\alpha, \beta)=0$ and thus α and β are orthogonal.

Example 7. *If V is an inner product space, then prove that*
(i) $\{0\}^\perp=V$, (ii) $V^\perp=\{0\}$.

Solution. (i) We shall show that $V\subseteq\{0\}^\perp$. Let $\alpha\in V$. Since $(\alpha, 0)=0$, therefore $\alpha\in\{0\}^\perp$. Thus $\alpha\in V\Rightarrow\alpha\in\{0\}^\perp$. Therefore $V\subseteq\{0\}^\perp$. But $\{0\}^\perp\subseteq V$. Hence $\{0\}^\perp=V$.

(ii) Let $\alpha\in V^\perp$. Then by def. of $V^\perp$, we have $(\alpha, \beta)=0$ $\forall\,\beta\in V$. Taking $\beta=\alpha$, we get $(\alpha, \alpha)=0$ which implies $\alpha=0$. Thus $\alpha\in V^\perp\Rightarrow\alpha=0$. Therefore $V^\perp=\{0\}$.

Example 8. *If V is an inner product space and S, S_1, S_2 are subsets of V, then*

(i) $S_1\subseteq S_2\Rightarrow S_2^\perp\subseteq S_1^\perp$

(ii) $S^\perp=[L(S)]^\perp$

(iii) $L(S)\subseteq S^{\perp\perp}$

(iv) $L(S)=S^{\perp\perp}$ *if V is finite dimensional.*

Solution. (i) Let $S_1\subseteq S_2$. We have

$\alpha\in S_2^\perp\Rightarrow\alpha$ is orthogonal to every vector in S_2

$\Rightarrow\alpha$ is orthogonal to every vector in S_1 because $S_1\subseteq S_2$

$\Rightarrow\alpha\in S_1^\perp$.

$\therefore\ S_2^\perp\subseteq S_1^\perp$.

(ii) We have $S\subseteq L(S)$. **[From (i)]**

$\therefore\ [L(S)]^\perp\subseteq S^\perp$.

Now let $\alpha\in S^\perp$. Then α is orthogonal to every vector in S. Let β be any vector in $L(S)$. Then β is a linear combination of finite number of vectors in S. Let $\beta=\sum_{i=1}^{n} a_i\alpha_i$ where each $\alpha_i\in S$.

We have $(\alpha, \beta)=\left(\alpha, \sum_{i=1}^{n} a_i\alpha_i\right)=\sum_{i=1}^{n}\bar{a}_i(\alpha, \alpha_i)$

$=0$, since α is orthogonal to each α_i.

Scanned with CamScanner

Thus α is orthogonal to every vector β in $L(S)$. Therefore $\alpha \in [L(S)]^{\perp}$.

$\therefore \quad S^{\perp} \subseteq [L(S)]^{\perp}$.

Hence $S^{\perp}=[L(S)]^{\perp}$.

(iii) Let $\alpha \in L(S)$. If β is any vector in $S^{\perp}$, then β is orthogonal to every vector in S. Consequently β is orthogonal to α which is nothing but a linear combination of a finite number of vectors in S. Thus

$$\alpha \in L(S) \Rightarrow \alpha \text{ is orthogonal to every vector } \beta \text{ in } S^{\perp}$$
$$\Rightarrow \alpha \in (S^{\perp})^{\perp}.$$

$\therefore \quad L(S) \subseteq S^{\perp\perp}$.

(iv) We have

$S^{\perp}=[L(S)]^{\perp}$ [as proved in (ii)]

$\therefore \quad (S^{\perp})^{\perp}=[(L(S))^{\perp}]^{\perp}$

$\Rightarrow \quad S^{\perp\perp}=[L(S)]^{\perp\perp}$

$\Rightarrow \quad S^{\perp\perp}=L(S)$ [$\because$ $L(S)$ is a subspace of V. If V is finite dimensional and W is a subspace of V, then $W^{\perp\perp}=W$]

Example 9. *If S is a subset of an inner product space V, then prove that $S^{\perp}=S^{\perp\perp\perp}$.*

Solution We know that $S \subseteq S^{\perp\perp}$. Taking $S^{\perp}$ in place of S, we see that

$$S^{\perp} \subseteq (S^{\perp})^{\perp\perp}$$

i.e., $\quad S^{\perp} \subseteq S^{\perp\perp\perp}$. ...(1)

Also $\quad S \subseteq S^{\perp\perp}$

$\Rightarrow (S^{\perp\perp})^{\perp} \subseteq S^{\perp}$

$\Rightarrow \quad S^{\perp\perp\perp} \subseteq S^{\perp}$. [$\because$ $S_1 \subseteq S_2 \Rightarrow S_2^{\perp} \subseteq S_1^{\perp}$] ...(2)

From (1) and (2), we get $S^{\perp}=S^{\perp\perp\perp}$.

Example 10. *Let V be a finite-dimensional inner product space of dimension n. If $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ is an orthonormal set in V, prove that there exist vectors $\alpha_{m+1}, \ldots, \alpha_n$ such that $\{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ is an orthonormal basis for V.*

Solution Let $B=\{\alpha_1, \ldots, \alpha_m\}$. If $m=n$, then B is a complete orthonormal set in V. Therefore B will form an orthonormal basis for V. If $m<n$, then B is not a complete orthonormal set and so B can be enlarged by adding one more vector to it so that the resulting set is also an orthonormal set. This process can be continued till B becomes a complete orthonormal set and it will happen only when the number of vectors in B will become n. Thus if $m<n$, then we can find vectors $\alpha_{m+1}, \ldots, \alpha_n$ such that $B_1=\{\alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ is a complete orthonormal set in V and so is a basis for V.

Example 11. *Let W be a subspace of an inner product space V. If $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for W, then*

$$\beta \in W^{\perp} \text{ if and only if } (\beta, \alpha_i)=0 \ \forall \ i=1, 2, \ldots, n.$$

Solution. Suppose $\beta \in W^{\perp}$. Then by definition of $W^{\perp}$, we have $(\beta, \alpha)=0 \ \forall \ \alpha \in W$. Since $\alpha_1, \ldots, \alpha_n \in W$, therefore we must have $(\beta, \alpha_i)=0 \ \forall \ i=1, 2, \ldots, n$.

Conversely suppose that $(\beta, \alpha_i)=0 \ \forall \ i=1, \ldots, n$. Then to prove that $\beta \in W^{\perp}$. Let α be any vector in W. Then α can be expressed as a linear combination of the vectors belonging to the basis $\{\alpha_1, \ldots, \alpha_n\}$ of W. Therefore we can write $\alpha=\Sigma c_i\alpha_i$. Now

$$(\beta, \alpha)=(\beta, \Sigma c_i\alpha_i)=(\beta, c_1\alpha_1+\ldots+c_n\alpha_n)$$
$$=\bar{c}_1(\beta, \alpha_1)+\ldots+\bar{c}_n(\beta, \alpha_n)$$
$$=0, \text{ since } (\beta, \alpha_i)=0 \ \forall \ i=1, \ldots, n.$$

Thus $(\beta, \alpha)=0 \ \forall \ \alpha \in W$. Therefore $\beta \in W^{\perp}$.

Example 12. *Let W be a finite dimensional proper subspace of an inner product space V. Let $\alpha \in V$ and $\alpha \notin W$. Show that there is a vector $\beta \in W$ such that $\alpha-\beta \perp W$.* **(Meerut 1979)**

Solution. We know that every finite dimensional inner product space possesses an orthonormal basis. Here W is a finite dimensional inner product space. So let $\{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis of W. Consider the vector

$$\beta=(\alpha, \alpha_1)\alpha_1+\ldots+(\alpha, \alpha_n)\alpha_n=\sum_{i=1}^{n}(\alpha, \alpha_i)\alpha_i.$$

Since β is a linear combination of the vectors belonging to a basis of W, therefore $\beta \in W$. We shall show that $\alpha-\beta \perp W$.

We have for each k where $1 \leqslant k \leqslant n$,

$$(\alpha-\beta, \alpha_k)=\left(\alpha-\sum_{i=1}^{n}(\alpha, \alpha_i)\alpha_i, \alpha_k\right)$$
$$=(\alpha, \alpha_k)-\left(\sum_{i=1}^{n}(\alpha, \alpha_i)\alpha_i, \alpha_k\right) \quad \text{[by linearity of inner product]}$$

Scanned with CamScanner

$$=(\alpha, \alpha_k)-\sum_{i=1}^{n}(\alpha, \alpha_i)(\alpha_i, \alpha_k) \quad \text{[by linearity of inner product]}$$

$$=(\alpha, \alpha_k)-\sum_{i=1}^{n}(\alpha, \alpha_i)\,\delta_{ik}$$

[∵ α_i, α_k belong to an orthonormal basis]

$$=(\alpha, \alpha_k)-(\alpha, \alpha_k) \quad [\because\ \delta_{ik}=1 \text{ if } i=k \text{ and } \delta_{ik}=0 \text{ if } i\neq k]$$

$=0$.

Thus $\alpha-\beta$ is orthogonal to every vector α_k belonging to a basis of W. Hence $\alpha-\beta \perp W$. Thus we have found a vector β in W such that $\alpha-\beta \perp W$.

Example 13. *Let V be a finite-dimensional inner product space, and let $\{\alpha_1, \ldots, \alpha_n\}$ be an orthonormal basis for V. Show that for any vectors α, β in V*

$$(\alpha, \beta)=\sum_{k=1}^{n}(\alpha, \alpha_k)\,\overline{(\beta, \alpha_k)}.$$

(Meerut 1979)

Solution. Here $\{\alpha_1, \ldots, \alpha_n\}$ is an orthonormal basis for V. Since $\alpha, \beta \in V$, therefore we have

$$\alpha=\sum_{i=1}^{n}(\alpha, \alpha_i)\,\alpha_i, \text{ and } \beta=\sum_{j=1}^{n}(\beta, \alpha_j)\,\alpha_j.$$

[Refer theorem 3 on page 306]

Now $$(\alpha, \beta)=\left(\sum_{i=1}^{n}(\alpha, \alpha_i)\,\alpha_i, \sum_{j=1}^{n}(\beta, \alpha_j)\,\alpha_j\right)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha, \alpha_i)\,\overline{(\beta, \alpha_j)}\,(\alpha_i, \alpha_j)=\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha, \alpha_i)\,\overline{(\beta, \alpha_j)}\,\delta_{ij}$$

$$=\sum_{i=1}^{n}(\alpha, \alpha_i)\,\overline{(\beta, \alpha_i)}$$

[On summing with respect to j. We remember that $\delta_{ij}=1$ if $j=i$ and $\delta_{ij}=0$ if $j\neq i$]

$$=\sum_{k=1}^{n}(\alpha, \alpha_k)\,\overline{(\beta, \alpha_k)}.$$

Example 14. *If $A=\{\alpha_1, \ldots, \alpha_m\}$ is an orthonormal basis for subspace W of a finite dimensional inner product space V and $B=\{\beta_1, \ldots, \beta_t\}$ is an orthonormal basis for $W^{\perp}$, then prove that*

$$S=\{\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_t\}$$

is an orthonormal basis for V.

Solution. First we shall prove that the set S is an orthonormal set. Obviously each vector in S is a unit vector. So it remains to prove that two distinct vectors in S are orthogonal.

Now $(\alpha_i, \alpha_j)=0 \ \forall\ i=1, \ldots, m, j=1, \ldots, m, i\neq j$.

[∵ A is orthonormal]

Similarly $(\beta_i, \beta_j)=0 \ \forall\ i=1, \ldots, t, j=1, \ldots, t, i\neq j$.

Lastly we are to verify that $(\alpha_i, \beta_j)=0 \ \forall\ i=1, \ldots, m$ and $j=1, \ldots, t$. But this is true since $\alpha_i \in W$ and $\beta_j \in W^{\perp}$.

Hence the set S is an orthogonal set. Therefore it is a linearly independent set. So S will be a basis for V if $L(S)=V$.

Let $\alpha \in V$. Since $V=W \oplus W^{\perp}$, therefore we can write $\alpha=\gamma+\delta$ where $\gamma \in W$ and $\delta \in W^{\perp}$. Now $\gamma \in W$ can be expressed as a linear combination of the vectors belonging to the basis A of W. Similarly $\delta \in W^{\perp}$ can be expressed as a linear combination of the vectors belonging to the basis B of $W^{\perp}$. Therefore α can be expressed as a linear combination of the vectors belonging to $A \cup B$ *i.e.*, belonging to S. Therefore $$L(S)=V.$$

Hence S is a basis for V.

Example 15. *If W_1 and W_2 are subspaces of a finite-dimensional inner product space, then*

(*i*) $(W_1+W_2)^{\perp}=W_1^{\perp} \cap W_2^{\perp}$ **(Meerut 1989)**

(*ii*) $(W_1 \cap W_2)^{\perp}=W_1^{\perp}+W_2^{\perp}$. **(Meerut 1969, 75, 89)**

Solution. (*i*) We have $W_1 \subseteq W_1+W_2$.

$$\therefore \quad (W_1+W_2)^{\perp} \subseteq W_1^{\perp}. \qquad \ldots(1)$$

Also $W_2 \subseteq W_1+W_2$.

$$\therefore \quad (W_1+W_2)^{\perp} \subseteq W_2^{\perp}. \qquad \ldots(2)$$

From (1) and (2), we conclude that

$$(W_1+W_2)^{\perp} \subseteq W_1^{\perp} \cap W_2^{\perp}. \qquad \ldots(3)$$

Now we shall show that $W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1+W_2)^{\perp}$.

Let $\alpha \in W_1^{\perp} \cap W_2^{\perp}$. Then $\alpha \in W_1^{\perp}$ and $\alpha \in W_2^{\perp}$. Therefore α is orthogonal to every vector in W_1 and also to every vector in W_2.

Let β be any vector in W_1+W_2. Then we can write $\beta=\gamma_1+\gamma_2$ where $\gamma_1 \in W_1, \gamma_2 \in W_2$.

We have $$(\alpha, \beta)=(\alpha, \gamma_1+\gamma_2)=(\alpha, \gamma_1)+(\alpha, \gamma_2)$$
$$=0+0=0.$$

Scanned with CamScanner

Therefore α is orthogonal to every vector β in W_1+W_2. So $\alpha \in (W_1+W_2)^{\perp}$.

Thus $\alpha \in W_1^{\perp} \cap W_2^{\perp} \Rightarrow \alpha \in (W_1+W_2)^{\perp}$.

$$\therefore \quad W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1+W_2)^{\perp}. \qquad \ldots(4)$$

From (3) and (4), we have

$$(W_1+W_2)^{\perp}=W_1^{\perp} \cap W_2^{\perp}.$$

(ii) $W_1^{\perp}$ and $W_2^{\perp}$ are also subspaces of V. Taking $W_1^{\perp}$ in place of W_1 and $W_2^{\perp}$ in place of W_2 in the result (i), we get

$$(W_1^{\perp}+W_2^{\perp})^{\perp}=W_1^{\perp\perp} \cap W_2^{\perp\perp}$$

$\Rightarrow (W_1^{\perp}+W_2^{\perp})^{\perp}=W_1 \cap W_2$ [$\because$ V is finite dimensional and so $W_1^{\perp\perp}=W_1$ etc.]

$\Rightarrow (W_1^{\perp}+W_1^{\perp})^{\perp\perp}=(W_1 \cap W_2)^{\perp}$

$\Rightarrow W_1^{\perp}+W_2^{\perp}=(W_1 \cap W_2)^{\perp}$

$\Rightarrow (W_1 \cap W_2)^{\perp}=W_1^{\perp}+W_2^{\perp}$.

Example 16. *If $W_1, \ldots, W_k$ are pairwise orthogonal subspaces in an inner product space V, and if $\alpha=\alpha_1+\ldots+\alpha_k$ with α_i in W_i for $i=1, \ldots, k$, then $\| \alpha \|^2=\| \alpha_1 \|^2+\ldots+\| \alpha_k \|^2$.*

Solution. We have

$$\| \alpha \|^2=(\alpha, \alpha)=\left(\sum_{i=1}^{k} \alpha_i, \sum_{j=1}^{k} \alpha_j \right)=\sum_{i=1}^{k} \sum_{j=1}^{k} (\alpha_i, \alpha_j)$$

$=\sum_{i=1}^{k} (\alpha_i, \alpha_i)$ [On summing with respect to j and remembering that α_j is orthogonal to each α_i if $j \neq i$]

$=\| \alpha_1 \|^2+\ldots+\| \alpha_k \|^2$.

Example 17. *Find a vector of unit length which is orthogonal to the vector $\alpha=(2, -1, 6)$ of $V_3(\mathbf{R})$ with respect to standard inner product.*

Solution. Let $\beta=(x, y, z)$ be the required vector so that

$$\alpha \cdot \beta=(2, -1, 6).(x, y, z)=2x-y+6z=0.$$

Any solution of this equation, for example,

$$\beta=(2, -2, -1),$$

gives a vector orthogonal to α. But

$$\| \beta \|=[2^2+(-2)^2+(-1)^2]^{1/2}=3.$$

Hence the vector $\frac{1}{3}\beta=(\frac{2}{3}, -\frac{2}{3}, -\frac{1}{3})$ has length 1 and is orthogonal to α.

Example 18. *Find two mutually orthogonal vectors each of which is orthogonal to the vector $\alpha=(4, 2, 3)$ of $V_3(\mathbf{R})$ with respect to standard inner product.*

Solution. Let $\beta=(x_1, x_2, x_3)$ be any vector orthogonal to the vector $(4, 2, 3)$. Then $4x_1+2x_2+3x_3=0$.

Obviously $\beta=(3, -3, -2)$ is a solution of this equation. We now require a third vector $\gamma=(y_1, y_2, y_3)$ orthogonal to both α and β. This means γ must be a solution vector of the system of equations

$$4y_1+2y_2+3y_3=0, \; 3y_1-3y_2-2y_3=0.$$

Obviously $\gamma=(5, 17, -18)$ is a solution of these equations. Thus, β and γ are orthogonal to each other and to α. The solution is, of course, by no means unique.

Example 19. *Let $V_3(\mathbf{R})$ be the inner product space with respect to the standard inner product and let W be the subspace of $V_3(\mathbf{R})$ spanned by the vector $\alpha=(2, -1, 6)$. Find the projections of the vector $\beta=(4, 1, 2)$ on W and $W^{\perp}$.*

Solution. Let $\beta=\alpha_1+\alpha_2$ where $\alpha_1 \in W$ and $\alpha_2 \in W^{\perp}$. But every vector of W is a scalar multiple of α. So let $\alpha_1=k\alpha$. Then

$$\beta=k\alpha+\alpha_2. \qquad \ldots(1)$$

Since $\alpha_2 \in W^{\perp}$ and $\alpha \in W$, therefore $\alpha_2.\alpha=0$. From (1), we get

$$\beta.\alpha=k\,(\alpha.\alpha)+\alpha_2.\alpha=k\,(\alpha.\alpha).$$

But $\quad \beta.\alpha=(4, 1, 2).(2, -1, 6)=8-1+12=19$

and $\quad \alpha.\alpha=(2, -1, 6).(2, -1, 6)=4+1+36=41.$

$\therefore \quad 19=41k \quad$ or $\quad k=\frac{19}{41}$.

$\therefore \quad \alpha_1=k\alpha=(\frac{38}{41}, -\frac{19}{41}, \frac{114}{41})$.

Again, from (1)

$$\alpha_2=\beta-\alpha_1=(4, 1, 2)-(\tfrac{38}{41}, -\tfrac{19}{41}, \tfrac{114}{41})=(\tfrac{126}{41}, \tfrac{60}{41}, -\tfrac{32}{41}).$$

Exercises

1. Let $V_3(\mathbf{R})$ be the inner product space relative to the standard inner product. Then find
   (*a*) two linearly independent vectors each of which is orthogonal to the vector (1, 1, 2).
   (*b*) two mutually orthogonal vectors, each of which is orthogonal to (5, 2, −1).
   (*c*) two mutually orthogonal unit vectors, each of which is orthogonal to (2, −1, 3).
   (*d*) the projections of the vector (3, 4, 1) onto the space spanned by (1, 1, 1) and on its orthogonal complement.
2. Verify that the vectors $(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}), (\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$ and $(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3})$ form an orthonormal basis for $V_3(\mathbf{R})$ relative to the standard inner product.
3. Given the basis (2, 0, 1), (3, −1, 5), and (0, 4, 2) for $V_3(\mathbf{R})$, construct from it by the Gram-Schmidt process an orthonormal basis relative to the standard inner product.
4. Given the basis (1, 0, 0), (1, 1, 0), (1, 1, 1) for $V_3(\mathbf{R})$, construct from it by the Gram-Schmidt process an orthonormal basis relative to the standard inner product.

Scanned with CamScanner

5. Let P be the vector space over the field $\mathbf{R}$ consisting of all polynomials in x of degree $\leqslant 2$ with real coefficients. Define an inner product P by

$$(p, q)=\int_0^1 p(x)\, q(x)\, dx.$$

(*a*) Verify that this does define an inner product.

(*b*) Apply the Gram-Schmidt process to the basis $1, x, x^2$ of P to obtain an orthonormal basis, relative to this inner product.

6. If W is a finite-dimensional subspace of an inner product space V, then show that $V=W\oplus W^{\perp}$. Hence show that for an orthogonal set $\{\alpha_1, ..., \alpha_m\}$ of non-zero vectors in V and for any arbitrary β in V,

$$\sum_{k=1}^{m} \frac{|(\beta, \alpha_k)|^2}{\|\alpha_k\|^2} \leqslant \|\beta\|^2.$$

Prove that $\sum_{k=1}^{m} \frac{|(\beta, \alpha_k)|^2}{\|\alpha_k\|^2}=\|\beta\|^2$ if and only if

$$\beta=\sum_{k=1}^{m} \frac{(\beta, \alpha_k)}{\|\alpha_k\|^2}\alpha_k.$$

(Meerut 1979)

Answers

1. (a), (b), (c). Check yourself.
(d) $(\frac{8}{3}, \frac{8}{3}, \frac{8}{3})$ and $(\frac{1}{3}, \frac{4}{3}, -\frac{5}{3})$.

3. $\frac{1}{\sqrt{5}}(2, 0, 1)$, $\frac{1}{\sqrt{(270)}}(-7, -5, 14)$, $\frac{1}{3\sqrt{6}}(-1, 7, 2)$.

4. $(1, 0, 0), (0, 1, 0), (0, 0, 1)$.

5. (b) $1, \sqrt{3}\,(2x-1), \sqrt{5}\,(6x^2-6x+1)$.

§ 4. Linear functionals and adjoints.

Theorem 1. *Let V be a finite-dimensional inner product space, and f a linear functional on V. Then there exists a unique vector β in V such that $f(\alpha)=(\alpha, \beta)$ for all α in V.* **(Meerut 1974, 79, 83, 85)**

Proof. Suppose $B=\{\alpha_1, \alpha_2, ..., \alpha_n\}$ is an orthonormal basis for V and f is linear functional on V. Let

$$\beta=\sum_{j=1}^{n} \overline{f(\alpha_j)}\, \alpha_j. \quad ...(1)$$

Then β is a vector in V. Let g be a function from V to F defined by

$$g(\alpha)=(\alpha, \beta) \;\forall\; \alpha\in V. \quad ...(2)$$

We claim that g is a linear functional on V. If $a, b\in F$ and $\gamma_1, \gamma_2\in V$, then we have

$$g(a\gamma_1+b\gamma_2)=(a\gamma_1+b\gamma_2, \beta)$$
$$=a(\gamma_1, \beta)+b(\gamma_2, \beta) \quad \text{[from (2)]}$$
$$=a\, g(\gamma_1)+bg(\gamma_2).$$

Thus g is a linear functional on V. [from (2)]

Now we shall show that $g=f$. If $\alpha_k\in B$, then

$$g(\alpha_k)=(\alpha_k, \beta)$$
$$=\left(\alpha_k, \sum_{j=1}^{n} \overline{f(\alpha_j)}\, \alpha_j\right) \quad \text{[Putting } \beta \text{ from (1)]}$$
$$=\sum_{j=1}^{n} [\overline{\overline{f(\alpha_j)}}]\, (\alpha_k, \alpha_j)=\sum_{j=1}^{n} f(\alpha_j)\,(\alpha_k, \alpha_j)$$
$$=f(\alpha_k)$$

[On summing with respect to j. Remember that $(\alpha_k, \alpha_j)=1$ if $j=k$ and $(\alpha_k, \alpha_j)=0$ if $j\neq k$].

Thus g and f agree on a basis for V. Therefore $g=f$. Therefore corresponding to a linear functional f on V there exists a vector β in V such that $f(\alpha)=(\alpha, \beta) \;\forall\; \alpha\in V$.

Now to show that β is unique.

Let γ be a vector in V such that $f(\alpha)=(\alpha, \gamma) \;\forall\; \alpha\in V$.

Then $(\alpha, \beta)=(\alpha, \gamma) \;\forall\; \alpha\in V$

$$\Rightarrow (\alpha, \beta)-(\alpha, \gamma)=0 \;\forall\; \alpha\in V$$
$$\Rightarrow (\alpha, \beta-\gamma)=0 \;\forall\; \alpha\in V$$
$$\Rightarrow (\beta-\gamma, \beta-\gamma)=0 \quad \text{[taking } \alpha=\beta-\gamma]$$
$$\Rightarrow \beta-\gamma=0$$
$$\Rightarrow \beta=\gamma.$$

Thus β is unique. Hence the theorem.

Theorem 2. *For any linear operator T on a finite-dimensional inner product space V, there exists a unique linear operator T^* on V such that*

$$(T\alpha, \beta)=(\alpha, T^*\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

(Meerut 1973, 74, 78, 81, 83P, 88, 91, 93P)

Proof. Let T be a linear operator on a finite-dimensional inner product space V over the field F. Let β be a vector in V. Let f be a function from V into F defined by

$$f(\alpha)=(T\alpha, \beta) \;\forall\; \alpha\in V. \quad ...(1)$$

Here $T\alpha$ stands for $T(\alpha)$. We claim that f is a linear functional on V. Let $a, b\in F$ and $\alpha_1, \alpha_2\in V$. Then

$$f(a\alpha_1+b\alpha_2)=(T(a\alpha_1+b\alpha_2), \beta) \quad \text{[from (1)]}$$

Scanned with CamScanner

$=(aT\alpha_1+bT\alpha_2, \beta)$

$=a\,(T\alpha_1, \beta)+b\,(T\alpha_2, \beta)$ [$\because$ T is linear]

$=af\,(\alpha_1)+bf\,(\alpha_2).$

Thus f is a linear functional on V. Therefore by theorem 1, [from (1)] there exists a unique vector β' in V such that

$$f(\alpha)=(\alpha, \beta') \;\forall\; \alpha\in V. \quad \ldots(2)$$

From (1) and (2) we see that if T is a linear operator on V, then corresponding to every vector β in V there is a uniquely determined vector β' in V such that

$$(T\alpha, \beta)=(\alpha, \beta') \;\forall\; \alpha\in V.$$

Let us denote by T^* the rule which associates β with β' *i.e.* let $T^*\beta=\beta'$. Then T^* is a function from V into V and is such that

$$(T\alpha, \beta)=(\alpha, T^*\beta) \text{ for all } \alpha, \beta \text{ in } V. \quad \ldots(3)$$

Now we shall show that T^* is a linear operator on V. Let $a, b\in F$ and $\beta_1, \beta_2\in V$. Then for every α in V, we have

$(\alpha, T^*(a\beta_1+b\beta_2))=(T\alpha, a\beta_1+b\beta_2)$

$=\bar{a}\,(T\alpha, \beta_1)+\bar{b}\,(T\alpha, \beta_2)$ [from (3)]

$=\bar{a}\,(\alpha, T^*\beta_1)+\bar{b}\,(\alpha, T^*\beta_2)$

$=(\alpha, aT^*\beta_1)+(\alpha, bT^*\beta_2)$ [from (3)]

$=(\alpha, aT^*\beta_1+bT^*\beta_2).$

$\therefore$ $T^*(a\beta_1+b\beta_2)=aT^*\beta_1+bT^*\beta_2.$ [Note that in an inner product space V if $(\alpha, \beta)=(\alpha, \gamma)$ for every α in V, then $\beta=\gamma$].

Thus T^* is a linear operator on V. Therefore corresponding to a linear operator T on V there exists a linear operator T^* on V such that

$$(T\alpha, \beta)=(\alpha, T^*\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

Now to show that T^* is unique. Let S be a linear operator on V such that

$$(T\alpha, \beta)=(\alpha, S\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

Then $(\alpha, T^*\beta)=(\alpha, S\beta)$ for all α, β in V

$\Rightarrow T^*\beta=S\beta$ for every β in V

$\Rightarrow T^*=S.$

Therefore T^* is unique.

Hence the theorem.

Adjoint. Definition. *Let T be a linear operator on an inner product space V (finite-dimensional or not). We say that T has an adjoint T^* if there exists a linear operator T^* on V such that*

$$(T\alpha, \beta)=(\alpha, T^*\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

In theorem 2, we have proved that every linear operator on a finite-dimensional inner product space possesses an adjoint. But it should be noted that if V is not finite-dimensional then some linear operator on V may possess an adjoint while the other may not. In any case if T possesses an adjoint T^*, then it must be unique as we have proved in the last part of theorem 2. Also mark that the adjoint of T depends not only upon T but also on the inner product on V.

Theorem 3 *Let V be a finite-dimensional inner product space and let $B=\{\alpha_1,\ldots, \alpha_n\}$ be an ordered orthonormal basis for V. Let T be a linear operator on V and let $A=[a_{ij}]_{n\times n}$ be the matrix of T with respect to the ordered basis B. Then $a_{ij}=(T\alpha_j, \alpha_i)$.*

Proof. Since B is an orthonormal basis for V, therefore if β is any vector in V, then

$$\beta=\sum_{i=1}^{n} (\beta, \alpha_i)\,\alpha_i.$$

Taking $T\alpha_j$ in place of β, we have

$$T\alpha_j=\sum_{i=1}^{n} (T\alpha_j, \alpha_i)\,\alpha_i,\; j=1, 2,\ldots, n. \quad \ldots(1)$$

Now if $A=[a_{ij}]_{n\times n}$ be the matrix of T in the ordered basis B, then we have

$$T\alpha_j=\sum_{i=1}^{n} a_{ij}\alpha_i,\; j=1,\ldots, n. \quad \ldots(2)$$

Since the expression for $T\alpha_j$ as a linear combination of vectors in B is unique, therefore from (1) and (2) we get

$$a_{ij}=(T\alpha_j, \alpha_i),\; i=1,\ldots, n \text{ and } j=1,\ldots, n.$$

Corollary. *Let V be a finite-dimensional inner product space and let T be a linear operator on V. In any orthonormal basis for V, the matrix of T^* is the conjugate transpose of the matrix of T.*

Proof. Let $B=\{\alpha_1,\ldots, \alpha_n\}$ be an orthonormal basis for V. Let $A=[a_{ij}]_{n\times n}$ be the matrix of T in the ordered basis B. Then

$$a_{ij}=(T\alpha_j, \alpha_i). \quad \ldots(1)$$

Now T^* is also a linear operator on V. Let $C=[c_{ij}]_{n\times n}$ be the matrix of T^* in the ordered basis B. Then

$$c_{ij}=(T^*\alpha_j, \alpha_i). \quad \ldots(2)$$

Scanned with CamScanner

We have

$$c_{ij}=(T^*\alpha_j, \alpha_i)$$
$$=\overline{(\alpha_i, T^*\alpha_j)} \qquad [\because \ (\alpha, \beta)=\overline{(\beta, \alpha)}]$$
$$=\overline{(T\alpha_i, \alpha_j)} \qquad \text{[by def. of } T^*]$$
$$=\bar{a}_{ji}. \qquad \text{[from (1)]}$$

$\therefore$ $C=[\bar{a}_{ji}]_{n\times n}$. Hence $C=A^*$, where A^* is the conjugate transpose of the matrix A.

Note. It should be marked that in this corollary the basis B is an orthonormal basis and not an ordinary basis.

Theorem 4. *Suppose S and T are linear operators on an inner product space V and c is a scalar. If S and T possess adjoints, the operators S+T, cT, ST, T* will also possess adjoints. Also we have*

(i) $(S+T)^*=S^*+T^*$ **(Meerut 1972, 76, 79, 87)**

(ii) $(cT^*)=\bar{c}T^*$ **(Meerut 1970, 71, 76, 87)**

(iii) $(ST)^*=T^*S^*$ **(Meerut 1972, 78, 79, 87, 91)**

(iv) $(T^*)^*=T$. **(Meerut 1970, 87)**

Proof. (i) Since S and T are linear operators on V, therefore $S+T$ is also a linear operator on V. For every α, β in V, we have

$$((S+T)\alpha, \beta)=(S\alpha+T\alpha, \beta)=(S\alpha, \beta)+(T\alpha, \beta)$$
$$=(\alpha, S^*\beta)+(\alpha, T^*\beta) \qquad \text{[by def. of adjoint]}$$
$$=(\alpha, S^*\beta+T^*\beta)=(\alpha, (S^*+T^*)\beta).$$

Thus for the linear operator $S+T$ on V there exists a linear operator S^*+T^* on V such that

$$((S+T)\alpha, \beta)=(\alpha, (S^*+T^*)\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

Therefore the linear operator $S+T$ has an adjoint. By the definition and by the uniqueness of adjoint, we get

$$(S+T)^*=S^*+T^*.$$

(ii) Since T is a linear operator on V, therefore cT is also linear operator on V. For every α, β in V, we have

$$((cT)\alpha, \beta)=(cT\alpha, \beta)=c\,(T\alpha, \beta)=c\,(\alpha, T^*\beta)$$
$$=(\alpha, \bar{c}T^*\beta)=(\alpha, (\bar{c}T^*)\beta).$$

Thus for the linear operator cT on V there exists a linear operator $\bar{c}T^*$ on V such that

$$((cT)\alpha, \beta)=(\alpha, (\bar{c}T^*)\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

Therefore the linear operator cT possesses an adjoint. By the definition and by the uniqueness of adjoint, we get

$$(cT)^*=\bar{c}T^*.$$

(iii) ST is a linear operator on V. For every α, β in V, we have

$$((ST)\alpha, \beta)=(ST\alpha, \beta) \qquad \text{[by the def. of product of two operators]}$$
$$=(T\alpha, S^*\beta) \qquad \text{[by def. of adjoint]}$$
$$=(\alpha, T^*S^*\beta) \qquad \text{[by def. of adjoint]}$$
$$=(\alpha, (T^*S^*)\beta).$$

Thus for the linear operator ST on V, there exists a linear operator T^*S^* on V such that

$$((ST)\alpha, \beta)=(\alpha, (T^*S^*)\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

Therefore the linear operator ST has an adjoint. By the definition and by the uniqueness of adjoint, we get

$$(ST)^*=T^*S^*.$$

(iv) The adjoint of T *i.e.*, T^* is a linear operator on V. For every α, β in V, we have

$$(T^*\alpha, \beta)=\overline{(\beta, T^*\alpha)} \qquad [\because \ (\alpha, \beta)=\overline{(\beta, \alpha)}]$$
$$=\overline{(T\beta, \alpha)} \qquad \text{[by def. of adjoint]}$$
$$=(\alpha, T\beta). \qquad [\because \ (\alpha, \beta)=\overline{(\beta, \alpha)}]$$

Thus for the linear operator T^* on V, there exists a linear operator T on V such that

$$(T^*\alpha, \beta)=(\alpha, T\beta) \text{ for all } \alpha, \beta \text{ in } V.$$

Therefore the linear operator T^* has an adjoint. By the definition and by the uniqueness of adjoint, we have

$$(T^*)^*=T.$$

Note 1. If in the above theorem the vector space V is finite-dimensional, then the results will be true for arbitrary linear operators S and T. In a finite dimensional inner product space each linear operator possesses an adjoint.

Note 2. The operation of adjoint behaves like the operation of conjugation on complex numbers.

Self-adjoint transformation. Definition. *A linear operator T on an inner product space V is said to be self-adjoint if*

$$T^*=T.$$

A self-adjoint linear operator on a real inner product space is called *symmetric* while a self-adjoint linear operator on a complex inner product space is called *Hermitian*.

The zero operator $\hat{0}$ and the identity operator I on any inner

Scanned with CamScanner

product space V are self-adjoint operators. For every α, β in V, we have

$$(\hat{0}\alpha, \beta)=(0, \beta)=0=(\alpha, 0)=(\alpha, \hat{0}\beta).$$

$\therefore \quad \hat{0}^*=\hat{0}$. (Meerut 1976)

Similarly for every α, β in V, we have

$$(I\alpha, \beta)=(\alpha, \beta)=(\alpha, I\beta).$$

$\therefore \quad I^*=I$. (Meerut 1976)

Skew-symmetric or Skew-Hermitian operators. Definition. *If a linear operator T on an inner product space V is such that*

$$T^*=-T$$

then T is called skew-symmetric or skew-Hermitian according as the vector space V is real or complex.

Theorem 5. *Every linear operator T on a finite dimensional complex inner product space V can be uniquely expressed as*

$$T=T_1+iT_2$$

where T_1 and T_2 are self-adjoint linear operators on V.

Proof Let $T_1=\dfrac{T+T^*}{2}$ and $T_2=\dfrac{1}{2i}(T-T^*)$.

Then $\quad T=T_1+iT_2. \quad \ldots(1)$

Now $\quad T_1^*=[\frac{1}{2}(T+T^*)]^*=\frac{1}{2}(T+T^*)^*=\frac{1}{2}[T^*+(T^*)^*]$

$\qquad =\frac{1}{2}[T^*+T]=\frac{1}{2}(T+T^*)=T_1.$

$\therefore \quad T_1$ is self adjoint.

Also $\quad T_2^*=\left[\dfrac{1}{2i}(T-T^*)\right]^*=\overline{\left(\dfrac{1}{2i}\right)}(T-T^*)^*$

$$=-\frac{1}{2i}[T^*-(T^*)^*]=-\frac{1}{2i}(T^*-T)=\frac{1}{2i}(T-T^*)=T_2.$$

$\therefore \quad T_2$ is self-adjoint.

Thus T can be expressed in the form (1) where T_1 and T_2 are both self adjoint operators.

To show that the expression (1) for T is unique.

Let $T=U_1+iU_2$ where U_1 and U_2 are both self-adjoint.

We have $T^*=(U_1+iU_2)^*=U_1^*+(iU_2)^*$

$\qquad =U_1^*+\bar{i}\,U_2^*=U_1^*-iU_2^*$

$\qquad =U_1-iU_2 \qquad [\because U_1$ and U_2 are self-adjoint]

$\therefore \quad T+T^*=(U_1+iU_2)+(U_1-iU_2)=2U_1.$

This gives $U_1=\frac{1}{2}(T+T^*)=T_1.$

Also $\quad T-T^*=(U_1+iU_2)-(U_1-iU_2)=2iU_2.$

This gives $\quad U_2=\dfrac{1}{2i}(T-T^*)=T_2.$

Hence the expression (1) for T is unique.

Note. If T is a linear operator on a complex inner product space V which is not finite-dimensional, then the above result will be still true provided it is given that T possesses adjoint. Also in the resolution $T=T_1+iT_2$, T_1 is called the *real part* of T and T_2 is called the *imaginary part* of T.

Theorem 6. *Every linear operator T on a finite-dimensional inner product space V can be uniquely expressed as*

$$T=T_1+T_2$$

where T_1 is self-adjoint and T_2 is skew.

Proof. Let $T_1=\frac{1}{2}(T+T^*)$ and $T_2=\frac{1}{2}(T-T^*)$.

Then $\quad T=T_1+T_2. \quad \ldots(1)$

Now $\quad T_1^*=[\frac{1}{2}(T+T^*)]^*=\frac{1}{2}(T+T^*)^*=\frac{1}{2}(T^*+T)$

$\qquad =\frac{1}{2}(T+T^*)=T_1.$

$\therefore \quad T_1$ is self-adjoint.

Also $\quad T_2^*=[\frac{1}{2}(T-T^*)]^*=\frac{1}{2}(T-T^*)^*=\frac{1}{2}(T^*-T)$

$\qquad =-\frac{1}{2}(T-T^*)=-T_2.$

$\therefore \quad T_2$ is skew.

Thus T can be expressed in the form (1) where T_1 is self-adjoint and T_2 is skew.

Now to show that the expression (1) for T is unique. Let

$$T=U_1+U_2,$$

where U_1 is self-adjoint and U_2 is skew.

Then $\quad T^*=(U_1+U_2)^*=U_1^*+U_2^*$

$\qquad =U_1-U_2 \qquad [\because U_1$ is self-adjoint and U_2 is skew]

$\therefore \quad \frac{1}{2}(T+T^*)=U_1=T_1$

and $\quad \frac{1}{2}(T-T^*)=U_2=T_2.$

Hence the expression (1) for T is unique.

Note. If T is a linear operator on an inner product space V which is not finite-dimensional, then the above result will be still true provided it is given that T possesses adjoint.

Theorem 7. *A necessary and sufficient condition that a linear transformation T on an inner product space V be $\hat{0}$ is that $(T\alpha, \beta)=0$ for all α and β in V.*

Scanned with CamScanner

Proof. Let $T=\hat{0}$. Then for all α and β, we have

$$(T\alpha, \beta)=(\hat{0}\alpha, \beta)=(0, \beta)=0.$$

Hence the condition is necessary.

Conversely, let $(T\alpha, \beta)=0$ for all α and β

Taking $\beta=T\alpha$, we get

$$(T\alpha, T\alpha)=0 \ \forall \ \alpha$$

$$\Rightarrow \quad T\alpha=0 \ \forall \ \alpha \qquad [\because \ (\alpha, \alpha)=0 \Rightarrow \alpha=0]$$

$$\Rightarrow \quad T=\hat{0}.$$

Hence the condition is sufficient.

Theorem 8. *A necessary and sufficient condition that a linear transformation T on a unitary space (complex inner product space) be* $\hat{0}$ *is that* $(T\alpha, \alpha)=0$ *for all* α *in V.*

Proof Let V be a complex inner product space.

Let $T=\hat{0}$. Then for all α in V, we have

$$(T\alpha, \alpha)=(\hat{0}\alpha, \alpha)=(0, \alpha)=0.$$

Hence the condition is necessary.

Conversely, let $(T\alpha, \alpha)=0$ for all α in V.

Then for every α, β in V, we have

$$(T(\alpha+\beta), \alpha+\beta)=0 \qquad [\text{taking } \alpha+\beta \text{ in place of } \alpha]$$

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

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

$$\Rightarrow (T\alpha, \beta)+(T\beta, \alpha)=0.$$

$$[\because \ (T\alpha, \alpha)=0 \text{ and } (T\beta, \beta)=0]$$

Thus for every α, β in V, we have

$$(T\alpha, \beta)+(T\beta, \alpha)=0. \qquad ...(1)$$

Since the result (1) is true for every β in V, therefore taking $i\beta$ in place of β, we get

$$(T\alpha, i\beta)+(Ti\beta, \alpha)=0$$

$$\Rightarrow \bar{i}(T\alpha, \beta)+(iT\beta, \alpha)=0$$

$$\Rightarrow -i(T\alpha, \beta)+i(T\beta, \alpha)=0$$

$$\Rightarrow -i[(T\alpha, \beta)-(T\beta, \alpha)]=0$$

$$\Rightarrow (T\alpha, \beta)-(T\beta, \alpha)=0 \qquad [\because \ i\neq 0] \qquad ...(2)$$

Adding (1) and (2), we get

$$2(T\alpha, \beta)=0$$

$$\Rightarrow (T\alpha, \beta)=0 \ \forall \ \alpha, \beta \text{ in } V$$

$$\Rightarrow (T\alpha, T\alpha)=0 \ \forall \ \alpha\in V \qquad [\text{Taking } \beta=T\alpha]$$

$$\Rightarrow T\alpha=0 \ \forall \ \alpha\in V$$

$$\Rightarrow T=\hat{0}.$$

Hence the condition is sufficient.

Note. If V is a real inner product space, then the above theorem may fail. For example consider the vector space $V_2(\mathbf{R})$ with standard inner product defined on it. Let T be the linear operator on $V_2(\mathbf{R})$ defined as

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

Then $T\neq\hat{0}$. But

$$(T(a, b), (a, b))=((b, -a), (a, b)) \qquad [\text{by def. of } T]$$

$$=ba-ab \qquad [\text{by def. of standard inner product in } V_2(\mathbf{R})]$$

$$=0.$$

Thus $(T\alpha, \alpha)=0 \ \forall \ \alpha\in V_2(\mathbf{R})$ and yet $T\neq\hat{0}$.

However if T is self-adjoint, then the above theorem is true for real inner product spaces also. Thus we have the following theorem.

Theorem 9. *A necessary and sufficient condition that a self-adjoint linear transformation T on an inner product space V be* $\hat{0}$ *is that* $(T\alpha, \alpha)=0$ *for all* α *in V.*

Proof. T is a self-adjoint linear operator on an inner product space V *i.e.* $T^*=T$.

Let $T=\hat{0}$. Then for all α in V, we have

$$(T\alpha, \alpha)=(\hat{0}\alpha, \alpha)=(0, \alpha)=0.$$

Hence the condition is necessary.

Conversely, let $(T\alpha, \alpha)=0$ for all α in V.

Then for every α, β in V, we have

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

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

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

$$\Rightarrow (T\alpha, \beta)+(T\beta, \alpha)=0 \qquad [\because \ (T\beta, \alpha)=(\beta, T^*\alpha)]$$

$$\Rightarrow (T\alpha, \beta)+(\beta, T^*\alpha)=0 \qquad [\because \ T=T^*] \qquad ...(1)$$

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

Now two cases arise.

Case I. V is a real vector space.

In this case $(\beta, T\alpha)=(T\alpha, \beta)$

$$[\because \text{ in a Euclidean space } (\alpha, \beta)=(\beta, \alpha)]$$

$\therefore$ (1) gives

$$2\,(T\alpha, \beta)=0$$
$$\Rightarrow (T\alpha, \beta)=0 \text{ for all } \alpha, \beta \text{ in } V$$
$$\Rightarrow (T\alpha, T\alpha)=0 \text{ for all } \alpha \text{ in } V$$
$$\Rightarrow T\alpha=0 \text{ for every } \alpha \text{ in } V$$
$$\Rightarrow T=\hat{0}.$$

Case II. V is a complex vector space.

Now proceed as in theorem 8. Repeat the steps after the result (1).

Theorem 10. *A necessary and sufficient condition that a linear transformation T on a complex inner product space V (unitary space) be self-adjoint (Hermitian) is that $(T\alpha, \alpha)$ be real for all α.* **(Meerut 1978, 91, 93)**

Proof. Let V be a complex inner product space. Suppose T is a self-adjoint linear operator on V *i.e.* $T^*=T$. Then for every α in V, we have

$$(T\alpha, \alpha)=(\alpha, T^*\alpha)=(\alpha, T\alpha)=\overline{(T\alpha, \alpha)}.$$

Thus $(T\alpha, \alpha)$ is equal to its own conjugate and is therefore real. Hence the condition is necessary.

Conversely, let $(T\alpha, \alpha)$ be real for every α in V. Then to prove that T is self-adjoint. For this we should show that

$$(T\alpha, \beta)=(\alpha, T\beta), \text{ for all } \alpha, \beta \text{ in } V.$$

For every α, β in V, we have

$$(T(\alpha+\beta), \alpha+\beta)=(T\alpha+T\beta, \alpha+\beta)$$
$$=(T\alpha, \alpha)+(T\alpha, \beta)+(T\beta, \alpha)+(T\beta, \beta). \quad \ldots(1)$$

Since $(T(\alpha+\beta), \alpha+\beta)$, $(T\alpha, \alpha)$, and $(T\beta, \beta)$ are all real, therefore from (1) we see that $(T\alpha, \beta)+(T\beta, \alpha)$ is also real. So equating it to its complex conjugate, we get

$$(T\alpha, \beta)+(T\beta, \alpha)=\overline{(T\alpha, \beta)+(T\beta, \alpha)}=\overline{(T\alpha, \beta)}+\overline{(T\beta, \alpha)}$$
$$=(\beta, T\alpha)+(\alpha, T\beta).$$

Thus for all α, β in V, we have

$$(T\alpha, \beta)+(T\beta, \alpha)=(\beta, T\alpha)+(\alpha, T\beta) \quad \ldots(2)$$

Taking $i\beta$ in place of β in (2), we get

$$(T\alpha, i\beta)+(Ti\beta, \alpha)=(i\beta, T\alpha)+(\alpha, Ti\beta)$$
or $$-i\,(T\alpha, \beta)+(iT\beta, \alpha)=i\,(\beta, T\alpha)+(\alpha, iT\beta)$$
or $$-i\,(T\alpha, \beta)+i\,(T\beta, \alpha)=i\,(\beta, T\alpha)-i\,(\alpha, T\beta). \quad \ldots(3)$$

Multiplying (3) by i and adding to (2), we get

$$2\,(T\alpha, \beta)=2\,(\alpha, T\beta)$$
or $$(T\alpha, \beta)=(\alpha, T\beta).$$

$\therefore$ T is self-adjoint.

Note. If V is finite-dimensional, then we can take advantage of the fact that T must possess adjoint. So in that case the converse part of the theorem can be easily proved as follows:

Since $(T\alpha, \alpha)$ is real for all α in V, therefore

$$(T\alpha, \alpha)=\overline{(T\alpha, \alpha)}=\overline{(\alpha, T^*\alpha)}=(T^*\alpha, \alpha).$$

From this, we get for every α in V

$$(T\alpha-T^*\alpha, \alpha)=0$$
$$\Rightarrow ((T-T^*)\,\alpha, \alpha)=0$$
$$\Rightarrow T-T^*=\hat{0} \quad \text{[by theorem 8]}$$
$$\Rightarrow T=T^*.$$

Solved Examples

Example 1. *Let V be the vector space $V_2(\mathbf{C})$, with the standard inner product. Let T be the linear operator defined by*

$$T(1, 0)=(1, -2),\ T(0, 1)=(i, -1).$$

If $\alpha=(a, b)$, find $T^*\alpha$. **(Meerut 1984P)**

Solution. Let $B=\{(1, 0), (0, 1)\}$. Then B is the standard ordered basis for V. It is an orthonormal basis. Let us find $[T]_B$ *i.e.* the matrix of T in the ordered basis B.

We have $T(1, 0)=(1, -2)=1\,(1, 0)-2\,(0, 1)$

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

$$\therefore\ [T]_B=\begin{bmatrix} 1 & i \\ -2 & -1 \end{bmatrix}.$$

The matrix of T^* in the ordered basis B is the conjugate transpose of the matrix $[T]_B$.

$$\therefore\ [T^*]_B=\begin{bmatrix} 1 & -2 \\ -i & -1 \end{bmatrix}.$$

Now $(a, b)=a\,(1, 0)+b\,(0, 1)$.

$\therefore$ the coordinate matrix of $T^*(a, b)$ in the basis B

$$=\begin{bmatrix} 1 & -2 \\ -i & -1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix}=\begin{bmatrix} a-2b \\ -ia-b \end{bmatrix}. \quad \text{[See page 161]}$$

$$\therefore\ T^*(a, b)=(a-2b)\,(1, 0)+(-ia-b)\,(0, 1)$$
$$=(a-2b, -ia-b).$$

Example 2. *A linear operator on $\mathbf{R}^2$ is defined by*

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

Find the adjoint T^* if the inner product is standard one. **(Meerut 1977)**

Solution. Let $B=\{(1, 0), (0, 1)\}$. Then B is the standard ordered basis for V. It is an orthonormal basis. Let us find $[T]_B$.

We have $T(1, 0)=(1, 1)$, and $T(0, 1)=(2, -1)$.

$\therefore\ [T]_B=\begin{bmatrix}1 & 2\\ 1 & -1\end{bmatrix}$.

The matrix of T^* in the ordered basis B is the transpose of the matrix $[T]_B$. Note that $\mathbf{R}^2$ is a real inner product space.

$\therefore\ [T^*]_B=\begin{bmatrix}1 & 1\\ 2 & -1\end{bmatrix}$.

The coordinate matrix of $T^*(x, y)$ in the basis B

$$=\begin{bmatrix}1 & 1\\ 2 & -2\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}=\begin{bmatrix}x+y\\ 2x-y\end{bmatrix}.$$

$\therefore\ T^*(x, y)=(x+y, 2x-y)$.

Example 3. *Let T be the linear operator on V_2 (C) defined by*

$$T(1, 0)=(1+i, 2),\ T(0, 1)=(i, i).$$

Using the standard inner product, find the matrix of T^* in the standard ordered basis. Does T commute with T^*?

Solution. Let $B=\{(1, 0), (0, 1)\}$. Then B is the standard ordered basis for V. It is an orthonormal basis. We have

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

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

$\therefore\ [T]_B=\begin{bmatrix}1+i & i\\ 2 & i\end{bmatrix}$.

$\therefore\ [T^*]_B=$the conjugate transpose of the matrix $[T]_B$

$$=\begin{bmatrix}1-i & 2\\ -i & -i\end{bmatrix}.$$

We have $[T]_B\,[T^*]_B=\begin{bmatrix}1+i & i\\ 2 & i\end{bmatrix}\begin{bmatrix}1-i & 2\\ -i & -i\end{bmatrix}=\begin{bmatrix}3 & 3+2i\\ 3-2i & 5\end{bmatrix}$.

Also $[T^*]_B\,[T]_B=\begin{bmatrix}1-i & 2\\ -i & -i\end{bmatrix}\begin{bmatrix}1+i & i\\ 2 & i\end{bmatrix}=\begin{bmatrix}6 & 3i+1\\ -3i+1 & 2\end{bmatrix}$.

Now $[T]_B\,[T^*]_B\neq[T^*]_B\,[T_B]\ \Rightarrow\ [TT^*]_B\neq[T^*T]_B$

$$\Rightarrow TT^*\neq T^*T.$$

Example 4. *If β is a vector in an inner product space, if T is a linear transformation on that space, and if $f(\alpha)=\overline{(\beta, T\alpha)}$ for every vector α, then f is a linear functional; find a vector β' such that $f(\alpha)=(\alpha, \beta')$ for every α.*

Solution. It is given that $f(\alpha)=\overline{(\beta, T\alpha)}\ \forall\ \alpha\in V$.

$\therefore\ f$ is a function from V into F.

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

$$f(a\alpha_1+b\alpha_2)=\overline{(\beta, T(a\alpha_1+b\alpha_2))}=(T(a\alpha_1+b\alpha_2), \beta)$$
$$=(aT\alpha_1+bT\alpha_2, \beta)=a(T\alpha_1, \beta)+b(T\alpha_2, \beta)$$
$$=a\overline{(\beta, T\alpha_1)}+b\overline{(\beta, T\alpha_2)}=af(\alpha_1)+bf(\alpha_2).$$

Hence f is a linear functional on V. If V is finite-dimensional then there will exist a unique vector β' such that $f(\alpha)=(\alpha, \beta')$ for every α.

We have $f(\alpha)=\overline{(\beta, T\alpha)}=(T\alpha, \beta)=(\alpha, T^*\beta)$ for every α.

$\therefore$ if $f(\alpha)=(\alpha, \beta')$ for every α, then

$(\alpha, T^*\beta)=(\alpha, \beta')$ for every α.

Hence $\beta'=T^*\beta$.

Example 5. *If T_1 and T_2 are self-adjoint linear operators on an inner product space V, then*

(i) T_1+T_2 is self-adjoint.

(ii) If $T_1\neq\hat{0}$ and a is a non-zero scalar, then aT_1 is self-adjoint iff a is real.

Solution. (i) It is given that $T_1^*=T_1$, $T_2^*=T_2$.

We have $(T_1+T_2)^*=T_1^*+T_2^*=T_1+T_2$.

$\therefore\ T_1+T_2$ is self-adjoint.

(ii) Let a be real. Then

$$(aT_1)^*=\bar{a}T_1^*$$
$$=aT_1. \qquad [\because\ a \text{ is real and } T_1^*=T_1]$$

$\therefore\ aT_1$ is self-adjoint.

Conversely, let aT_1 be self-adjoint. Then

$$(aT_1)^*=aT_1$$
$$\Rightarrow \bar{a}T_1^*=aT_1$$
$$\Rightarrow \bar{a}T_1=aT_1 \qquad [\because\ T_1^*=T_1]$$
$$\Rightarrow (\bar{a}-a)\,T_1=\hat{0}$$
$$\Rightarrow \bar{a}-a=0 \qquad [\because\ T_1\neq\hat{0}]$$
$$\Rightarrow a=\bar{a}\ \Rightarrow a \text{ is real.}$$

Example 6. *Show that the product of two self-adjoint operators on an inner product space is self-adjoint iff the two operators commute.* (Meerut 1977, 79, 80, 82, 89)

Solution. Let S and T be two self-adjoint operators on an inner product space V. Suppose S and T commute *i.e.*, $ST=TS$. Then to prove that ST is self-adjoint.

We have $(ST)^*=T^*S^*$ $\qquad$ [$\because$ S and T are both self-adjoint]

$$=TS \qquad [\because\ ST=TS]$$
$$=ST$$

$\therefore\ ST$ is self-adjoint.

Scanned with CamScanner

Conversely, suppose that ST is self-adjoint.

Then $(ST)^*=ST$

$\Rightarrow T^*S^*=ST$

$\Rightarrow TS=ST$ [$\because$ S and T are both self-adjoint]

$\Rightarrow$ S and T commute.

Example 7. *Let V be a finite-dimensional inner product space and T a linear operator on V. If T is invertible, show that T^* is invertible and $(T^*)^{-1}=(T^{-1})^*$.* **(Meerut 1976, 81, 85)**

Solution. Suppose T is invertible. Then there exists a linear operator T^{-1} on V such that

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

$$\Rightarrow (T^{-1}T)^*=I^*=(TT^{-1})^*$$

$$\Rightarrow T^*(T^{-1})^*=I=(T^{-1})^*T^* \qquad [\because \ I \text{ is self-adjoint}]$$

$\therefore$ T^* is invertible and $(T^*)^{-1}=(T^{-1})^*$.

Example 8. *Let T be a linear operator on a finite-dimensional inner product space V. Then T is self-adjoint iff its matrix in every orthonormal basis is a self-adjoint matrix.*

Solution Let B be any orthonormal basis for T. Then

$$[T^*]_B=[T]_B^*. \qquad ...(1)$$

If T is self-adjoint, then $T^*=T$. Therefore from (1), we get

$$[T]_B=[T]_B^*.$$

$\therefore$ $[T]_B$ is a self-adjoint matrix.

Conversely let $[T]_B$ be a self-adjoint matrix. Then

$$[T]_B=[T]_B^*$$

$$=[T^*]_B. \qquad \text{[from (1)]}$$

$\therefore$ $T=T^*$ *i.e.* T is self adjoint.

Example 9. *If T is self-adjoint, then S^*TS is self-adjoint for all S; if S is invertible and S^*TS is self-adjoint, then T is self-adjoint.*

Solution. T is a self-adjoint operator. Let S be any operator. Then $(S^*TS)^*=S^*T^*(S^*)^*=S^*TS$.

$\therefore$ S^*TS is self-adjoint.

Conversely, let S be invertible. Then S^* is also invertible. If S^*TS is self-adjoint, then

$$(S^*TS)^*=S^*TS$$

$$\Rightarrow S^*T^*(S^*)^*=S^*TS \Rightarrow S^*T^*S=S^*TS$$

$$\Rightarrow (S^*)^{-1}(S^*T^*S)S^{-1}=(S^*)^{-1}(S^*TS)S^{-1}$$

$$\Rightarrow IT^*I=ITI \Rightarrow T^*=T \Rightarrow T \text{ is self-adjoint.}$$

Example 10. *If T is a self-adjoint linear operator on a finite-dimensional inner product space V, then $\det T$ is real.*

Solution. Let B be any orthonormal basis for V. Then

$$[T^*]_B=[T]_B^*.$$

Since $T^*=T$, therefore

$$[T]_B=[T]_B^*. \qquad ...(1)$$

Let $[T]_B=A$. Then from (1), we get

$$A=A^*$$

$$\Rightarrow \det A=\det A^*$$

$$\Rightarrow \det A=\overline{(\det A)}$$

[$\because$ $\det A^*$ = conjugate complex of $\det A$]

$\Rightarrow$ $\det A$ is real

$\Rightarrow$ $\det T$ is real. [$\because$ $\det T=\det [T]_B=\det A$]

Example 11. *Let V be a finite-dimensional inner product space, and let T be any linear operator on V. Suppose W is a subspace of V which is invariant under T. Then the orthogonal complement of W is invariant under T^*.*

Solution. It is given that W is invariant under T. To prove that $W^{\perp}$ is invariant under T^*.

Let β be any vector in $W^{\perp}$. Then to prove that $T^*\beta$ is in $W^{\perp}$ *i.e.* $T^*\beta$ is orthogonal to every vector in W.

Let α be any vector in W. Then

$$(\alpha, T^*\beta)=(T\alpha, \beta)$$

$$=0 \qquad [\because \ \alpha \in W \Rightarrow T\alpha \in W. \text{ Also } \beta \text{ is orthogonal to every vector in } W]$$

$\therefore$ $T^*\beta$ is orthogonal to every vector α in W.

Hence $T^*\beta$ is in $W^{\perp}$.

$\therefore$ $W^{\perp}$ is invariant under T^*.

Example 12. *Let V be a finite-dimensional inner product space, and let E be an idempotent linear operator on V, i.e. $E^2=E$. Prove that E is self-adjoint iff $EE^*=E^*E$.*

Solution. E is idempotent *i.e.*, $E^2=E$.

Let E be self-adjoint *i.e.*, $E^*=E$.

Then $EE^*=E^*E^*$ [putting E^* in place of E in L.H.S.] [$\because$ $E^*=E$]

$=E^*E$.

Conversely, let $EE^*=E^*E$. Then to prove that $E^*=E$.

For every vector β in V, we have

$$(E\beta, E\beta)=(\beta, E^*E\beta) \qquad [\because \ E^*E=EE^*]$$

$$=(\beta, EE^*\beta) \qquad [\because \ (E^*)^*=E]$$

$$=(E^*\beta, E^*\beta)$$

From this we conclude that $E\beta=0$ iff $E^*\beta=0$.

Scanned with CamScanner

Now let α be any vector in V. Let $\beta=\alpha-E\alpha$.
Then $E\beta=E(\alpha-E\alpha)=E\alpha-E^2\alpha=E\alpha-E\alpha=0$.

$\therefore\quad 0=E^*\beta=E^*(\alpha-E\alpha)=E^*\alpha-E^*E\alpha$.

This gives $E^*\alpha=E^*E\alpha$ for all α in V.

Therefore $E^*=E^*E$.

Now $E=(E^*)^*=(E^*E)^*=E^*E=E^*$.

$\therefore$ E is self-adjoint.

Example 13. *If T is skew, does it follow that so is T^2? How about T^3?*

Solution. T is skew $\Rightarrow T^*=-T$.

We have $(T^2)^*=(TT)^*=T^*T^*=(-T)(-T)=T^2$.

$\therefore$ T^2 is self-adjoint and not skew.

Again $(T^3)^*=T^*T^*T^*=(-T)(-T)(-T)=-(T^3)$.

$\therefore$ T^3 is skew

Example 14. *If T is self-adjoint, or skew, and if $T^2\alpha=0$, then $T\alpha=0$.* **(Meerut 1979)**

Solution. (i) Let T be self-adjoint *i.e.* $T^*=T$.

Suppose $T^2\alpha=0$. Then for every β in V, we have

$$(T^2\alpha,\beta)=0$$
$$\Rightarrow (TT\alpha,\beta)=0 \Rightarrow (T\alpha,T^*\beta)=0 \Rightarrow (T\alpha,T\beta)=0.$$

Taking $\beta=\alpha$, we get

$$(T\alpha,T\alpha)=0 \Rightarrow T\alpha=0.$$

(ii) Let T be skew *i.e.* $T^*=-T$.

Suppose $T^2\alpha=0$. Then for every β in V, we have

$$(T^2\alpha,\beta)=0$$
$$\Rightarrow (TT\alpha,\beta)=0 \Rightarrow (T\alpha,T^*\beta)=0 \Rightarrow (T\alpha,-T\beta)=0$$
$$\Rightarrow (-1)(T\alpha,T\beta)=0 \Rightarrow (T\alpha,T\beta)=0.$$

Taking $\beta=\alpha$, we get

$$(T\alpha,T\alpha)=0 \Rightarrow T\alpha=0.$$

Example 15. *If T is a skew-symmetric transformation on a Euclidean space, then $(T\alpha,\alpha)=0$ for every vector α.*

Solution. It is given that T is a linear operator on a real inner product space V and $T^*=-T$. For every vector α in V, we have

$$(T\alpha,\alpha)=(\alpha,T^*\alpha)=(\alpha,-T\alpha)=-(\alpha,T\alpha)=-(T\alpha,\alpha).$$

$$\therefore\quad 2(T\alpha,\alpha)=0 \Rightarrow (T\alpha,\alpha)=0.$$

Example 16. *Let V be a finite dimensional inner product space and T a self-adjoint linear operator on V. Prove that the range of T is the orthogonal complement of the null space of T i.e.*

$$R(T)=[N(T)]^{\perp}.$$

Solution. We shall prove that $R(T)$ is a subspace of $[N(T)]^{\perp}$ and $\dim R(T)=\dim [N(T)]^{\perp}$. Then we must have

$$R(T)=[N(T)]^{\perp}.$$

Let $\alpha\in R(T)$. Then there exists a vector $\beta\in V$ such that $\alpha=T\beta$. Let γ be an arbitrary vector of $[N(T)]^{\perp}$. Then $T\gamma=0$.

We have

$$(\alpha,\gamma)=(T\beta,\gamma)=(\beta,T^*\gamma)=(\beta,T\gamma) \qquad [\because T^*=T]$$
$$=(\beta,0)=0.$$

Thus $(\alpha,\gamma)=0\ \forall\ \gamma\in N(T)$. Therefore $\alpha\in[N(T)]^{\perp}$. Thus we have proved that $\alpha\in R(T) \Rightarrow \alpha\in[N(T)]^{\perp}$. Therefore $R(T)\subseteq[N(T)]^{\perp}$.

Again we know that

$$\dim R(T)+\dim N(T)=\dim V. \qquad \ldots(1)$$

Also $\qquad V=N(T)\oplus[N(T)]^{\perp}$.

$$\therefore\quad \dim N(T)+\dim[N(T)]^{\perp}=\dim V. \qquad \ldots(2)$$

From (1) and (2), we get

$$\dim R(T)=\dim[N(T)]^{\perp}.$$

Since $R(T)\subseteq[N(T)]^{\perp}$ and $\dim R(T)=\dim[N(T)]^{\perp}$, therefore

$$R(T)=[N(T)]^{\perp}.$$

Exercises

1. Let β be a fixed vector in an inner product space V over a field F. If $f_\beta: V\to F$ is defined by $f_\beta(\alpha)=(\alpha,\beta)$ for all $\alpha\in V$, then show that f_β is a linear functional on V. If V is finite dimensional, then prove that each functional on V arises in this way from some β. **(Meerut 1973)**
2. Suppose T is a self-adjoint linear operator on a finite dimensional inner product space V. If a subspace W of V is invariant under T, then prove that $W^{\perp}$ is also invariant under T.
3. If both T_1 and T_2 are self-adjoint, or else if both are skew, then $T_1T_2+T_2T_1$ is self-adjoint and $T_1T_2-T_2T_1$ is skew. What happens if one of T_1 and T_2 is self-adjoint and the other skew?
4. State whether the following statements are true or false:
   (i) The product of two symmetric matrices is symmetric. **(Meerut 1977)**
   (ii) The product of two self-adjoint operators on an inner product space is a self-adjoint operator.

Ans. (i) False; (ii) false.

Scanned with CamScanner

§ 5. **Positive Operators.**

Positive Operator. Definition. *A linear operator T on an inner product space V is called* **positive**, *in symbols $T > 0$, if it is self-adjoint and if $(T\alpha, \alpha) > 0$ whenever $\alpha \neq 0$.* **(Meerut 1972, 82)**

If $\alpha = 0$, then $(T\alpha, \alpha) = 0$. Thus if T is positive, then $(T\alpha, \alpha) \geq 0$ for all α and $(T\alpha, \alpha) = 0 \Rightarrow \alpha = 0$. Also if T is self-adjoint and if $(T\alpha, \alpha) \geq 0$ for all α, and $(T\alpha, \alpha) = 0 \Rightarrow \alpha = 0$, then T is positive. If V is a complex inner product space, then by theorem 10 of § 4, $(T\alpha, \alpha) \geq 0$ for every α implies that T must be self-adjoint. Therefore a linear operator T on a complex inner product space is positive if and only if $(T\alpha, \alpha) > 0$ whenever $\alpha \neq 0$.

Non-negative operator. Definition. *A linear operator T on an inner product space V is called* **non-negative**, *in symbols $T \geq 0$, if it is self-adjoint and if $(T\alpha, \alpha) \geq 0$ for all α in V.*

Every positive operator is also a non-negative operator. If T is a non-neagtive operator, then $(T\alpha, \alpha) = 0$ is possible even if $\alpha \neq 0$. Therefore a non-negative operator may or may not be a positive operator.

If S and T are two linear operators on an inner product space V, then we define $S > T$ (or $T < S$) if $S - T > 0$.

Note. Some authors call a positive operator by the name '*strictly positive*' or '*positive definite*'. Also they use the phrase '*positive operator*' in place of '*non-negative operator*'.

Theorem 1. *Let V be an inner product space, and let T be a linear operator on V. Let p be the function defined on ordered pairs of vectors α, β in V by*

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

Show that the function p is an inner product on V if and only if T is a positive operator. **(Meerut 1969, 88)**

Proof. The function p obviously satisfies linearity property. If $a, b \in F$ and $\alpha_1, \alpha_2 \in V$, then

$p(a\alpha_1 + b\alpha_2, \beta) = (T(a\alpha_1 + b\alpha_2), \beta) = (aT\alpha_1 + bT\alpha_2, \beta)$
$= a(T\alpha_1, \beta) + b(T\alpha_2, \beta) = ap(\alpha_1, \beta) + bp(\alpha_2, \beta).$

$\therefore$ the function p satisfies linearity property.

Now the function p will be an inner product on V if and only if $p(\alpha, \beta) = \overline{p(\beta, \alpha)}$ and $p(\alpha, \alpha) > 0$ if $\alpha \neq 0$.

We have $p(\alpha, \beta) = (T\alpha, \beta)$ and $\overline{p(\beta, \alpha)} = \overline{(T\beta, \alpha)} = (\alpha, T\beta)$. Also $p(\alpha, \alpha) = (T\alpha, \alpha)$.

$\therefore$ the function p will be an inner product on V if and only if
(i) $(T\alpha, \beta) = (\alpha, T\beta)\ \forall\ \alpha, \beta \in V$ *i.e.* T is self-adjoint.
(ii) $(T\alpha, \alpha) > 0$, if $\alpha \neq 0$.

Hence the function p will be an inner product on V if and only if the linear operator T is positive.

Now we shall show that if V is finite-dimensional, then every inner product on V is of the type just described.

Theorem 2. *Let $V(F)$ be a finite-dimensional inner product space with inner product $(\ ,\ )$. If p is any inner product on V, there is a unique positive linear operator T on V such that $p(\alpha, \beta) = (T\alpha, \beta)$ for all α, β in V.* **(Meerut 1980)**

Proof. Fix a vector β in V. Let f be a mapping from V to F defined as $f(\alpha) = p(\alpha, \beta)\ \forall\ \alpha \in V$. Since the inner product p satisfies linearity property, therefore f is a linear functional on V. By theorem 1 on page 330, there axists a unique vector β' in V such that $f(\alpha) = (\alpha, \beta')$ for all α in V *i.e.* $p(\alpha, \beta) = (\alpha, \beta')$ for all α in V.

Now let us define a mapping T from V into V by the rule $T\beta = \beta'$. Then we have

$$p(\alpha, \beta) = (\alpha, \beta') = (\alpha, T\beta) \text{ for all } \alpha, \beta \text{ in } V. \quad \ldots(1)$$

We also have

$p(\alpha, \beta) = (\alpha, T\beta)$

$= \overline{p(\beta, \alpha)}$ [$\because$ by conjugate property of inner product p, we have $p(\alpha, \beta) = \overline{p(\beta, \alpha)}$]

$= \overline{(\beta, T\alpha)}$ [by (1)]

$= (T\alpha, \beta).$ [by conjugate property of inner product $(\ ,\ )$]

Thus we have

$$p(\alpha, \beta) = (T\alpha, \beta) \text{ for all } \alpha, \beta \text{ in } V. \quad \ldots(2)$$

Now we shall show that T is **linear**. Let $\alpha_1, \alpha_2 \in V$ and $a_1, a_2 \in F$. Then for all γ in V, we have

$(T(a_1\alpha_1 + a_2\alpha_2), \gamma) = p(a_1\alpha_1 + a_2\alpha_2, \gamma)$ [by (2)]

$= a_1 p(\alpha_1, \gamma) + a_2 p(\alpha_2, \gamma)$ [by linearity of p]

$= a_1(T\alpha_1, \gamma) + a_2(T\alpha_2, \gamma)$ [by (2)]

$= (a_1T\alpha_1 + a_2T\alpha_2, \gamma)$ [by linearity of inner product $(\ ,\ )$]

Therefore, we have $T(a_1\alpha_1 + a_2\alpha_2) = a_1T\alpha_1 + a_2T\alpha_2$. Hence T is a linear operator. Thus we have proved the existence of a linear operator T with $p(\alpha, \beta) = (T\alpha, \beta)$. Since p is an inner product, therefore by theorem 1, T is positive.

Scanned with CamScanner

Uniqueness of T. Suppose there are two linear operators T and U such that $p(\alpha, \beta)=(T\alpha, \beta)=(U\alpha, \beta)$ for all α, β in V. Then we have

$$(T\alpha-U\alpha, \beta)=0 \ \forall \ \alpha, \beta \in V. \qquad \ldots(3)$$

Suppose we keep α fixed. Then from (3), we see that the vector $T\alpha-U\alpha$ is orthognnal to every vector β in V. Therefore $T\alpha-U\alpha$ is a zero vector Thus we have $T\alpha-U\alpha=0 \ \forall \ \alpha \in V$ Therefore $T\alpha=U\alpha \ \forall \ \alpha \in V$ and so $T=U$ Hence T is unique.

Theorem 3. *Let V be a finite-dimensional inner product space and T a linear operator on V. Then T is positive if and only if there is an invertible linear operator U on V such that $T=U^*U$.*

(Meerut 1973, 76, 78, 82, 85, 87, 89)

Proof. Let $T=U^*U$, where U is an invertible linear operator on V. We have $T^*=(U^*U)^*=U^*(U^*)^*=U^*U=T$. Therefore T is self-adjoint.

Also $(T\alpha, \alpha)=(U^*U\alpha, \alpha)=(U\alpha, U^{**}\alpha)=(U\alpha, U\alpha) \geqslant 0$.

Further $(T\alpha, \alpha)=0 \Rightarrow (U\alpha, U\alpha)=0 \Rightarrow U\alpha=0$

$\Rightarrow \alpha=0$ [$\because$ U is invertible and V is finite-dimensional implies that U is non-singular]

Therefore if $\alpha \neq 0$, then $(T\alpha, \alpha) > 0$. Hence T is positive.

Conversely, suppose that T is positive. Then by theorem 1, $p(\alpha, \beta)=(T\alpha, \beta)$ is an inner product on V. Let $\{\alpha_1, \ldots, \alpha_n\}$ be a basis for V which is orthonormal with respect to the inner product (,) and let $\{\beta_1, \ldots, \beta_n\}$ be a basis orthonormal with respect to the inner product p. Then

$$p(\beta_i, \beta_j)=\delta_{ij}=(\alpha_i, \alpha_j).$$

Now let U be the unique linear operator on V such that $U\beta_i=\alpha_i$, $i=1, \ldots n$ Obviously U is invertible because it carries a basis onto a basis We have

$$p(\beta_i, \beta_j)=(\alpha_i, \alpha_j)=(U\beta_i, U\beta_j).$$

Now let α, β be any two vectors in V.

Let $\alpha=\sum_{i=1}^{n} x_i \beta_i$ and $\beta=\sum_{j=1}^{n} y_j \beta_j$. Then

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

$$=p\left(\sum_{i=1}^{n} x_i \beta_i, \sum_{j=1}^{n} y_j \beta_j\right)=\sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j \, p(\beta_i, \beta_j)$$

$$=\sum_{i=1}^{n} \sum_{j=1}^{n} x_i \bar{y}_j (U\beta_i, U\beta_j)=\left(\sum_{i=1}^{n} x_i U\beta_i, \sum_{j=1}^{n} y_j U\beta_j\right)$$

$$=\left(U \sum_{i=1}^{n} x_i \beta_i, U \sum_{j=1}^{n} y_j \beta_j\right)=(U\alpha, U\beta)=(U^*U\alpha, \beta)$$

Thus for all α, β in V, we have

$$(T\alpha, \beta)=(U^*U\alpha, \beta).$$

$$\therefore \quad T=U^*U.$$

Hence the theorem

Theorem 4. *Let T be a linear operator on a finite-dimensional inner product space V. Let $A=[a_{ij}]_{n \times n}$ be the matrix of T relative to an ordered orthonormal basis $B=\{\alpha_1, \ldots, \alpha_n\}$. Then T is positive if and only if the matrix A satisfies the following two conditions :*

(*i*) *$A=A^*$ i.e. A is self-adjoint,*

(*ii*) $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \bar{x}_i x_j > 0$ *where $x_1, \ldots, x_n$ are any n scalars not all zero.*

Proof. Let α be any vector in V. Let $\alpha=x_1\alpha_1+\ldots+x_n\alpha_n$.

Then $(T\alpha, \alpha)=\left(T \sum_{j=1}^{n} x_j \alpha_j, \sum_{i=1}^{n} x_i \alpha_i\right)$

$$=\left(\sum_{j=1}^{n} x_j T\alpha_j, \sum_{i=1}^{n} x_i \alpha_i\right)=\sum_{j=1}^{n} \sum_{i=1}^{n} x_j \bar{x}_i (T\alpha_j, \alpha_i)$$

$$=\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \bar{x}_i x_j \qquad [\because \text{ by theorem 3 of § 4, } (T\alpha_j, \alpha_i)=a_{ij}]$$

Now suppose T is positive. Then $T=T^*$. Therefore $A=A^*$. If $x_1, \ldots\ldots, x_n$ are any n scalars not all zero, then $\alpha=x_1 \alpha_1+\ldots+x_n \alpha_n$ is a non-zero vector in V. Since T is positive, therefore $(T\alpha, \alpha) > 0$. Hence $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \bar{x}_i x_j > 0$.

Conversely, suppose that the conditions (i) and (ii) of the theorem hold. Then $A=A^* \Rightarrow T=T^*$. Also (iii) implies that $(T\alpha, \alpha) > 0$ if $\alpha \neq 0$. Note that if $0 \neq \alpha \in V$, then we can write

Scanned with CamScanner

$\alpha = x_1\alpha_1 + \dots + x_n\alpha_n$ where $x_1, \dots, x_n$ are scalars not all zero. Hence T is positive.

Positive matrix. Definition. *Let $A = [a_{ij}]_{n \times n}$ be a square matrix of order n over the field of real or complex numbers. Then A is said to be positive if*

(*i*) $A = A^*$,

and (*ii*) $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, \bar{x}_i\, x_j > 0$ *where $x_1, \dots, x_n$ are any n scalars not all zero.*

If the field is real, then the bars may be omitted. If the field is complex then the condition (i) will automatically follow from the condition (ii) and so it may be omitted.

Now the theorem 8 may be stated as follows :

Let V be a finite-dimensional inner product space and B an ordered orthonormal basis for V. If T is a linear operator on V, then T is positive if and only if the matrix of T in the ordered basis B is positive.

Principal minors of a matrix.

Definition. *Let $A = [a_{ij}]_{n \times n}$ be a square matrix of order n over an arbitrary field F. The principal minors of A are the n scalars defined as*

$$\det A^{(k)} = \det \begin{bmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{bmatrix},\ k = 1, \dots, n$$

We shall now give without proof a criterion for a matrix to be positive.

Let A be an $n \times n$ self-adjoint matrix over the field of real or complex numbers. Then A is positive if and only if the principal minors of A are all positive.

If det A is not positive, then the matrix A is not positive.

Solved Examples

Example 1. *Suppose S and T are two positive linear operators on an inner product space V. Then show that $S+T$ is also positive.*

Solution. Since S and T are both positive, therefore $S^* = S$ and $T^* = T$.

We have $(S+T)^* = S^* + T^* = S + T$.

$\therefore$ $S+T$ is also self-adjoint.

Also if α is any vector in V, then

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

Since S and T are both positive, therefore

$(S\alpha, \alpha) > 0$, and $(T\alpha, \alpha) > 0$ if $\alpha \neq 0$.

So $((S+T)\,\alpha, \alpha) > 0$ if $\alpha \neq 0$.

Hence $S+T$ is positive.

Example 2. *Which of the following matrices are positive ?*

(*i*) $\begin{bmatrix} 1 & 1+i \\ 1-i & 3 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

(*iii*) $\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$.

Solution. (i) Let $A = \begin{bmatrix} 1 & 1+i \\ 1-i & 3 \end{bmatrix}$.

We have A^* = the conjugate transpose of A

$$= \begin{bmatrix} 1 & 1+i \\ 1-i & 3 \end{bmatrix} = A.$$

$\therefore$ A is self-adjoint.

Now the principal minors of A are 1, $\begin{vmatrix} 1 & 1+i \\ 1-i & 8 \end{vmatrix}$.

We have $\begin{vmatrix} 1 & 1+i \\ 1-i & 3 \end{vmatrix} = 3 - 2 = 1$.

Since A is self-adjoint and the principal minors of A are all positive, therefore A is a positive matrix.

(ii) The given matrix is not self-adjoint because its transpose is not equal to itself. Hence it is not positive.

(iii) Let A denote the given matrix. Obviously $A = A^*$ i.e. A = the transpose of A.

The principal minors of A are

$$1,\ \begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{vmatrix},\ \begin{vmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{vmatrix}.$$

All these are positive as can be easily seen. Hence A is positive.

Example 3. *Prove that every entry on the main diagonal of a positive matrix is positive.* (Meerut 1976, 83)

Solution. Let $A = [a_{ij}]_{n \times n}$ be a positive matrix. Then ...(1)

$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, \bar{x}_i\, x_j > 0$,

Scanned with CamScanner

where $x_1, \ldots, x_n$ are any n scalars not all zero. Now suppose that out of n scalars $x_1, \ldots, x_n$ we take $x_i = 1$ and each of the remaining $n-1$ scalars is taken as 0. Then from (1) we conclude that $a_{ii} > 0$. Then $a_{ii} > 0$ for each $i = 1, \ldots, n$. Hence each entry on the main diagonal of a positive matrix is positive.

§ 6. Unitary operators.

Definitions. *Let U and V be two inner product spaces over the same field F and let T be a linear transformation from U into V. We say that*

(*i*) T **preserves inner products** *if* $(T\alpha, T\beta) = (\alpha, \beta)$ *for all* α, β *in* U.

(*ii*) T **preserves norms** if $\| T\alpha \| = \| \alpha \| \; \forall \; \alpha \in U$.

(*iii*) *T is an* **isometry** *if T preserves distances, i.e., if* $\| T\alpha - T\beta \| = \| \alpha - \beta \|$ *for all* α, β *in* U.

Note that $d(T\alpha, T\beta) = \| T\alpha - T\beta \|$.

Theorem 1. *Let U and V be two inner product spaces over the same field F and let T be a linear operator from U into V. Then the following three conditions on T are equivalent:*

(*i*) *T preserves inner products.*

(*ii*) *T preserves norms.*

(*iii*) *T is an isometry.*

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

It is given that T preserves inner products. Therefore $(T\alpha, T\beta) = (\alpha, \beta)$ for all α, β in U. Taking $\beta = \alpha$, we get $(T\alpha, T\alpha) = (\alpha, \alpha)$ i.e. $\| T\alpha \|^2 = \| \alpha \|^2$. Thus $\| T\alpha \| = \| \alpha \|$ for every α in U. Hence T preserves norms.

(ii) $\Rightarrow$ (iii).

It is given that $\| T\alpha \| = \| \alpha \|$ for every α in U. Therefore for all α, β in U, we have

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

$$\Rightarrow \| T\alpha - T\beta \| = \| \alpha - \beta \| \quad \text{(taking } \alpha - \beta \text{ in place of } \alpha)$$

$\Rightarrow T$ is an isometry.

(iii) $\Rightarrow$ (i).

It is given that $\| T\alpha - T\beta \| = \| \alpha - \beta \|$ for all α, β in U, taking $\beta = 0$, we see that

$$\| T\alpha \| = \| \alpha \| \text{ for every } \alpha \text{ in } U$$

$$\Rightarrow (T\alpha, T\alpha) = (\alpha, \alpha) \text{ for all } \alpha \text{ in } U.$$

Now let $\alpha, \beta \in U$. Then

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

$$\Rightarrow (T\alpha + T\beta, T\alpha + T\beta) = (\alpha + \beta, \alpha + \beta)$$

$$\Rightarrow (T\alpha, T\alpha) + (T\alpha, T\beta) + (T\beta, T\alpha) + (T\beta, T\beta) = (\alpha, \alpha) + (\alpha, \beta) + (\beta, \alpha) + (\beta, \beta)$$

$$\Rightarrow (T\alpha, T\beta) + (T\beta, T\alpha) = (\alpha, \beta) + (\beta, \alpha)$$

$$[\because \; (T\alpha, T\alpha) = (\alpha, \alpha) \text{ and } (T\beta, T\beta) = (\beta, \beta)]$$

Thus if α, β are any vectors in U, then

$(T\alpha, T\beta) + (T\beta, T\alpha) = (\alpha, \beta) + (\beta, \alpha)$. ...(1)

If F is the field of real numbers, then

$$(T\beta, T\alpha) = (T\alpha, T\beta) \text{ and } (\beta, \alpha) = (\alpha, \beta).$$

$\therefore$ (1) gives $2(T\alpha, T\beta) = 2(\alpha, \beta)$

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

If F is the field of complex numbers, then $\beta \in U \Rightarrow i\beta \in U$. So replacing β by $i\beta$ in (1), we get

$$(T\alpha, Ti\beta) + (Ti\beta, T\alpha) = (\alpha, i\beta) + (i\beta, \alpha)$$

$$\Rightarrow (T\alpha, iT\beta) + (iT\beta, T\alpha) = \bar{i}\,(\alpha, \beta) + i\,(\beta, \alpha)$$

$$\Rightarrow \bar{i}\,(T\alpha, T\beta) + i\,(T\beta, T\alpha) = -i\,(\alpha, \beta) + i\,(\beta, \alpha)$$

$$\Rightarrow -i\,(T\alpha, T\beta) + i\,(T\beta, T\alpha) = -i\,(\alpha, \beta) + i\,(\beta, \alpha)$$

$$\Rightarrow -(T\alpha, T\beta) + (T\beta, T\alpha) = -(\alpha, \beta) + (\beta, \alpha). \quad \text{...(2)}$$

Subtracting (2) from (1), we get

$$2\,(T\alpha, T\beta) = 2\,(\alpha, \beta)$$

$$\Rightarrow (T\alpha, T\beta) = (\alpha, \beta) \text{ for all } \alpha, \beta \text{ in } U.$$

Hence the theorem.

Inner product space isomorphism.

Definition. *Let U and V be inner product spaces over the same field F, and let T be a linear transformation from U into V. Then T is said to be an* **inner product space isomorphism** *if*

(*i*) *T is invertible i.e. T is one-one and onto,*

(*ii*) *T preserves inner products.*

Also the inner product spaces U and V are then said to be **isomorphic** *and we write $U \cong V$.*

If T preserves inner products, then $\| T\alpha \| = \| \alpha \|$ for every α in U i.e. $(T\alpha, T\alpha) = (\alpha, \alpha)$ for every α in U. So

$$T\alpha = 0 \Rightarrow (\alpha, \alpha) = 0 \Rightarrow \alpha = 0.$$

$\therefore$ T is non-singular i.e. T is one-one.

Therefore an inner product space isomorphism from U onto V can also be defined as a linear transformation from U onto V which preserves inner products.

Theorem 2. *Let V be any unitary vector space of dimension n with inner product $(\,,\,)$. Then V is inner product space isomorphic to V_n* **(C)** *with standard inner product.*

Scanned with CamScanner

Proof. Let $B=\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthonormal basis for the vector space V. Then

$(\alpha_i, \alpha_j)=\delta_{ij}$, where $\delta_{ij}=0$ if $i \neq j$ and $\delta_{ij}=1$ if $i=j$.

Now let β_1 be any vector in V and let $(x_1, x_2, \ldots, x_n)$ be the coordinate vector of β_1 relative to the basis B. Then

$$\beta_1=x_1\alpha_1+x_2\alpha_2+\ldots+x_n\alpha_n.$$

Let f be a mapping from V to V_n (C) defined as

$$f(\beta_1)=(x_1, x_2, \ldots, x_n).$$

Then we know that f is an isomorphism from V onto V_n (C).

Now if $\beta_2=y_1\alpha_1+y_2\alpha_2+\ldots+y_n\alpha_n$ be any vector in V, then

$(\beta_1, \beta_2)=(\Sigma\, x_i\alpha_i, \Sigma\, y_j\alpha_j)$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j\,(\alpha_i, \alpha_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\,\bar{y}_j\,\delta_{ij}$$

$$=\sum_{i=1}^{n} x_i\,\bar{y}_i, \text{ on summing with respect to } j$$

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

$=f(\beta_1).f(\beta_2)$

$=$the standard inner product of $f(\beta_1)$ and $f(\beta_2)$ in V_n (C).

Hence V is an inner product space isomorphic to V_n (C) with standard inner product.

Corollary. *In an n-dimensional unitary vector space V the dot product of the coordinate vectors of any two vectors of V is invariant under transformation from one orthonormal basis to another.*

Proof. Let $B_1=\{\alpha_1, \ldots, \alpha_n\}$ and $B_2=\{\gamma_1, \ldots, \gamma_n\}$ be two orthonormal bases for V and let β_1, β_2 be any two vectors in V. Suppose X_1, X_2 are the coordinate vectors of β_1, β_2 relative to the basis B_1 and Y_1, Y_2 are their coordinate vectors relative to the basis B_2. Then

$$(\beta_1, \beta_2)=X_1.X_2$$

and

$$(\beta_1, \beta_2)=Y_1.Y_2.$$

$$\therefore \quad X_1.X_2=Y_1.Y_2.$$

Hence the result.

Unitary operator.

Definition. *A linear operator T on an inner product space V over the field F is said to be a* **unitary operator** *if T is an inner product space isomorphism of V onto itself.*

In other words T is a unitary operator on an inner product space V if T is invertible and if T preserves inner products.

Theorem 3. *Let T be a linear operator on an inner product space V. Then T is unitary if and only if the adjoint T^* of T exists and $TT^*=T^*T=I$.* (Meerut 1974, 76, 83, 87, 89)

Proof. Suppose T is unitary. Then T is invertible.

For all α, β in V, we have

$(T\alpha, \beta)=(T\alpha, I\beta)$ $\qquad$ [$\because$ $I\beta=\beta$]

$=(T\alpha, TT^{-1}\beta)$ $\qquad$ [$\because$ $TT^{-1}=I$]

$=(\alpha, T^{-1}\beta)$ $\qquad$ [$\because$ T is unitary $\Rightarrow$ T preserves inner products]

$\therefore$ T^{-1} is the adjoint of T. Thus T is unitary implies that T^* exists and $TT^*=I=T^*T$.

Conversely, suppose that T^* exists and $TT^*=T^*T=I$. Then T is invertible and $T^{-1}=T^*$. So T will be unitary if we show that T preserves inner products.

For all α, β in V, we have

$$(T\alpha, T\beta)=(\alpha, T^*T\beta)=(\alpha, I\beta)=(\alpha, \beta).$$

$\therefore$ T is unitary.

Theorem 4. *A linear operator T on a finite-dimensional inner product space is unitary iff $T^*T=I$.* (Meerut 1990)

Proof. Since T is a linear operator on a finite dimensional inner product space, therefore T^* exists. Also $T^*T=I$ implies that T is invertible.

Now give the same proof as in theorem 3.

Theorem 5. *A linear operator T on a finite dimensional inner product space V is unitary if and only if T preserves inner products.* (Meerut 1977)

Proof. T is a linear operator on a finite dimensional inner product space V.

Suppose T is unitary. Then T preserves inner products.

Conversely suppose that T preserves inner products. Then T will be unitary if we prove that T is invertible. Since T preserves inner products, therefore $(T\alpha, T\beta)=(\alpha, \beta)$ for every α, β in V. So Taking $\beta=\alpha$, we get $(T\alpha, T\alpha)=(\alpha, \alpha)$ for every α in V.

$$T\alpha=0 \Rightarrow (\alpha, \alpha)=0 \Rightarrow \alpha=0.$$

$\therefore$ T is non-singular *i.e.*, T is one-one. Now V is finite dimensional. Therefore T is one-one implies that T is onto. Hence T is invertible.

$\therefore$ T is unitary.

Scanned with CamScanner

Theorem 6. *A linear operator T on a finite dimensional inner product space V is unitary if and only if it takes an orthonormal basis of V onto an orthonormal basis of V.* **(Meerut 1976, 88)**

Proof. Suppose T is a unitary operator on a finite-dimensional inner product space V. Then T preserves inner products. If $B=\{\alpha_1,\ldots,\alpha_n\}$ is an orthonormal basis of V, then to show that $\{T\alpha_1,\ldots,T\alpha_n\}$ is an orthonormal basis of V. For $i=1,\ldots,n$, and $j=1,\ldots,n$, we have

$$(T\alpha_i, T\alpha_j)=(\alpha_i, \alpha_j) \qquad [\because\ T \text{ preserves inner products}]$$

$$=\delta_{ij} \qquad [\because\ \alpha_i, \alpha_j \in \text{an orthonormal set } B]$$

$\therefore$ $\{T\alpha_1,\ldots,T\alpha_n\}$ is an orthonormal set in V. It will be a basis of V because it is linearly independent and it contains n vectors which is the dimension of V.

Conversely, suppose that T is a linear operator on V such that both

$$\{\alpha_1,\ldots,\alpha_n\} \text{ and } \{T\alpha_1,\ldots,T\alpha_n\}$$

are orthonormal bases of V. Then $(\alpha_i, \alpha_j)=\delta_{ij}=(T\alpha_i, T\alpha_j)$. Since T carries a basis of V onto a basis of V, therefore T is invertible.

Now T will be unitary if T preserves inner products.

For any $\alpha=\sum_{i=1}^{n} x_i\alpha_i$, $\beta=\sum_{j=1}^{n} y_j\alpha_j$ in V, we have

$$(\alpha, \beta)=\left(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j\,(\alpha_i, \alpha_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j\,\delta_{ij}=\sum_{i=1}^{n} x_i\bar{y}_i.$$

Also $$(T\alpha, T\beta)=\left(T\sum_{i=1}^{n} x_i\alpha_i, T\sum_{j=1}^{n} y_j\alpha_j\right)$$

$$=\left(\sum_{i=1}^{n} x_i\,T\alpha_i, \sum_{j=1}^{n} y_j\,T\alpha_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j\,(T\alpha_i, T\alpha_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i\bar{y}_j\,\delta_{ij}=\sum_{i=1}^{n} x_i\bar{y}_i=(\alpha, \beta).$$

$\therefore$ T preserves inner products. Hence T is unitary.

Unitary or Isometric matrix. Definition. (Meerut 1980). *Let A be an $n\times n$ matrix over the field of real or complex numbers. Then A is said to be* **unitary** *or* **Isometric** *if $A^*A=I$.*

Orthogonal matrix. Definition. (Meerut 1980). *A real or complex $n\times n$ matrix A is said to be orthogonal, if $A'A=I$ where A' is the transpose of A.*

A real orthogonal matrix is unitary and a unitary matrix over the real field is orthogonal. Therefore sometimes a unitary matrix over the real field is also called an orthogonal matrix.

Unitarily Equivalent matrices. Definition. *Let A and B be complex $n\times n$ matrices. We say that B is* **unitarily equivalent to** *A if there is an $n\times n$ unitary matrix P such that $B=P^*AP$.*

Orthogonally Equivalent matrices. Definition. *Let A and B be complex $n\times n$ matrices. We say that B is* **orthogonally equivalent to** *A if there is an $n\times n$ orthogonal matrix P such that*

$$B=P^*\,AP.$$

Unitarily Equivalent matrices are also sometimes called unitarily similar matrices. Similarly orthogonally equivalent matrices are also sometimes called orthogonally similar matrices.

Theorem 7. *Let V be a finite-dimensional inner product space and let U be a linear operator on V. Then U is unitary if and only if the matrix of U in some (or every) ordered orthonormal basis is a unitary matrix.* **(Meerut 1977)**

Proof. Let $B=\{\alpha_1, \alpha_2,\ldots,\alpha_n\}$ be an ordered orthonormal basis of V. Let A be the matrix of U relative to the basis B *i.e.*, let

$$[U]_B=A. \quad \text{Then } [U^*]_B=A^*.$$

Now suppose U is unitary. Then

$$U^*U=I$$

$$\Rightarrow [U^*U]_B=[I]_B \Rightarrow [U^*]_B\,[U_B]=I$$

$$\Rightarrow A^*A=I \Rightarrow \text{the matrix } A \text{ is unitary.}$$

Conversely suppose that the matrix A is unitary. Then

$$A^*A=I$$

$$\Rightarrow [U^*]_B\,[U]_B=I \Rightarrow [U^*U]_B=I$$

$$\Rightarrow U^*U=I \Rightarrow \text{the linear operator } U \text{ is unitary.}$$

Solved Examples

Example 1. *Show that the set of all unitary operators on an inner product space V is a group under the operation of composition.*

Solution. The product of two unitary operators is unitary. If T_1, T_2 are two unitary operators, then both of them are invertible.

Scanned with CamScanner

Therefore T_1T_2 is also invertible. Also $\| T_1T_2\alpha \| = \| T_2\alpha \| = \| \alpha \|$ for each α. Hence T_1T_2 is unitary.

The composition of linear operators is an associative operation.

The identity operator I is invertible and $\| I\alpha \| = \| \alpha \|$ for each α. Hence I is also unitary.

Finally, if T is a unitary operator, then T is invertible. Now to show that T^{-1} is also unitary. Let α be any vector in V and let $T^{-1}\alpha = \beta$. Then $T\beta = \alpha$. We have

$$\begin{aligned} \| T^{-1}\alpha \| &= \| \beta \| \\ &= \| T\beta \| \qquad [\because\ T \text{ is unitary}] \\ &= \| \alpha \|. \end{aligned}$$

$\therefore$ T^{-1} is unitary. Hence the result.

Example 2. *Show that the determinant of a unitary operator has absolute value 1.*

Solution. Let T be a unitary operator on a finite-dimensional vector space V. Then $T^*T = I$. Let B be an ordered orthonormal basis for V. Let A be the matrix of T relative to B. Then

$$\det T = \det A.$$

Also A^* will be the matrix of T^* with respect to B.

Now $T^*T = I$

$$\Rightarrow [T^*T]_B = [I]_B \Rightarrow [T^*]_B\,[T]_B = I \Rightarrow A^*A = I$$
$$\Rightarrow \det (A^*A) = \det I \Rightarrow (\det A^*).(\det A) = 1$$
$$\Rightarrow \overline{(\det A)}\,(\det A) = 1 \Rightarrow |\det A|^2 = 1 \Rightarrow |\det A| = 1$$
$$\Rightarrow \det A \text{ has absolute value 1}$$
$$\Rightarrow \det T \text{ has absolute value 1}.$$

Example 3. *For which values of a are the following matrices isometric :*

(*i*) $\begin{bmatrix} a & \frac{1}{2} \\ -\frac{1}{2} & a \end{bmatrix}$ (*ii*) $\begin{bmatrix} a & 0 \\ 1 & 1 \end{bmatrix}$

Solution. (*i*) Let $A = \begin{bmatrix} a & \frac{1}{2} \\ -\frac{1}{2} & a \end{bmatrix}$.

Then $A^* = \begin{bmatrix} \bar{a} & -\frac{1}{2} \\ \frac{1}{2} & \bar{a} \end{bmatrix}$ = the conjugate transpose of A.

Now A will be isometric if $A^*A = I$ i.e. if

$$\begin{bmatrix} \bar{a} & -\frac{1}{2} \\ \frac{1}{2} & \bar{a} \end{bmatrix}\begin{bmatrix} a & \frac{1}{2} \\ -\frac{1}{2} & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \bar{a}a + \frac{1}{4} & \frac{\bar{a}}{2} - \frac{a}{2} \\ \frac{a}{2} - \frac{\bar{a}}{2} & \frac{1}{4} + \bar{a}a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\bar{a}a + \frac{1}{4} = 1,\ \frac{\bar{a}}{2} - \frac{a}{2} = 0$$

$$\frac{a}{2} - \frac{\bar{a}}{2} = 0,\ \frac{1}{4} + a\bar{a} = 1.$$

From these equations, we get $a = \bar{a}$. Therefore a must be real.

Then we get $a^2 = \frac{3}{4}$. This gives $a = \pm \frac{\sqrt{3}}{2}$.

$\therefore$ A is isometric if $a = \pm \frac{\sqrt{3}}{2}$.

(ii) Proceed as in part (i).

Example 4. *Show that the following three conditions on a linear operator T on an inner product space V are equivalent :*

(*i*) $T^*T = I$,

(*ii*) $(T\alpha, T\beta) = (\alpha, \beta)$ *for all* α *and* β,

(*iii*) $\| T\alpha \| = \| \alpha \|$ *for all* α. (Meerut 1978, 91)

Solution (i) $\Rightarrow$ (ii). It is given that $T^*T = I$. We have

$$(T\alpha, T\beta) = (\alpha, T^*T\beta)$$
$$= (\alpha, I\beta) = (\alpha, \beta) \text{ for all } \alpha \text{ and } \beta.$$

(ii) $\Rightarrow$ (iii). It is given that $(T\alpha, T\beta) = (\alpha, \beta)$ for all α and β. Taking $\beta = \alpha$, we get

$$(T\alpha, T\alpha) = (\alpha, \alpha)$$
$$\Rightarrow \| T\alpha \|^2 = \| \alpha \|^2$$
$$\Rightarrow \| T\alpha \| = \| \alpha \| \text{ for all } \alpha.$$

(iii) $\Rightarrow$ (i). It is given that $\| T\alpha \| = \| \alpha \|$ for all α. Therefore

$$(T\alpha, T\alpha) = (\alpha, \alpha)$$
$$\Rightarrow (T^*T\alpha, \alpha) = (\alpha, \alpha) \qquad \text{...(1)}$$
$$\Rightarrow ((T^*T - I)\,\alpha, \alpha) = 0 \text{ for all } \alpha.$$

Now $(T^*T - I)^* = (T^*T)^* - I^* = T^*T - I$.

$\therefore$ $T^*T - I$ is self-adjoint.

Hence from (1), we get $(T^*T) - I = \hat{0}$ i.e. $T^*T = I$.

Example 5. *If $B_1 = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an orthonormal basis of an n-dimensional unitary space V and if $B_2 = \{\beta_1, \ldots, \beta_n\}$ is a second*

Scanned with CamScanner

basis of V, then the basis B_2 is orthonormal if and only if the transition matrix from the basis B_1 to the basis B_2 is unitary.

Solution. Let $P=[p_{ij}]_{n\times n}$ be the transition matrix from the basis B_1 to the basis B_2. Then

$$\beta_j=p_{1j}\alpha_1+p_{2j}\alpha_2+\ldots+p_{nj}\alpha_n,\ j=1,\ 2,\ \ldots,\ n.$$

We have

$$(\beta_i,\ \beta_j)=(p_{1i}\alpha_1+p_{2i}\alpha_2+\ldots+p_{ni}\alpha_n,\ p_{1j}\alpha_1+p_{2j}\alpha_2+\ldots+p_{nj}\alpha_n)$$

$$=p_{1i}\bar{p}_{1j}+p_{2i}\bar{p}_{2j}+\ldots+p_{ni}\bar{p}_{nj},$$

since $(\alpha_i,\ \alpha_j)=\delta_{ij}$, B_1 being an orthonormal basis.

Now suppose B_2 is an orthonormal basis. Then $(\beta_i,\ \beta_j)=\delta_{ij}$.

$\therefore\quad p_{1i}\bar{p}_{1j}+\ldots+p_{ni}\bar{p}_{nj}=\delta_{ij}$

$\Rightarrow P^*P=[\delta_{ij}]_{n\times n}=$ unit matrix

$\Rightarrow P$ is a unitary matrix.

Conversely suppose that P is a unitary matrix. Then

$P^*P=$ unit matrix

$\Rightarrow p_{1i}\bar{p}_{1j}+\ldots+p_{ni}\bar{p}_{nj}=\delta_{ij}$

$\Rightarrow (\beta_i,\ \beta_j)=\delta_{ij}$

$\Rightarrow B_2$ is an orthonormal basis.

Example 6. *Let B and B' be two ordered orthonormal bases for a finite dimensional complex inner product space V. Prove that for each linear operator T on V, the matrix $[T]_{B'}$ is unitarily equivalent to the matrix $[T]_B$.*

Solution. Let P be the transition matrix from the basis B to the basis B'. Since B and B' are orthonormal bases, therefore P is a unitary matrix.

Therefore

$$P^*P=I$$

$$\Rightarrow P^*=P^{-1}.$$

Now $[T]_{B'}=P^{-1}\ [T]_B\ P=P^*[T]_B\ P.$

$\therefore\quad [T]_{B'}$ is unitarily equivalent to the matrix $[T]_B$.

Exercises

1. Let V be any Euclidean vector space of dimension n with inner product (,). Then V is inner product space isomorphic to V_n (R) with standard inner product.
2. Prove that in an n-dimensional Euclidean space V the dot product of the coordinate vectors of any two vectors of V is invariant under transformation from one orthonormal basis to another.

3. If $B_1=\{\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n\}$ is an orthonormal basis of an n-dimensional Euclidean space V and if $B_2=\{\beta_1,\ \ldots,\ \beta_n\}$ is a second basis of V, then the basis B_2 is orthonormal if and only if the transition matrix from the basis B_1 to the basis B_2 is orthogonal.
4. If an isometric matrix is triangular, then it is diagonal. **(Meerut 1969)**
5. If P and Q are orthogonal matrices, then PQ, P^T, and P^{-1} are orthogonal and det $P=\pm 1$.
6. If P and Q are unitary matrices, then so is PQ.
7. Let B and B' be two ordered orthonormal bases for a finite dimensional real inner product space V. Prove that for each linear operator T on V, the matrix $[T]_{B'}$ is orthogonally equivalent to the matrix $[T]_B$.
8. Prove that the relation of being 'unitarily similar' is an equivalence relation in the set of all $n\times n$ complex matrices.
9. Fill up the blanks in the following statements:
   (i) Transition Matrices expressing a change of orthonormal bases are...
   (ii) Matrices of a linear transformation expressing a change of orthonormal bases are...

Ans. (i) unitary, (ii) unitarily equivalent.

§ 7. Normal Operators. Definition. *Let T be a linear operator on an inner product space V. Then T is said to be* **normal** *if it commutes with its adjoint i.e. if $TT^*=T^*T$.* **(Meerut 1972, 87)**

If V is finite-dimensional, then T^* will definitely exist. If V is not finite-dimensional, then the above definition will make sense only if T possesses adjoint.

Note 1. *Every self-adjoint operator is normal.*

Suppose T is a self-adjoint operator i.e. $T^*=T$. Then obviously $T^*T=TT^*$. Therefore T is normal.

Note 2. *Every unitary operator is normal.*

Suppose T is a unitary operator. Then the adjoint T^* of T exists and we have $T^*T=TT^*=I$. Therefore T is normal.

Theorem 1. *Let T be a normal operator on an inner product space V. Then a necessary and sufficient condition that α be a characteristic vector of T^* is that it be a characteristic vector of T.* **(Meerut 1977, 78, 81, 84P, 88, 91)**

Scanned with CamScanner

Proof. Suppose T is a normal operator on an inner product space V. Then $TT^*=T^*T$. If α is any vector in V, then we have

$$\| T\alpha \|^2=(T\alpha, T\alpha)=(\alpha, T^*T\alpha)=(\alpha, TT^*\alpha)$$
$$=(T^*\alpha, T^*\alpha)=\| T^*\alpha \|^2.$$

$\therefore$ If T is normal and if α is any vector in V, then

$$\| T\alpha \|=\| T^*\alpha \|. \quad \ldots(1)$$

Further if c is any scalar, then

$$(T-cI)^*=T^*-\bar{c}I^*=T^*-\bar{c}I.$$

We shall show that $T-cI$ is normal.

We have $(T-cI)(T-cI)^*=(T-cI)(T^*-\bar{c}I)$

$$=TT^*-\bar{c}T-cT^*+c\bar{c}I.$$

Also $(T-cI)^*(T-cI)=(T^*-\bar{c}I)(T-cI)$

$$=T^*T-cT^*-\bar{c}T+\bar{c}cI.$$

Since $T^*T=TT^*$, therefore

$$(T-cI)(T-cI)^*=(T-cI)^*(T-cI).$$

Thus $T-cI$ is normal. Therefore from (1), we have

$$\| (T-cI)\alpha \|=\| (T-cI)^*\alpha \| \ \forall\ \alpha\in V$$

i.e. $$\| (T-cI)\alpha \|=\| (T^*-\bar{c}I)\alpha \| \ \forall\ \alpha\in V. \quad \ldots(2)$$

From (2), we conclude that

$$(T-cI)\alpha=0 \text{ iff } (T^*-\bar{c}I)\alpha=0$$

i.e. $$T\alpha=c\alpha \text{ iff } T^*\alpha=\bar{c}\alpha.$$

Thus α is a characteristic vector of T with characteristic value c iff it is a characteristic vector of T^* with characteristic value $\bar{c}$.

Theorem 2. *If T is a normal operator on an inner product space V, then the characteristic vectors for T belonging to distinct characteristic values are orthogonal.*

Proof. Suppose T is a normal operator on an inner product space V. Let α, β be the characteristic vectors for T corresponding to the characteristic values c_1 and c_2 where $c_1\neq c_2$. Then $T\alpha=c_1\alpha$ and $T\beta=c_2\beta$. Also $T^*\beta=\bar{c}_2\beta$. We have

$$c_1(\alpha, \beta)=(c_1\alpha, \beta)=(T\alpha, \beta)=(\alpha, T^*\beta)=(\alpha, \bar{c}_2\beta)=c_2(\alpha, \beta).$$

$\therefore$ $(c_1-c_2)(\alpha, \beta)=0$

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

$\Rightarrow \alpha, \beta$ are orthogonal. $[\because\ c_1\neq c_2]$

Corollary. *Characteristic spaces of a normal operator are pair-wise orthogonal.* **(Meerut 1975, 78)**

Proof. Let W_1, W_2 be characteristic spaces of a normal operator T corresponding to the distinct characteristic values c_1 and c_2. Then to prove that W_1 is orthogonal to W_2. Let $\alpha\in W_1, \beta\in W_2$. Then $T\alpha=c_1\alpha$, $T\beta=c_2\beta$. By theorem 2, $(\alpha, \beta)=0$. Thus every vector $\alpha\in W_1$ is orthogonal to every vector β belonging to W_2. Therefore W_1 is orthogonal to W_2.

Theorem 3 *Let V be a finite-dimensional complex inner product space, and let T be a normal operator on V. Then V has an orthonormal basis B, each vector of which is a characteristic vector for T and consequently the matrix of T with respect to B is a diagonal matrix.*

Proof. Since T is a linear operator on a finite dimensional complex inner product space V, therefore T must have a characteristic value and so T must have a characteristic vector.

Let $0\neq\alpha$ be a characteristic vector for T. Let $\alpha_1=\dfrac{\alpha}{\|\alpha\|}$. Then α_1 is also a characteristic vector for T and $\|\alpha_1\|=1$. If dim $V=1$, then $\{\alpha_1\}$ is an orthonormal basis for V and α_1 is a characteristic vector for T. Thus the theorem is true if dim $V=1$. Now we proceed by induction on the dimension of V. Suppose the theorem is true for inner product spaces of dimension less than dim V. Then we shall prove that it is true for V and the proof will be complete by induction.

Let W be the one-dimensional subspace of V spanned by the characteristic vector α_1 for T. Let α_1 be the characteristic vector corresponding to the characteristic value c. Then $T\alpha_1=c\alpha_1$. If β is any vector in W, then $\beta=k\alpha_1$ where k is some scalar. We have $T\beta=Tk\alpha_1=kT\alpha_1=k(c\alpha_1)=(kc)\alpha_1$. Therefore $T\beta\in W$. Thus W is invariant under T. Therefore $W^\perp$ is invariant under T^*. Now T is normal. Therefore if α_1 is a characteristic vector of T, then α_1 is also a characteristic vector of T^*. Therefore by the same argument as above W is also invariant under T^*. So $W^\perp$ is invariant under $(T^*)^*$ i e. $W^\perp$ is invariant under T. If dim $V=n$, then dim $W^\perp=$dim $V-$dim $W=n-1$. Therefore $W^\perp$ with the inner product from V is a complex inner product space of dimension one less than the dimension of V.

Suppose U is the linear operator induced by T on $W^\perp$ i.e U is the restriction of T to $W^\perp$. Then $U\gamma=T\gamma\ \forall\ \gamma\in W^\perp$. The restriction of T^* to $W^\perp$ will be the adjoint U^* of U. Now U is normal operator on $W^\perp$. For if γ is any vector in $W^\perp$, then

$$(UU^*)\,\gamma=U(U^*\gamma)=U(T^*\gamma)=T(T^*\gamma)=(TT^*)\,\gamma$$
$$=(T^*T)\,\gamma=T^*\,(T\gamma)=T^*(U\gamma)=U^*(U\gamma)=(U^*U)\,\gamma.$$

$\therefore\ UU^*=U^*U$ and thus U is a normal operator on $W^\perp$ whose dimension is less than dimension of V.

Therefore by our induction hypothesis, $W^\perp$ has an orthonormal basis $\{\alpha_2,\ldots,\alpha_n\}$ consisting of characteristic vectors for U. Suppose α_i is the characteristic vector for U corresponding to the characteristic value c_i. Then $U\alpha_i=c_i\alpha_i \Rightarrow T\alpha_i=c_i\alpha_i$. Therefore α_i is also a characteristic vector for T. Thus $\alpha_2,\ldots,\alpha_n$ are also characteristic vectors for T. Since $V=W\oplus W^\perp$, therefore $B=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ is an orthonormal basis for T each vector of which is a characteristic vector for T. The matrix of T relative to B will be a diagonal matrix.

Hence the theorem.

Normal matrix Definition. *A complex $n\times n$ matrix A is said to be normal if*

$$AA^*=A^*A.$$

If D is a diagonal matrix, then obviously

$$DD^*=D^*D.$$

Therefore every diagonal matrix is necessarily a normal matrix.

Theorem 4. *Let A be an $n\times n$ matrix with complex entries. There exists a unitary matrix P such that P^*AP is diagonal if and only if*

$$AA^*=A^*A.$$

In other words, A is unitarily equivalent to a diagonal matrix if and only if A is normal.

Proof. Let V be the vector space $\mathbf{C}^n$, with the standard inner product. Let B denote the standard ordered basis for V and let T be the linear operator on V which is represented in the standard ordered basis by the matrix A. Then

$$[T]_B=A.$$

Also $\quad [T^*]_B=A^*.$

We have

$$[TT^*]_B=[T]_B\,[T^*]_B=AA^*$$

and $\quad [T^*T]_B=A^*A.$

If A is a normal matrix, then

$$AA^*=A^*A.$$

$$\therefore\quad [TT^*]_B=[T^*T]_B$$
$$\Rightarrow\quad TT^*=T^*T \Rightarrow T \text{ is normal.}$$

Since T is a normal operator on a finite dimensional complex inner product space V, therefore there exists an orthonormal basis, say B', for V each vector of which is a characteristic vector for T. Consequently $[T]_{B'}$ will be a diagonal matrix. Now let P be the transition matrix from the basis B to the basis B'. Since B and B' are orthonormal bases, therefore P is a unitary matrix. [Note that the standard ordered basis is an orthonormal basis]. Now P is a unitary matrix implies that $P^*P=I$ and therefore

$$P^*=P^{-1}.$$

We have $[T]_{B'}=P^{-1}\,[T]_B\,P=P^*\,AP.$

Thus there exists a unitary matrix P such that

$$P^*AP=[T]_{B'}=\text{a diagonal matrix.}$$

Hence A is unitarily similar to a diagonal matrix.

Conversely suppose that A is unitarily similar to a diagonal matrix. Then there exists a unitary matrix P such that $P^*AP=D$ where D is a diagonal matrix. Since P is unitary, therefore $P^*=P^{-1}$. Therefore, we have

$$P^{-1}AP=D$$
$$\Rightarrow\quad A=PDP^{-1}.$$

We have $\quad AA^*=(PDP^{-1})\,(PDP^{-1})^*$

$=(PDP^*)\,(PDP^*)^*$ $\quad [\because\ P^{-1}=P^*]$

$=PDP^*\,(P^*)^*\,D^*P^*=PDP^*\,PD^*P^*$

$=PDID^*P^*$ $\quad [\because\ P^*P=I]$

$=PD^*DP^*$ $\quad [\because\ D$ is diagonal $\Rightarrow D$ is normal$]$

$=PD^*P^*PDP^*$ $\quad [\because\ P^*P=I]$

$=(PD^*P^*)\,(PDP^*)=(PDP^*)^*\,(PDP^*)=(PDP^{-1})^*\,(PDP^{-1})$

$=A^*A.$

$\therefore\quad A$ is normal.

Solved Examples

Example 1. *Let T be a normal operator on an inner product space V. If c is a scalar, prove that cT is also normal.*

Solution. It is given that T is normal. Therefore $TT^*=T^*T$.

We have $\quad (cT)^*=\bar{c}T^*.$

Now $\quad (cT)\,(cT)^*=(cT)\,(\bar{c}T^*)=c\bar{c}\,(TT^*).$

Also $\quad (cT)^*\,(cT)=(\bar{c}T^*)\,(cT)=(\bar{c}c)\,(T^*T)$

Scanned with CamScanner

$$=(c\bar{c})\,(TT^*).$$

$$\therefore \quad (cT)\,(cT)^*=(cT)^*\,(cT).$$

Hence cT is normal.

Example 2. *If T_1, T_2 are normal operators on an inner product space with the property that either commutes with the adjoint of the other, then prove that T_1+T_2 and T_1T_2 are also normal operators.* **(Meerut 1977, 89)**

Solution. Since T_1, T_2 are normal operators, therefore

$$T_1T_1^*=T_1^*T_1 \text{ and } T_2T_2^*=T_2^*T_2.$$

Also it is given that

$$T_1T_2^*=T_2^*T_1 \text{ and } T_2T_1^*=T_1^*T_2. \quad \ldots(2)$$

To prove that T_1+T_2 is normal. T_1+T_2 will be normal if

$$(T_1+T_2)\,(T_1+T_2)^*=(T_1+T_2)^*\,(T_1+T_2).$$

We have $(T_1+T_2)\,(T_1+T_2)^*=(T_1+T_2)\,(T_1^*+T_2^*)$

$$=T_1T_1^*+T_1T_2^*+T_2T_1^*+T_2T_2^*$$

$$=T_1^*T_1+T_2^*T_1+T_1^*T_2+T_2^*T_2 \quad \text{[from (1) and (2)]}$$

$$=T_1^*\,(T_1+T_2)+T_2^*(T_1+T_2)=(T_1^*+T_2^*)\,(T_1+T_2)$$

$$=(T_1+T_2)^*\,(T_1+T_2).$$

$\therefore$ T_1+T_2 is normal.

Also to prove that T_1T_2 is normal.

We have

$$(T_1T_2)\,(T_1T_2)^*=T_1T_2T_2^*T_1^*=T_1\,(T_2T_2^*)\,T_1^*$$

$$=T_1\,(T_2^*T_2)\,T_1^* \quad \text{[by (1)]}$$

$$=(T_1T_2^*)\,(T_2T_1^*)$$

$$=(T_2^*T_1)\,(T_1^*T_2) \quad \text{[by (2)]}$$

$$=T_2^*\,(T_1T_1^*)\,T_2$$

$$=T_2^*\,(T_1^*T_1)\,T_2 \quad \text{[from (1)]}$$

$$=(T_2^*T_1^*)\,(T_1T_2)=(T_1T_2)^*\,(T_1T_2).$$

$\therefore$ T_1T_2 is normal.

Example 3. *Let T be a linear operator on a finite-dimensional complex inner product space V. If $\| T\alpha \|=\| T^*\alpha \|$ for all α in V, then T is normal.* **(Meerut 1973, 90)**

Solution. If $\alpha \in V$, then we have

$$\| T\alpha \|=\| T^*\alpha \| \Rightarrow \| T\alpha \|^2=\| T^*\alpha \|^2$$

$$\Rightarrow (T\alpha, T\alpha)=(T^*\alpha, T^*\alpha) \Rightarrow (T^*T\alpha, \alpha)=(TT^*\alpha, \alpha)$$

$$\Rightarrow (T^*T\alpha, \alpha)-(TT^*\alpha, \alpha)=0 \quad ((T^*T-TT^*)\,\alpha, \alpha)=0.$$

Thus if $\| T\alpha \|=\| T^*\alpha \|$ for all α, then $((T^*T-TT^*)\,\alpha, \alpha)=0$

for all α. Since V is a complex inner product space, therefore from this we get $T^*T-TT^*=\hat{0}$

$$\Rightarrow T^*T=TT^* \Rightarrow T \text{ is normal.}$$

Example 4. *Let T be a normal operator on an inner product space V. If $\alpha \in V$, then $T\alpha=0 \Leftrightarrow T^*\alpha=0$.*

Solution. Since T is a normal operator on an inner product space V, therefore $\| T\alpha \|=\| T^*\alpha \|$ for all α in V.

Here $T\alpha=0 \Leftrightarrow \| T\alpha \|=0 \Leftrightarrow \| T^*\alpha \|=0 \Leftrightarrow T^*\alpha=0$.

Example 5. *Let T be a linear operator on a finite-dimensional complex inner product space. Prove that T is normal if and only if its real and imaginary parts commute.* **(Meerut 1974)**

Solution. Let $T=T_1+iT_2$. Then $T_1^*=T_1$ and $T_2^*=T_2$.

Suppose $T_1T_2=T_2T_1$. Then to prove that T is normal.

We have $\quad T^*=(T_1+iT_2)^*=T_1^*+\bar{i}\,T_2^*=T_1-iT_2.$

$$\therefore \quad TT^*=(T_1+iT_2)\,(T_1-iT_2)=T_1^2-iT_1T_2+iT_2T_1+T_2^2$$

$$=T_1^2+T_2^2 \quad [\because \; T_1T_2=T_2T_1]$$

Also $T^*T=(T_1-iT_2)(T_1+iT_2)=T_1^2+iT_1T_2-iT_2T_1+T_2^2$

$$=T_1^2+T_2^2.$$

$\therefore$ $TT^*=T^*T$. Hence T is normal.

Conversely, suppose that T is normal.

Then $\quad TT^*=T^*T$

$$\Rightarrow T_1^2-iT_1T_2+iT_2T_1+T_2^2=T_1^2+iT_1T_2-iT_2T_1+T_2^2$$

$$\Rightarrow 2i\,(T_1T_2-T_2T_1)=\hat{0}$$

$$\Rightarrow T_1T_2-T_2T_1=\hat{0} \quad [\because \; 2i \neq 0]$$

$$\Rightarrow T_1T_2=T_2T_1.$$

Example 6. *If T is an arbitrary linear operator on a finite-dimensional complex inner product space, and a and b are complex numbers such that $|a|=|b|=1$, then $aT+bT^*$ is normal.*

Solution. We have $(aT+bT^*)^*=(aT)^*+(bT^*)^*=\bar{a}T^*+\bar{b}T.$

Now $\quad (aT+bT^*)\,(aT+bT^*)^*=(aT+bT^*)\,(\bar{a}T^*+\bar{b}T)$

$$=a\bar{a}TT^*+a\bar{b}T^2+b\bar{a}\,(T^*)^2+b\bar{b}T^*T$$

$$=|a|^2\,TT^*+a\bar{b}T^2+b\bar{a}\,(T^*)^2+|b|^2\,T^*T$$

$$=TT^*+a\bar{b}T^2+b\bar{a}\,(T^*)^2+T^*T.$$

Also $(aT+bT^*)^*\,(aT+bT^*)=(\bar{a}T^*+\bar{b}T)\,(aT+bT^*)$

$$=|a|^2\,T^*T+\bar{a}b\,(T^*)^2+\bar{b}aT^2+|b|^2\,TT^*$$

$$=T^*T+\bar{a}b\,(T^*)^2+\bar{b}aT^2+TT^*$$

Scanned with CamScanner

$=TT^*+a\bar{b}T^2+b\bar{a}\,(T^*)^2+T^*T.$

$\therefore\quad (aT+bT^*)\,(aT+bT^*)^*=(aT+bT^*)^*\,(aT+bT^*).$

Hence $aT+bT^*$ is normal.

Example 7. *If T is normal, c is a characteristic value of T, and W is the characteristic space of c i.e. W is the set of all solutions of $T\alpha=c\alpha$, then both W and $W^\perp$ are invariant under T.*

Solution. Let $\alpha \in W$. Then $T\alpha=c\alpha \in W$.

$\therefore$ W is invariant under T.

Now in order to prove that $W^\perp$ is invariant under T, we shall first prove that W is invariant under T^*.

Let $\alpha \in W$. Then $T\alpha=c\alpha$. We have

$$T\,(T^*\alpha)=(TT^*)\,\alpha=(T^*T)\,\alpha$$
$$=T^*\,(T\alpha)=T^*\,(c\alpha)=c\,(T^*\alpha).$$

Now $T(T^*\alpha)=c\,(T^*\alpha)$ implies that $T^*\alpha$ is in W.

Thus $\alpha \in W \Rightarrow T^*\alpha \in W$. Therefore W is invariant under T^*. Hence $W^\perp$ is invariant under $(T^*)^*$ *i.e.*, T.

Example 8 *Suppose T is a linear operator on a finite-dimensional inner product space V and suppose there exists an orthonormal basis $B=\{\alpha_1,\ldots,\alpha_n\}$ for V such that each vector in B is a characteristic vector for T. Then prove that T is normal.*

Solution. If $\alpha_i \in B$, then it is given that α_i is a characteristic vector for T. So let

$$T\alpha_i=c_i\alpha_i,\ i=1,\ldots,n.$$

Then $[T^*]_B$ is a diagonal matrix with diagonal elements $c_1,\ldots,c_n$. Since $[T^*]_B=[T]_B{}^*$, therefore $[T^*]_B$ is also a diagonal matrix with diagonal elements $\bar{c}_1,\ldots,\bar{c}_n$. Now two diagonal matrices commute. Therefore

$$[T]_B\,[T^*]_B=[T^*]_B\,[T]_B$$
$$\Rightarrow\quad [TT^*]_B=[T^*T]_B$$
$$\Rightarrow\quad TT^*=T^*T \Rightarrow T \text{ is normal.}$$

Example 9. *Let T be a normal operator and α a vector such that $T^2\alpha=0$. Then $T\alpha=0$.*

Hence show that the range and null space of a normal operator are disjoint.

Solution Let α be a vector such that $T^2\alpha=0$.

Let $T\alpha=\beta$. Then to prove that $\beta=0$.

We have $T\alpha=\beta$

$$\Rightarrow T\,(T\alpha)=T\beta \Rightarrow T^2\alpha=T\beta \Rightarrow 0=T\beta.$$

Now T is normal. Therefore $\|T\beta\|=\|T^*\beta\|$ Hence
$$T\beta=0 \Rightarrow T^*\beta=0.$$

Now $(\beta,\beta)=(\beta,T\alpha)=(T^*\beta,\alpha)=(0,\alpha)=0.$

$\therefore\quad \beta=0$ i.e. $T\alpha=0$.

Second Part. Let $R(T)$ denote the range of T and $N(T)$ denote the null space of T. Let $\alpha \in R(T) \cap N(T)$. Then $\alpha \in R(T)$ and $\alpha \in N(T)$. Since $\alpha \in N(T)$, therefore $T\alpha=0$. Also $\alpha \in R(T)$ $\Rightarrow \alpha=T\beta$ for some vector β. Now

$$T\beta=\alpha \Rightarrow T(T\beta)=T\alpha$$
$$\Rightarrow T^2\beta=T\alpha=0.$$

But T is a normal operator, therefore by first part

$$T^2\beta=0 \Rightarrow T\beta=0$$
$$\Rightarrow \alpha=0.$$

Thus $\alpha \in R(T) \cap N(T) \Rightarrow \alpha=0$ Therefore $R(T) \cap N(T)=\{0\}$ *i.e.* $R(T)$ and $N(T)$ are disjoint.

Example 10. *Let T be a normal operator on a finite-dimensional complex inner product space V and f a polynomial with complex coefficients. Then the operator $f(T)$ is normal.* **(Meerut 1972)**

Solution. Let $f=a_0+a_1x+\ldots+a_mx^m$; then
$$f(T)=a_0I+a_1T+\ldots+a_mT^m.$$

We know that $(cT)^*=\bar{c}T^*$. Also if k is any positive integer, then $(T^k)^*=(TTT\ldots\text{upto } k \text{ times})^*$
$$=T^*T^*\ldots\text{upto } k \text{ times}=T^{*k}.$$

We have $[f(T)]^*=[a_0I+a_1T+\ldots+a_mT^m]^*$
$$=\bar{a}_0I^*+\bar{a}_1T^*+\ldots+\bar{a}_m(T^m)^*=\bar{a}_0I+\bar{a}_1T^*+\ldots+\bar{a}_mT^{*m}.$$

Now $f(T)$ will be a normal operator if we prove that
$$[f(T)]^*\,f(T)=f(T)\,[f(T)]^*.$$

Since T is normal, therefore $TT^*=T^*T$. First we shall prove that if p is any positive integer, then $TT^{*p}=T^{*p}T$. We shall prove it by induction on p. Obviously the result is true when $p=1$. Assume as our induction hypothesis that the result is true for $p=k$ *i.e.* $TT^{*k}=T^{*k}T$. Then

$$TT^{*k+1}=(TT^*)\,T^{*k}=(T^*T)\,T^{*k}=T^*\,(TT^{*k})$$
$$=T^*\,(T^{*k}T),\quad \text{since the result is true for } p=k$$
$$=T^{*k+1}T.$$

Thus the result is true for $p=k+1$ if it is true for $p=k$. Hence, it is true for all positive integers p.

Now we shall prove that if p is any fixed positive integer, then $T^{*p}\,T^q=T^q\,T^{*p}$ for all positive integers q. Obviously this

Scanned with CamScanner

result is true for $q=1$, as we have just proved it. Now assume that this result is true for $q=k$ *i.e.*, $T^{*p}T^k=T^kT^{*p}$. Then

$$T^{*p}T^{k+1}=(T^{*p}T^k)\,T=(T^kT^{*p})\,T=T^k\,(T^{*p}T)$$
$$=T^k\,(TT^{*p})=T^{k+1}T^{*p}.$$

Thus the result is true for $q=k+1$ if it is true for $q=k$. Hence it is true for all positive intergers q.

So far we have proved that $T^{*p}T^q=T^qT^{*p}$ for all positive integers p and q. Now if p is any positive integer, then

$$T^{*p}\,[f(T)]=T^{*p}\,[a_0I+a_1T+\ldots+a_mT^m]$$
$$=a_0T^{*p}I+a_1T^{*p}T+\ldots+a_mT^{*p}T^m$$
$$=a_0IT^{*p}+a_1TT^{*p}+\ldots+a_mT^mT^{*p}$$
$$=(a_0I+a_1T+\ldots+a_mT^m)\,T^{*p}=[f(T)]\,T^{*p}.$$

Now $[f(T)]^*f(T)=[\bar{a}_0I+\bar{a}_1T^*+\ldots+\bar{a}_mT^{*m}]\,f(T)$
$$=\bar{a}_0I\,f(T)+\bar{a}_1T^*\,f(T)+\ldots+\bar{a}_mT^{*m}\,f(T)$$
$$=\bar{a}_0\,f(T)\,I+\bar{a}_1\,f(T)\,T^*+\ldots+\bar{a}_m\,f(T)\,T^{*m}$$
$$=f(T)\,[\bar{a}_0I+\bar{a}_1T^*+\ldots+\bar{a}_mT^{*m}]=f(T)\,[f(T)]^*.$$

Hence $f(T)$ is normal.

Example 11. *Show that the minimal polynomial of a normal operator on a finite-dimensional inner product space has distinct roots.* (Meerut 1976, 80)

Solution. Let $p(x)$ be the minimal polynomial of a normal operator T on a finite dimensional inner product space V. Then $p(x)$ is the monic polynomial of lowest degree that annihilates T *i.e.*, for which $p(T)=\hat{0}$. We shall show that $p(x)=(x-c_1)\ldots(x-c_k)$ where $c_1,\ldots,c_k$ are distinct complex numbers. Suppose $c_1,\ldots,c_k$ are not all distinct, *i.e*, suppose some root c of $p(x)$ is repeated. Then $p(x)=(x-c)^2\,g(x)$ for some polynomial $g(x)$.

Now $p(T)=\hat{0}\Rightarrow(T-cI)^2\,g(T)=\hat{0}$
$$\Rightarrow(T-cI)^2\,g(T)\,\alpha=0,\ \forall\ \alpha\in V.$$

Let us set $U=T-cI$. Since the operator T is normal and $x-c$ is a polynomial with complex coefficients, therefore $T-cI=U$ is a normal operator. [Refer Ex. 10 above]

Now let α be any vector in V and let $\beta=g(T)\,\alpha$. Then
$$U^2\beta=U^2g(T)\,\alpha=(T-cI)^2\,g(T)\,\alpha=0.$$

Since the operator U is normal, therefore

$U^2\beta=0\Rightarrow U\beta=0$ [See Ex. 9]
$$\Rightarrow(T-cI)\,\beta=0$$
$$\Rightarrow(T-cI)\,g(T)\,\alpha=0,\ \forall\ \alpha\in V$$
$$\Rightarrow(T-cI)\,g(T)=\hat{0}.$$

Thus the operator T annihilates the monic polynomial $(x-c)g(x)$ and the degree of this polynomial is less than that of $p(x)$. Therefore it contradicts the assumption that $p(x)$ is the minimal polynomial for T. Hence no root c of $p(x)$ is repeated. Therefore $p(x)$ has distinct roots.

§ 8. Characterization of Spectra.

Theorem 1. *Let T be a self-adjoint linear operator on an inner product space V. Then each characteristic value of T is real. Also if T is positive, or non-negative, then every characteristic value of T is positive, or non-negative, respectively.*

Proof. Suppose c is a characteristic value of T. Then $T\alpha=c\alpha$ for some non-zero vector α. We have

$$c\,(\alpha,\alpha)=(c\alpha,\alpha)=(T\alpha,\alpha)=(\alpha,T^*\alpha)$$
$$=(\alpha,T\alpha)\qquad[\because\ T^*=T]$$
$$=(\alpha,c\alpha)=\bar{c}\,(\alpha,\alpha).$$
$$\therefore\quad(c-\bar{c})\,(\alpha,\alpha)=0$$
$$\Rightarrow\quad c-\bar{c}=0\qquad[\because\ \alpha\neq0\Rightarrow(\alpha,\alpha)\neq0]$$
$$\Rightarrow\quad c=\bar{c}\Rightarrow c\text{ is real.}$$

Also $(T\alpha,\alpha)=(c\alpha,\alpha)=c\,(\alpha,\alpha)=c\,\|\alpha\|^2$.
$$\therefore\quad c=\frac{(T\alpha,\alpha)}{\|\alpha\|^2}.$$

If T is positive, then $(T\alpha,\alpha)>0$. Therefore $c>0$ *i.e.* c is positive.

If T is non-negative, then $(T\alpha,\alpha)\geqslant0$. Therefore $c\geqslant0$ *i.e.*, c is non-negative.

Theorem 2. *Every characteristic value of an isometry has absolute value one.*

Proof. Let T be an isometry and let c be a characteristic value of T. Then $T\alpha=c\alpha$ for some non-zero vector α.

Since T is an isometry, therefore
$$\|\alpha\|=\|T\alpha\|=\|c\alpha\|=|c|\,.\,\|\alpha\|.$$
$$\because\quad\|\alpha\|\neq0,\text{ therefore }|c|=1.$$

Theorem 3. *Let T be either self-adjoint or isometric, then characteristic vectors of T belonging to distinct characteristic values are orthogonal.* (Meerut 1968, 69)

Proof. Let T be a linear operator on an inner product space V. Let c_1 and c_2 be two distinct characteristic values of T. Let α.

Scanned with CamScanner

β be characteristic vectors for T corresponding to the characteristic values c_1 and c_2 respectively. Then

$$T\alpha=c_1\alpha,\ T\beta=c_2\beta.$$

Case I. If T is self adjoint, then $T^*=T$. Also by theorem 1 both c_1 and c_2 are real. We have

$$c_1(\alpha,\beta)=(c_1\alpha,\beta)=(T\alpha,\beta)=(\alpha,T^*\beta)$$
$$=(\alpha,T\beta)=(\alpha,c_2\beta)=\bar{c}_2(\alpha,\beta)=c_2(\alpha,\beta).$$

$\therefore\ (c_1-c_2)(\alpha,\beta)=0$

$\Rightarrow (\alpha,\beta)=0$ $\qquad [\because\ c_1\neq c_2]$

$\Rightarrow \alpha$ and β are orthogonal.

Case II. If T is an isometry, then $(\alpha,\beta)=(T\alpha,T\beta)$. Also by theorem 2, $|c_2|=1$ *i.e.*, $|c_2|^2=1$ *i.e.*, $c_2\bar{c}_2=1$ *i.e.*, $\bar{c}_2=1/c_2$. We have

$$(\alpha,\beta)=(T\alpha,T\beta)=(c_1\alpha,c_2\beta)=c_1\bar{c}_2(\alpha,\beta)=\frac{c_1}{c_2}(\alpha,\beta).$$

$\therefore\ \left(1-\dfrac{c_1}{c_2}\right)(\alpha,\beta)=0$

$\Rightarrow (\alpha,\beta)=0$ $\qquad [\because\ c_1\neq c_2]$

$\Rightarrow \alpha$ and β are orthogonal.

Theorem 4. *Every root of the characteristic equation of a self-adjoint operator on a finite-dimensional inner product space is real.*

Proof. Suppose T is a self-adjoint linear operator on a finite dimensional inner product space V. If V is a complex vector space, then every root of the characteristic equation of T is also a characteristic value of T and so is real by theorem 1.

If V is a real vector space, then it is possible to show that there exists a Hermitian operator T^+ on some complex inner product space such that T and T^+ have the same characteristic equation. Now every root of the characteristic equation of T^+ is also a characteristic value of T^+. So it must be real. Hence every root of the characteristic equation of T must also be real.

Corollary. *Every self-adjoint operator on a finite-dimensional inner product space has a characteristic value and consequently a characteristic vector.* **(Meerut 1973)**

Theorem 5. *Let V be a finite-dimensional inner product space and let T be a self-adjoint linear operator on V. Then there is an orthonormal basis B for V, each vector of which is a characteristic vector for T and consequently the matrix of T with respect to B is a diagonal matrix.*

Proof. Since T is a self-adjoint linear operator on a finite-dimensional inner product space V, therefore T must have a characteristic value and so T must have a characteristic vector.

Let $0\neq\alpha$ be a characteristic vector for T.

Let $\alpha_1=\dfrac{\alpha}{\|\alpha\|}$. Then α_1 is also a characteristic vector for T and $\|\alpha_1\|=1$. If $\dim V=1$, then $\{\alpha_1\}$ is an orthonormal basis for V and α_1 is a characteristic vector for T. Thus the theorem is true if $\dim V=1$. Now we proceed by induction on the dimension of V. Suppose the theorem is true for inner product spaces of dimension less than $\dim V$. Then we shall prove that it is true for V and the proof will be complete by induction.

Let W be the one-dimensional subspace of V spanned by the characteristic vector α_1 for T. Let α_1 be the characteristic vector corresponding to the characteristic value c. Then $T\alpha_1=c\alpha_1$. If β is any vector in W, then $\beta=k\alpha_1$ where k is some scalar. We have

$$T\beta=Tk\alpha_1=kT\alpha_1=k(c\alpha_1)=(kc)\alpha_1.$$

Therefore $T\beta\in W$. Thus W is invariant under T. Therefore $W^\perp$ is invariant under T^*. But T is self-adjoint means that $T=T^*$. Therefore $W^\perp$ is invariant under T.

If $\dim V=n$, then $\dim W^\perp=\dim V-\dim W=n-1$.

Therefore $W^\perp$, with the inner product from V, is an inner product space of dimension one less than the dimension of V.

Suppose U is the linear operator induced by T on $W^\perp$ *i.e.*, U is the restriction of T to $W^\perp$. Then $U\gamma=T\gamma\ \forall\ \gamma\in W^\perp$. Then restriction of T^* to $W^\perp$ will be the adjoint U^* of U. Now U is a self-adjoint linear operator on $W^\perp$, because if γ is any vector in $W^\perp$, then

$$U^*\gamma=T^*\gamma=T\gamma=U\gamma.$$
$$\therefore\quad U^*=U.$$

Thus U is a self-adjoint linear operator on $W^\perp$ whose dimension is less than dimension of V.

Therefore by our induction hypothesis, $W^\perp$ has an orthonormal basis $\{\alpha_2,\ldots,\alpha_n\}$ consisting of characteristic vectors for U. Suppose α_i is the characteristic vector for U corresponding to the characteristic value c_i. Then $U\alpha_i=c_i\alpha_i\Rightarrow T\alpha_i=c_i\alpha_i$. Therefore α_i is also a characteristic vector of T. Thus $\alpha_2,\ldots,\alpha_n$ are also characteristic vectors for T. Since $V=W\oplus W^\perp$, therefore $B=\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ is

Scanned with CamScanner

an orthonormal basis for V each vector of which is a characteristic vector of T. The matrix of T relative to B will be a diagonal matrix.

Theorem 6. *Let A be an $n \times n$ Hermitian matrix. Then there exists a unitary matrix P such that $P^* AP$ is a diagonal matrix.*

Proof. Let V be the vector space $\mathbf{C}^n$, with the standard inner product. Let B denote the standard ordered basis for V and let T be the linear operator on V which is represented in the standard ordered basis by the matrix A.

Then $\quad [T]_B = A.$

Also $\quad [T^*]_B = A^*.$

Since A is a Hermitian matrix, therefore $A^* = A$. Consequently $[T^*]_B = [T]_B$ Hence $T^* = T$ and thus T is a self-adjoint linear operator.

Since T is a self-adjoint linear operator on a finite dimensional complex inner product space V, therefore there exists an orthonormal basis, say B', for V each vector of which is a characteristic vector for T. Consequently $[T]_{B'}$ will be a diagonal matrix.

Now let P be the transition matrix from the basis B'. Since B and B' are orthogonal bases, therefore P is a unitary matrix. Now P is a unitary matrix implies that $P^* = P^{-1}$.

We have $[T]_{B'} = P^{-1} [T]_B P = P^* AP$.

Thus there exists a unitary matrix P such that $P^* AP = [T]_{B'} =$ a diagonal matrix.

Theorem 7. *The self-adjoint linear transformation T on a finite dimensional inner product space V is non-negative if and only if all of its characteristic values are non-negative.*

Proof. Suppose that T is non-negative *i e.* $T \geqslant 0$.

Let λ be a characteristic value of T.

Then $T(\alpha) = \lambda\alpha$ for some $\alpha \neq 0$.

We have $(T\alpha, \alpha) = (\lambda\alpha, \alpha) = \lambda (\alpha, \alpha) = \lambda \| \alpha \|^2$.

$$\therefore \quad \lambda = \frac{(T\alpha, \alpha)}{\| \alpha \|^2}.$$

Since $T \geqslant 0$, therefore $(T\alpha, \alpha) \geqslant 0$.

Also $\| \alpha \|^2 > 0$. Therefore $\lambda \geqslant 0$.

Conversely suppose that T has all its characteristic values non-negative. Since T is self-adjoint, therefore we can find an orthonormal basis $\{\beta_1, \ldots, \beta_n\}$ consisting of characteristic vectors of T. For each β_i, we have $T\beta_i = \lambda_i \beta_i$, where $\lambda_i \geqslant 0$.

Now let γ be any vector in V. Let $\gamma = a_1\beta_1 + a_2\beta_2 + \ldots + a_n\beta_n$.

Then
$$T\gamma = a_1 T\beta_1 + a_2 T\beta_2 + \ldots + a_n T\beta_n$$
$$= a_1\lambda_1\beta_1 + a_2\lambda_2\beta_2 + \ldots + a_n\lambda_n\beta_n.$$

We have $(T\gamma, \gamma) = (a_1\lambda_1\beta_1 + a_2\lambda_2\beta_2 + \ldots + a_n\lambda_n\beta_n,$
$$a_1\beta_1 + a_2\beta_2 + \ldots + a_n\beta_n)$$
$$= a_1\bar{a}_1\lambda_1 + a_2\bar{a}_2\lambda_2 + \ldots + a_n\bar{a}_n\lambda_n$$
$$[\because \{\beta_1, \ldots, \beta_n\} \text{ is an orthonormal basis}]$$
$$= | a_1 |^2 \lambda_1 + | a_2 |^2 \lambda_2 + \ldots + | a_n |^2 \lambda_n$$
$$\geqslant 0, \text{ since each } \lambda_i \geqslant 0 \text{ and } | a_i | \geqslant 0.$$

Thus $(T\gamma, \gamma) \geqslant 0 \; \forall \; \gamma \in V$.

Hence $\quad T \geqslant 0$.

Exercises

1. Let A be an $n \times n$ real symmetric matrix. Then there exists a real orthogonal matrix P such that P^*AP is a diagonal matrix. **(Meerut 1988)**
2. The self-adjoint linear operator T on a finite dimensional inner product space V is positive if and only if all its characteristic values are positive.

§ 9. Perpendicular projections or Orthogonal projections.

Definition. *Suppose V is a finite dimensional inner product space and W is a subspace of V. Then $V = W \oplus W^\perp$. Let E be the projection on W along $W^\perp$. Then E is called a* **perpendicular projection or an orthogonal projection.** **(Meerut 1775)**

Since $W^\perp$ is uniquely determined by W, therefore there is no necessity of saying that E is a perpendicular projection on W along $W^\perp$. It is sufficient to say that E is a perpendicular projection on W and we shall write $E = P_W$.

If E is the perpendicular projection on W, then W is the range of E and $W^\perp$ is the null space of E. Also W is the set of all solutions of the equation $E\alpha = \alpha$ and $W^\perp$ is the set of all solutions of the equation $E\alpha = 0$. Further $I - E$ is the projection on $W^\perp$ along W. Therefore $I - E$ is the perpendicular projection on $W^\perp$.

Theorem 1. *A linear transformation E is a perpendicular projection if and only if $E = E^2 = E^*$. Perpendicular projections are non-negative linear transformations and have the property that* **(Meerut 1975, 78)**
$$\| E\alpha \| \leqslant \| \alpha \| \textit{ for all } \alpha.$$

Proof. Suppose E is the perpendicular projection on W. Then E is the projection on W along $W^\perp$. Since E is a projection, therefore $E^2 = E$. Also
$$R(E) = \text{the range of } E = W = \{\alpha \in V : E\alpha = \alpha\}$$
and
$$N(E) = \text{the null space of } E = W^\perp = \{\alpha \in V : E\alpha = 0\}.$$

Scanned with CamScanner

Let α, β be any two vectors in V. Since $V=W \oplus W^{\perp}$, so let $\alpha=\alpha_1+\alpha_2$ and $\beta=\beta_1+\beta_2$ where $\alpha_1, \beta_1 \in W$ and $\alpha_2, \beta_2 \in W^{\perp}$. Then $E\alpha=\alpha_1$ and $E\beta=\beta_1$. We have

$$(E\alpha, \beta)=(\alpha_1, \beta)=(\alpha_1, \beta_1+\beta_2)=(\alpha_1, \beta_1)+(\alpha_1, \beta_2)$$
$$=(\alpha_1, \beta_1) \quad [\because (\alpha_1, \beta_2)=0 \text{ because } \alpha_1 \in W \text{ and } \beta_2 \in W^{\perp}]$$
$$=(\alpha_1, \beta_1)+(\alpha_2, \beta_1) \quad [\because (\alpha_2, \beta_1)=0]$$
$$=(\alpha_1+\alpha_2, \beta_1)=(\alpha, \beta_1)=(\alpha, E\beta).$$

Thus $(E\alpha, \beta)=(\alpha, E\beta)$ for all α, β in V. Therefore E is the adjoint of E *i.e.* $E=E^*$.

Conversely, suppose $E=E^2=E^*$. Since $E=E^2$, therefore E is a projection on $R(E)$ along $N(E)$. In order to show that E is a perpendicular projection, we are simply to prove that the subspaces $R(E)$ and $N(E)$ are orthogonal. Let $\alpha \in R(E)$ and $\beta \in N(E)$. Then $E\alpha=\alpha$ and $E\beta=0$. We have

$$(\alpha, \beta)=(E\alpha, \beta)=(\alpha, E^*\beta)=(\alpha, E\beta)=(\alpha, 0)=0.$$

Thus $R(E)$ and $N(E)$ are orthogonal and consequently E is a perpendicular projection.

Now to show that if E is a perpendicular projection *i.e.* if $E=E^2=E^*$, then E is non-negative. We have

$$(E\alpha, \alpha)=(E^2\alpha, \alpha)=(EE\alpha, \alpha)=(E\alpha, E^*\alpha)=(E\alpha, E\alpha)=\| E\alpha \|^2.$$

Since $\| E\alpha \|^2 \geqslant 0$ for all α in V, therefore

$$(E\alpha, \alpha) \geqslant 0 \text{ for all } \alpha \text{ in } V.$$

Hence E is non-negative.

Now $I-E$ is also perpendicular projection. Therefore $I-E$ is non-negative. So for all α in V, we have

$$((I-E)\alpha, \alpha) \geqslant 0$$
$$\Rightarrow (\alpha-E\alpha, \alpha) \geqslant 0$$
$$\Rightarrow (\alpha, \alpha)-(E\alpha, \alpha) \geqslant 0$$
$$\Rightarrow \| \alpha \|^2-\| E\alpha \|^2 \geqslant 0 \quad [\because (E\alpha, \alpha)=\| E\alpha \|^2]$$
$$\Rightarrow \| \alpha \|^2 \geqslant \| E\alpha \|^2$$
$$\Rightarrow \| E\alpha \| \leqslant \| \alpha \|.$$

Hence the theorem.

Note. If E is a perpendicular projection, then $(E\alpha, \alpha)=\| E\alpha \|^2$ for all α.

Theorem 2. *If a linear transformation E is such that $E=E^2$ and $\| E\alpha \| \leqslant \| \alpha \|$ for all α, then $E=E^*$.* **(Meerut 1968)**

Proof. Let $R=R(E)$ and $N=N(E)$ be the range and the null space of E respectively. Since $E^2=E$, therefore E is the projection

on R along N. We shall show that E is the perpendicular projection on R. Then by theorem 1, we shall have $E=E^*$.

Now E will be perpendicular projection on R if R and N are orthogonal. For this we shall prove that $R=N^{\perp}$.

Let $\alpha \in N^{\perp}$. Put $\beta=E\alpha-\alpha$.

We have $E\beta=E(E\alpha-\alpha)=E^2\alpha-E\alpha=E\alpha-E\alpha=0$.

$$\therefore \beta \in N.$$

Now $\alpha \in N^{\perp}$ and $\beta \in N \Rightarrow (\alpha, \beta)=0$.

We have $\| E\alpha \|^2=(E\alpha, E\alpha)$

$$=(\alpha+\beta, \alpha+\beta) \quad [\because \beta=E\alpha-\alpha]$$
$$=(\alpha, \alpha)+(\alpha, \beta)+(\beta, \alpha)+(\beta, \beta)$$
$$=\| \alpha \|^2+\| \beta \|^2 \quad [\because (\alpha, \beta)=0 \text{ and so } (\beta, \alpha)=0]$$

Now it is given that $\| E\alpha \|^2 \leqslant \| \alpha \|^2$.

$$\therefore \| \alpha \|^2+\| \beta \|^2 \leqslant \| \alpha \|^2$$

or $\| \beta \|^2 \leqslant 0$

or $\| \beta \| \leqslant 0$

or $\| \beta \|=0$ $\quad [\because \| \beta \|$ cannot be negative$]$

or $\beta=0$.

Putting $\beta=0$ in the relation $\beta=E\alpha-\alpha$, we get

$$E\alpha=\alpha$$
$$\Rightarrow \alpha \in \text{the range of } E$$
$$\Rightarrow \alpha \in R.$$

Thus $\alpha \in N^{\perp} \Rightarrow \alpha \in R$. So $N^{\perp} \subseteq R$.

Conversely, let $\gamma \in R$. Then $E\gamma=\gamma$.

Since $V=N^{\perp} \oplus N$, therefore let $\gamma=\gamma_1+\gamma_2$ where $\gamma_1 \in N^{\perp}$ and $\gamma_2 \in N$. Since $\gamma_2 \in N$, therefore $E\gamma_2=0$. Also $\gamma_1 \in N^{\perp} \Rightarrow \gamma_1 \in R$ because we have just proved that $N^{\perp} \subseteq R$. Now $\gamma_1 \in R \Rightarrow E\gamma_1=\gamma_1$.

We have

$$\gamma=E\gamma=E(\gamma_1+\gamma_2)=E\gamma_1+E\gamma_2=\gamma_1+0=\gamma_1.$$

Since $\gamma_1 \in N^{\perp}$, therefore $\gamma \in N^{\perp}$.

Thus $\gamma \in R \Rightarrow \gamma \in N^{\perp}$. So $R \subseteq N^{\perp}$.

Hence $R=N^{\perp}$.

Definition. *Two perpendicular projections E and F are said to be orthogonal if $EF=\hat{0}$.*

Also if E, F are perpendicular projections, then

$$EF=\hat{0} \Leftrightarrow (EF)^*=(\hat{0})^* \Leftrightarrow F^*E^*=\hat{0} \Leftrightarrow FE=\hat{0}.$$

Scanned with CamScanner

Theorem 3. *Let E and F be two perpendicular projections on W_1 and W_2 respectively. Then E and F are orthogonal if and only if the subspaces W_1 and W_2 (that is, the ranges of E and F) are orthogonal.*

Proof. Let $EF=\hat{0}$ and let $\alpha \in W_1$ and $\beta \in W_2$. Since W_1 is the range of E, therefore $E\alpha=\alpha$. Also W_2 is the range of F. Therefore $F\beta=\beta$. We have

$$(\alpha, \beta)=(E\alpha, F\beta)=(\alpha, E^*F\beta)$$
$$=(\alpha, EF\beta) \qquad [\because \ E^*=E]$$
$$=(\alpha, \hat{0}\beta) \qquad [\because \ EF=\hat{0}]$$
$$=(\alpha, 0)=0.$$

$\therefore$ W_1 and W_2 are orthogonal.

Conversely, let W_1 and W_2 be orthogonal. If β is any vector in W_2, then β is orthogonal to every vector in W_1. So $\beta \in W_1^{\perp}$. Consequently $W_2 \subseteq W_1^{\perp}$.

Now let γ be any vector in V. Then $F\gamma$ is in the range of F i.e. $F\gamma$ is in W_2. Consequently $F\gamma$ is in $W_1^{\perp}$ which is the null space of E. Therefore $E(F\gamma)=0$.

Thus $EF\gamma=0$ for all γ in V. Hence $EF=\hat{0}$.

Theorem 4. *If $E_1, \ldots, E_n$ are perpendicular projections on the subspaces $W_1, \ldots, W_n$ respectively of an inner product space V then a necessary and sufficient condition that $E=E_1+\ldots+E_n$ be a perpendicular projection is that $E_i E_j=\hat{0}$, whenever $i \neq j$ (that is, that the E_i be pairwise orthogonal). Also then E is the perpendicular projection on $W=W_1+\ldots+W_n$.*

Proof. It is given that $E_1, \ldots, E_n$ are perpendicular projections. Therefore $E_i^2=E_i=E_i^*$ for each $i=1, \ldots, n$.

Let $E=E_1+\ldots+E_n$.

Then $E^*=(E_1+\ldots+E_n)^*=E_1^*+\ldots+E_n^*=E_1+\ldots+E_n=E.$

The condition is sufficient. Suppose that $E_iE_j=\hat{0}$ whenever $i \neq j$. Then to prove that E is a perpendicular projection.

We have $E^2=EE=(E_1+\ldots+E_n)(E_0+\ldots+E_n)$

$$=E_1^2+\ldots+E_n^2 \qquad [\because \ E_iE_j=\hat{0} \text{ if } i\neq j]$$
$$=E_1+\ldots+E_n=E.$$

Thus $E^2=E=E^*$. Therefore E is a perpendicular projection.

The condition is necessary. Suppose E is a perpendicular projection. Then $E^2=E=E^*$.

We are to prove that $E_iE_j=\hat{0}$ if $i \neq j$.

Let α belong to the range of some E_i. Then $E_i\alpha=\alpha$.

Therefore $\| \alpha \|^2=\| E_i\alpha \|^2$. ...(1)

Since E is a perpendicular projection, therefore by theorem 1, we get

$$\| \alpha \|^2 \geqslant \| E\alpha \|^2. \qquad ...(2)$$

From (1) and (2), we get

$$\| E_i\alpha \|^2 \geqslant \| E\alpha \|^2. \qquad ...(3)$$

Now $\| E\alpha \|^2=(E\alpha, \alpha)$ [See Note below theorem 1]

$$=\left(\left(\sum_{j=1}^{n} E_j\right)\alpha, \alpha\right)=\left(\sum_{j=1}^{n} E_j\,\alpha, \alpha\right)=\sum_{j=1}^{n}(E_j\,\alpha, \alpha)$$

$$=\sum_{j=1}^{n} \| E_j\alpha \|^2 \qquad [\because \ E_j \text{ is perpendicular projection}]$$

$$\Rightarrow \| E_j\alpha \|^2=(E_j\alpha, \alpha). \qquad \text{[See Note below theorem 1]}$$

$\therefore$ From (3), we get

$$\| E_i\alpha \|^2 \geqslant \sum_{j=1}^{n} \| E_j\alpha \|^2. \qquad ...(4)$$

But $$\| E_i\alpha \|^2 \leqslant \sum_{j=1}^{n} \| E_j\alpha \|^2. \qquad ...(5)$$

[$\because$ $\| E_i\alpha \|^2$ is one of the n terms in R.H.S. of (5)]

From (4) and (5), we get

$$\| E_i\alpha \|^2=\sum_{j=1}^{n} \| E_j\alpha \|^2$$

$\Rightarrow \| E_j\alpha \|^2=0$ if $j\neq i \Rightarrow \| E_j\alpha \|=0$ if $j \neq i$

$\Rightarrow E_j\alpha=0$ if $j\neq i \Rightarrow \alpha$ is in the null space of E_j if $j \neq i$

$\Rightarrow \alpha \in W_j^{\perp}$ if $j \neq i$

$\Rightarrow \alpha$ is orthogonal to the range W_j of every E_j with $j \neq i$.

Thus every vector α in the range of E_i $(i=1, \ldots, n)$ is orthogonal to the range of every E_j with $j \neq i$. Therefore the range of E_i is orthogonal to the range of every E_j with $j \neq i$. Hence by theorem 3, we have $E_iE_j=\hat{0}$ whenever $i \neq j$.

Scanned with CamScanner

Finally in order to show that E is the perpendicular projection on $W=\sum_{i=1}^{n} W_i$, we are to show that $R(E)=W$ where $R(E)$ is the range of E.

Let $\alpha \in R(E)$. Then

$\alpha=E\alpha=\left(\sum_{i=1}^{n} E_i\right)\alpha=\sum_{i=1}^{n} E_i\,\alpha \in W$ since $E_i\alpha \in W_i$ (i.e. the range of E_i) for each i.

$\therefore\ R(E) \subseteq W.$

Conversely, let $\alpha \in W$. Then $\alpha=\sum_{j=1}^{n} \alpha_j$ where $\alpha_j \in W_j$.

We have $E\alpha=\left(\sum_{i=1}^{n} E_i\right)\alpha=\sum_{i=1}^{n} E_i\alpha=\sum_{j=1}^{n}\sum_{i=1}^{n} E_i\alpha_j$

$=\sum_{j=1}^{n} E_j\alpha_j$ [$\because\ \alpha_j \in W_j \Rightarrow \alpha_j \in$ the null space of E_i if $i \neq j$]

$=\sum_{j=1}^{n} \alpha_j$ [$\because\ \alpha_j$ is in the range of E_j]

$=\alpha.$

$\therefore\ \alpha$ is in the range of E.

Hence $W \subseteq R(E)$.

$\therefore\ W=R(E)$.

This completes the proof of the theorem.

§ 10. The Spectral Theorem.

Theorem 1. (Spectral theorem for a normal operator). *To every normal operator T on a finite-dimensional complex inner product space V there correspond distinct complex numbers $c_1, \ldots, c_k$ and perpendicular projections $E_1, \ldots, E_k$ (where k is a strictly positive integer, and greater than the dimension of the space) so that*

(1) *the E_i are pairwise orthogonal and different from $\hat{0}$,*

(2) $E_1+\ldots+E_k=I$,

(3) $T=c_1E_1+\ldots+c_kE_k$.

(Meerut 1976, 78, 80, 83 P, 84, 91, 93, 93 P)

Proof. Suppose T is a normal operator on a finite-dimensional

complex inner product space V. Then T will definitely possess a characteristic value. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T. If dim $V=n$, then we must have $k \leqslant n$. Let $W_1, \ldots, W_k$ be the characteristic subspaces of the characteristic values $c_1, \ldots, c_k$ respectively. Then

$$W_i=\{\alpha \in V : T\alpha=c_i\alpha\},\ 1 \leqslant i \leqslant k.$$

Let $E_1, \ldots, E_k$ be the perpendicular projections on $W_1, \ldots, W_k$ respectively. Then W_i is the range of E_i for each $i=1, \ldots, k$.

(1) To show that $E_i \neq 0$ where $i=1, \ldots, k$. Let $0 \neq \alpha$ be a characteristic vector for T corresponding to the characteristic value c_i. Then $T\alpha=c_i\alpha$. Therefore $\alpha \in W_i$ which is the range of the perpendicular projection E_i So $E_i\alpha=\alpha$. Since $\alpha \neq 0$, therefore $E_i \neq \hat{0}$.

Further to show that the E_i are pairwise orthogonal. We know that if T is a normal operator, then the characteristic vectors of T corresponding to distinct characteristic values are orthogonal. Therefore each vector in W_i is orthogonal to each vector in W_j if $i \neq j$. Therefore the subspaces W_i and W_j are orthogonal. Hence by theorem 3 of § 9, we have $E_i E_j=\hat{0}$ if $i \neq j$.

(2) To prove that $E_1+\ldots+E_k=I$.

Let $E=E_1+\ldots+E_k$ and $W=W_1+\ldots+W_k$. Then by theorem 4 of § 9, E is a perpendicular projection on W and consequently $I-E$ is a perpendicular projection having $W^{\perp}$ as its range.

First we shall prove that $W^{\perp}$ is invariant under T. For this it is sufficient to show that W is invariant under T^*. Let $\alpha=\sum_{i=1}^{k} \alpha_i \in W$ where $\alpha_i \in W_i$ for each i. Then $T\alpha_i=c_i\alpha_i$ and $T^*\alpha_i=\bar{c}_i\alpha_i$ because T is normal. Therefore $T^*\alpha=\sum_{i=1}^{k} T^*\alpha_i=\sum_{i=1}^{k} \bar{c}_i\alpha_i \in W$ since $\bar{c}_i\alpha_i \in W_i$ for each i. Thus W is invariant under T^* and consequently $W^{\perp}$ is invariant under T^{**} i.e. T.

Let U be the restriction of T to $W^{\perp}$. Then $U\alpha=T\alpha$ for all α in $W^{\perp}$.

Suppose $E \neq I$ Then $I-E \neq \hat{0}$. Therefore the range $W^{\perp}$ of the perpendicular projection $I-E$ is not equal to the zero subspace of V. Since U is a linear operator on a finite dimensional non-zero complex space $W^{\perp}$, therefore U must have a characteristic

Scanned with CamScanner

value and consequently a characteristic vector. Suppose α is a characteristic vector for U corresponding to the characteristic value c. Then $U\alpha=c\alpha$. Therefore $T\alpha=c\alpha$ and thus α is also a characteristic vector for T. Therefore each characteristic vector for U is also a characteristic vector for T. But T has no characteristic vector in $W^{\perp}$ since all the characteristic vectors for T are in W. So U has no characteristic vector and thus we get a contradiction. Therefore we must have $E=I$ *i.e.*

$$E_1+\ldots+E_k=I.$$

(3) To prove that $T=c_1E_1+\ldots+c_kE_k$.

Let β be any vector in V. Let $E_i\beta=\beta_i$, $i=1, \ldots, k$. Then β_i is in W_i which is the range of E_i. Therefore by the definition of W_i, we have $T\beta_i=c_i\beta_i$.

$$\text{Now } T\beta=T(I\beta)=T\left(\sum_{i=1}^{k} E_i\beta\right)=T\sum_{i=1}^{k}\beta_i=\sum_{i=1}^{k}T\beta_i$$

$$=\sum_{i=1}^{k}c_i\beta_i=\sum_{i=1}^{k}c_iE_i\beta=\left(\sum_{i=1}^{k}c_iE_i\right)\beta.$$

$$\therefore \quad T=\sum_{i=1}^{k}c_iE_i=c_1E_1+\ldots+c_kE_k.$$

Note. We shall call the decomposition $T=c_1E_1+\ldots+c_kE_k$ the **Spectral resolution** of T.

Theorem 2. **(Spectral theorem for a self-adjoint operator).** *To every self-adjoint operator T on a finite-dimensional inner product space V there correspond distinct real numbers $c_1, \ldots, c_k$ and perpendicular projections $E_1, \ldots, E_k$, (where k is a strictly positive integer, not greater than the dimension of the space) so that*

(1) *the E_i are pairwise orthogonal and different from $\hat{0}$,*

(2) $E_1+\ldots+E_k=I$,

(3) $T=c_1E_1+\ldots+c_kE_k$. **(Meerut 1969)**

Proof. Suppose T is a self-adjoint operator on a finite-dimensional inner product space V. Then T will definitely possess a characteristic value and all the characteristic values of T will be real. Let $c_1, \ldots, c_k$ be the distinct characteristic values of T. If dim $V=n$, then we must have $k\leq n$. Let $W_1, \ldots, W_k$ be the characteristic subspaces of the characteristic values $c_1, \ldots, c_k$ respectively. Then $W_i=\{\alpha\in V : T\alpha=c_i\alpha\}$. Let $E_1, \ldots, E_k$ be the

perpendicular projections on $W_1, \ldots, W_k$ respectively. Then W_i is the range of E_i for each $i=1, \ldots, k$.

(1) To show that $E_i\neq\hat{0}$. Let $0\neq\alpha$ be a characteristic vector for T corresponding to the characteristic value c_i. Then $T\alpha=c_i\alpha$. Therefore $\alpha\in W_i$. So $E_i\alpha=\alpha$. Since $\alpha\neq 0$, therefore $E_i\neq\hat{0}$.

Further to show that the E_i are pairwise orthogonal. We know that if T is a self-adjoint operator, then the characteristic vectors of T belonging to distinct characteristic values are orthogonal. Therefore each vector in W_i is orthogonal to each vector in W_j if $i\neq j$. Therefore the subspaces W_i and W_j are orthogonal. But W_i and W_j are the ranges of E_i and E_j respectively. Therefore $E_iE_j=\hat{0}$ if $i\neq j$.

(2) To prove that $I=E_1+\ldots+E_k$.

Let $E=E_1+\ldots+E_k$ and $W=W_1+\ldots+W_k$. Then E is a perpendicular projection on W and consequently $I-E$ is a perpendicular projection having $W^{\perp}$ as its range.

First we shall prove that $W^{\perp}$ is invariant under T. Let $\alpha=\sum_{i=1}^{k}\alpha_i\in W$ where $\alpha_i\in W_i$ for each i. Then $T\alpha_i=c_i\alpha_i$. Therefore

$$T\alpha=T\sum_{i=1}^{k}\alpha_i=\sum_{i=1}^{k}T\alpha_i=\sum_{i=1}^{k}c_i\alpha_i\in W \text{ since } c_i\alpha_i\in W_i \text{ for each } i.$$

Thus W is invariant under T and consequently $W^{\perp}$ is invariant under $T^*=T$, T being self-adjoint. Let U be the restriction of T to $W^{\perp}$. Then $U\alpha=T\alpha$ for all α in $W^{\perp}$. Also U will be self-adjoint.

Suppose $E\neq I$ *i.e.*, $I-E\neq\hat{0}$. Then the range $W^{\perp}$ of $I-E$ is not equal to the zero subspace of V. Since U is a self-adjoint operator on a finite-dimensional non-zero space $W^{\perp}$, therefore U must have a characteristic vector. Obviously each characteristic vector for U is also a characteristic vector for T. But T has no characteristic vector in $W^{\perp}$ and so U has no characteristic vector. Thus we get a contradiction. Hence we must have $E=I$ *i.e.*,

$$E_1+\ldots+E_k=I.$$

(3) To prove that $T=c_1E_1+\ldots+c_kE_k$. Give the same proof here as we have given in the corresponding part of the spectral theorem for a normal operator.

Theorem 3. *Let T be a normal operator on a finite-dimensional complex inner product space V. Then V has an orthonormal basis B consisting of characteristic vectors for T. Consequently the matrix of T relative to B is a diagonal matrix.*

Proof. Let $T=c_1E_1+...+c_kE_k$ be the spectral resolution of T. Then $c_1,...,c_k$ are all the distinct characteristic values of T. Also $E_1,...,E_k$ are the perpendicular projections on $W_1,...,W_k$ which are the characteristic subspaces of the characteristic values $c_1,...,c_k$ respectively.

Thus W_i is the set of all vectors α which satisfy the equation

$$T\alpha=c_i\alpha.$$

Also $$I=E_1+...+E_k$$

and $$E_iE_j=\hat{0} \text{ for } i\neq j.$$

For each i, the subspace W_i is the range of E_i. For $i\neq j$, the subspaces W_i and W_j are orthogonal. For if $\alpha\in W_i$, $\beta\in W_j$, then

$$(\alpha,\beta)=(E_i\alpha, E_j\beta) \quad [\because \alpha\in \text{the range of } E_i \Rightarrow E_i\alpha=\alpha, \text{ etc.}]$$
$$=(\alpha, E_i^*E_j\beta)$$
$$=(\alpha, E_iE_j\beta) \quad [\because E_i^*=E_i, E_i \text{ being perpendicular projection}]$$
$$=(\alpha, \hat{0}\beta)=(\alpha, 0)=0.$$

Now let $B_1,...,B_k$ be orthonormal bases for the subspaces $W_1,...,W_k$ respectively. Then we claim that $B=\bigcup B_i$ *i.e.*, the union of the B_i is an orthonormal basis for V. Obviously B is an orthonormal set because each B_i is an orthonormal set and any vector in B_i is orthogonal to any vector in B_j, if $i\neq j$. Note that the vectors in B_i are some elements of W_i and the vectors in B_j are some elements of W_j. The subspaces W_i and W_j are orthogonal if $i\neq j$.

Since B is an orthonormal set, therefore B is linearly independent.

Now B will be a basis for V if we prove that B generates V. Let γ be any vector in V. Then

$$\gamma=I\gamma=(E_1+...+E_k)\,\gamma=E_1\gamma+...+E_k\gamma$$
$$=\alpha_1+...+\alpha_k, \text{ where } \alpha_i=E_i\gamma.$$

Since $E_i\gamma$ is in the range of E_i, therefore α_i is in W_i. So for each i the vector α_i can be expressed as a linear combination of the vectors in B_i which is a basis for W_i. Therefore γ can be expressed as a linear combination of the vectors in B. Hence V is generated by B. Therefore B is an orthonormal basis for V. Since each non-zero vector in W_i is a characteristic vector for T, therefore each vector in B_i is a characteristic vector for T. Consequently each vector in B is a characteristic vector for T. Thus there exists an orthonormal basis B for V such that each vector in B is a characteristic vector for T. Consequently the matrix of T relative to B will be a diagonal matrix.

Theorem 4. *Let T be a self-adjoint operator on a finite-dimensional inner product space V. Then V has an orthonormal basis B consisting of characteristic vectors for T.*

Proof. Give the same proof as in theorem 3.

Theorem 5. *Let T be a normal operator on the finite-dimensional complex inner product space V. Let $c_1,...,c_k$ be distinct complex numbers and $E_1,...,E_k$ be non-zero linear operators on V such that*

(a) $T=c_1E_1+...+c_kE_k$,

(b) $I=E_1+...+E_k$,

(c) $E_iE_j=\hat{0}$, *if* $i\neq j$.

Then $c_1,...,c_k$ are precisely the distinct characteristic values of T. Also for each i, E_i is a polynomial in T and is the perpendicular projection of V on the characteristic space of the characteristic value c_i. In short the decomposition of T given in (a) is the spectral resolution of T.

Proof. (i) First we shall prove that for each i, E_i is a projection.

We have $I=E_1+...+E_k$

$$\Rightarrow E_iI=E_i(E_1+...+E_k)$$
$$\Rightarrow E_i=E_iE_1+...+E_iE_k$$
$$\Rightarrow E_i=E_i^2 \qquad [\because E_iE_j=\hat{0}, \text{ if } i\neq j]$$
$$\Rightarrow E_i \text{ is a projection.}$$

(ii) Now we shall prove that $c_1,...,c_k$ are precisely the distinct characteristic values of T.

First we shall show that for each i, c_i is a characteristic value of T.

Since $E_i\neq\hat{0}$, therefore there exists a non-zero vector α in the range of E_i.

Since E_i is a projection, therefore $E_i\alpha=\alpha$.

Scanned with CamScanner

Now $T\alpha=(c_1E_1+\ldots+c_kE_k)\,\alpha$

$=(c_1E_1+\ldots+c_kE_k)\,E_i\alpha \quad [\because \alpha=E_i\alpha]$

$=c_1E_1E_i\alpha+\ldots+c_kE_kE_i\alpha$

$=c_iE_i^2\alpha \quad [\because E_iE_j=\hat{0}, \text{ if } i\neq j]$

$=c_iE_i\alpha \quad [\because E_i^2=E_i]$

$=c_i\alpha.$

$\therefore$ c_i is a characteristic value of T.

Since T is a linear operator on a finite-dimensional complex inner product space, therefore T must possess a characteristic value. Suppose c is a characteristic value of T. Then there exists a non-zero vector α such that

$$T\alpha=c\alpha$$

$$\Rightarrow T\alpha=cI\alpha$$

$$\Rightarrow (c_1E_1+\ldots+c_kE_k)\,\alpha=c\,(E_1+\ldots+E_k)\,\alpha$$

$$\Rightarrow (c_1-c)\,E_1\alpha+\ldots+(c_k-c)\,E_k\alpha=\mathbf{0}.$$

Operating on this with E_i and remembering that $E_i^2=E_i$ and $E_iE_j=\hat{0}$ if $i\neq j$, we obtain

$$(c_i-c)\,E_i\alpha=\mathbf{0} \text{ for } i=1,\ldots,k.$$

If $c_i\neq c$ for each i, then we have $E_i\alpha=\mathbf{0}$ for each i. Therefore

$$E_1\alpha+\ldots+E_k\alpha=\mathbf{0}$$

$$\Rightarrow (E_1+\ldots+E_k)\,\alpha=\mathbf{0}$$

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

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

This contradicts the fact that $\alpha\neq\mathbf{0}$. Hence c must be equal to c_i for some i.

(iii) Now we shall prove that E_i is a projection on W_i where W_i is the characteristic space of the characteristic value c_i. For this we shall prove that the range of E_i is W_i.

Remember that α is in the range of E_i iff $E_i\alpha=\alpha$ and α is in W_i iff $T\alpha=c_i\alpha$.

Let $\alpha\in W_i$. Then

$$T\alpha=c_i\alpha=c_iI\alpha=c_i\,(E_1+\ldots+E_k)\,\alpha$$

$$\Rightarrow (c_1E_1+\ldots+c_kE_k)\,\alpha=(c_iE_1+\ldots+c_iE_k)\,\alpha$$

$$\Rightarrow (c_1-c_i)\,E_1\alpha+\ldots+(c_k-c_i)\,E_k\alpha=\mathbf{0}.$$

Operating with E_j, we get

$$(c_1-c_i)\,E_jE_1\alpha+\ldots+(c_k-c_j)\,E_jE_k\alpha=\mathbf{0}$$

$$\Rightarrow (c_j-c_i)\,E_j^2\alpha=\mathbf{0}$$

$$\Rightarrow (c_j-c_i)\,E_j\alpha=\mathbf{0}$$

$$\Rightarrow E_j\alpha=\mathbf{0} \text{ if } j\neq i \quad [\because c_1,\ldots,c_k \text{ are all distinct}]$$

Then $\alpha=I\alpha=E_1\alpha+\ldots+E_k\alpha$

$=E_i\alpha.$

Thus $\alpha\in$ the range of E_i.

Conversely, suppose that α is in the range of E_i. Then $E_i\alpha=\alpha$. As proved in (ii), we shall get $T\alpha=c_i\alpha$. Thus $\alpha\in W_i$.

(iv) Now we shall prove that for each i, E_i is a polynomial in T.

We have $T^0=I=E_1+\ldots+E_k$.

$T=c_1E_1+\ldots+c_kE_k$.

$T^2=(c_1E_1+\ldots+c_kE_k)\,(c_1E_1+\ldots+c_kE_k)$

$=c_1^2E_1+\ldots+c_k^2E_k \quad [\because E_i^2=E_i,\ E_iE_j=\hat{0} \text{ if } i\neq j]$

Similarly $T^n=c_1^nE_1+\ldots+c_k^nE_k$ where n is any non-negative integer.

Therefore if g is any polynomial, then taking linear combinations of the above relations, we get

$$g(T)=g(c_1)\,E_1+\ldots+g(c_k)\,E_k=\sum_{j=1}^{k} g(c_j)\,E_j.$$

Now suppose p_i is a polynomial such that $p_i\,(c_j)=\delta_{ij}$. Then taking p_i in place of g, we get

$$p_i\,(T)=\sum_{j=1}^{k} p_i\,(c_j)\,E_j=\sum_{j=1}^{k}\delta_{ij}E_j=E_i.$$

Thus for each i, E_i is a polynomial in T. But we must show the existence of such a polynomial p_i over the field F. Obviously

$$p_i\,(t)=\frac{(t-c_1)\ldots(t-c_{i-1})\,(t-c_{i+1})\ldots(t-c_k)}{(c_i-c_1)\ldots(c_i-c_{i-1})\,(c_i-c_{i+1})\ldots(c_i-c_k)}$$

serves the purpose *i.e.*, $p_i\,(c_i)=1$ and $p_i\,(c_j)=0$ if $j\neq i$.

(v) Now we shall show that for each i, E_i is a perpendicular projection.

T is a normal operator. For each i, E_i is a polynomial in T. Therefore for each i, E_i is a normal operator. Since E_i is a projection, therefore we must have $E_i^*=E_i$. Hence E_i is a perpendicular projection.

{Note. If T is a self-adjoint operator on an inner product space, then also the above theorem is true. The only difference in

Scanned with CamScanner

proof will be in part (v). Since $c_1, \ldots, c_k$ are all real, therefore p_i is a polynomial with real coefficients.

We have $E_i = p_i(T)$.

Therefore $E_i^* = [p_i(T)]^* = p_i(T^*) = p_i(T) = E_i$.

Hence E_i is a perpendicular projection.

Theorem 6. *Let T be the normal operator on the finite-dimensional complex inner product space V. Then T is self-adjoint, positive, or unitary according as each characteristic value of T is real, positive, or of absolute value 1.* **(Meerut 1979)**

Proof. Let $T = c_1E_1 + \ldots + c_kE_k$ be the spectral resolution for T. Then $c_1, \ldots, c_k$ are all the distinct characteristic values for T and $E_1, \ldots, E_k$ are perpendicular projections.

Now $E_i \neq \hat{0}$ and $E_iE_j = \hat{0}$ for $i \neq j$.

(i) Suppose each characteristic value for T is real. Then to prove that T is self-adjoint.

We have
$$\begin{aligned} T^* &= (c_1E_1 + \ldots + c_kE_k)^* \\ &= \bar{c}_1E_1^* + \ldots + \bar{c}_kE_k^* \\ &= \bar{c}_1E_1 + \ldots + \bar{c}_kE_k \\ &= c_1E_1 + \ldots + c_kE_k \quad [\because \text{ each } c_i \text{ is real} \Rightarrow \bar{c}_i = c_i] \\ &= T. \end{aligned}$$

$\therefore$ T is self-adjoint.

(ii) Suppose each characteristic value for T is positive *i.e.* $c_i > 0$ for each i. Then to prove that T is positive.

Since $c_i > 0$ for each i, therefore each c_i is real. [Consequently by case (i) T is self-adjoint. Now

$$(T\alpha, \alpha) = (T\alpha, I\alpha)$$
$$= \left(\sum_{i=1}^{k} c_iE_i\alpha, \sum_{j=1}^{k} E_j\alpha\right) = \sum_{i=1}^{k}\sum_{j=1}^{k} c_i(E_i\alpha, E_j\alpha)$$
$$= \sum_{i=1}^{k}\sum_{j=1}^{k} c_i(\alpha, E_i^*E_j\alpha) = \sum_{i=1}^{k}\sum_{j=1}^{k} c_i(\alpha, E_iE_j\alpha)$$
$$= \sum_{i=1}^{k} c_i(\alpha, E_iE_i\alpha)$$ [Summing with respect to j. Note that $E_iE_j = \hat{0}, j \neq i$]
$$= \sum_{i=1}^{k} c_i(E_i^*\alpha, E_i\alpha) = \sum_{i=1}^{k} c_i(E_i\alpha, E_i\alpha)$$

$$= \sum_{i=1}^{k} c_i \| E_i\alpha \|^2.$$

Since $\| E_i\alpha \| \geqslant 0$, therefore if $c_i > 0$, then
$$(T\alpha, \alpha) \geqslant 0.$$

Also
$$\begin{aligned} (T\alpha, \alpha) = 0 &\Rightarrow \sum_{i=1}^{k} c_i \| E_i\alpha \|^2 = 0 \\ &\Rightarrow \| E_i\alpha \| = 0 \text{ for each } i \\ &\Rightarrow E_i\alpha = 0 \text{ for each } i \\ &\Rightarrow (E_1 + \ldots + E_k)\alpha = 0 \\ &\Rightarrow I\alpha = 0 \\ &\Rightarrow \alpha = 0. \end{aligned}$$

Thus T is self-adjoint. Also $(T\alpha, \alpha) \geqslant 0$ for each α and $(T\alpha, \alpha) = 0 \Rightarrow \alpha = 0$.

$\therefore$ T is positive.

(iii) Suppose each characteristic value of E is of absolute value 1. Then to show that E is unitary. We have
$$T^* = \bar{c}_1E_1 + \ldots + \bar{c}_kE_k.$$
$$\begin{aligned} \therefore \quad TT^* &= (c_1E_1 + \ldots + c_kE_k)(\bar{c}_1E_1 + \ldots + \bar{c}_kE_k) \\ &= |c_1|^2 E_1 + \ldots + |c_k|^2 E_k \quad [\because E_i^2 = E_i \text{ and } E_iE_j = \hat{0} \text{ for } i \neq j] \\ &= E_1 + \ldots + E_k \quad [\because |c_i| = 1 \text{ for each } i] \\ &= I. \end{aligned}$$

Since V is finite-dimensional, therefore $TT^* = I \Rightarrow T^*T = I$.

Hence T is unitary.

Solved Examples

Example 1. *Let T be a normal operator on the finite dimensional complex inner product space V. Then T is non-negative, invertible, or idempotent according as each characteristic value of T is non-negative, different from 0, or equal to zero or one.*

Solution Let $T = c_1E_1 + \ldots + c_kE_k$ be the spectral resolution for T.

(i) Suppose each characteristic value of T is non-negative *i.e.* $c_i \geqslant 0$ for each $i = 1, \ldots, k$. Then as in theorem 6, we have $T^* = T$. Also
$$(T\alpha, \alpha) = \sum_{i=1}^{k} c_i \| E_i\alpha \|^2.$$

Scanned with CamScanner

Since $\| E_i\alpha \| \geq 0$, therefore if $c_i \geq 0$, then $(T\alpha, \alpha) \geq 0$ for all α in V.

Hence T is non-negative.

(ii) Suppose each characteristic value of T is different from zero i.e. $c_i \neq 0$ for each i.

Consider the linear operator

$$S=\frac{1}{c_1} E_1+...+\frac{1}{c_k} E_k.$$

We have $TS=(c_1E_1+...+c_kE_k)\left(\frac{1}{c_1} E_1+...+\frac{1}{c_k} E_k\right)$

$=E_1+...+E_k$

$[\because \; E_i^2=E_i$ and $E_iE_j=\hat{0}$ for $i \neq j]$

$=I$.

$\therefore$ T is invertible.

(iii) Suppose each characteristic value of T is zero or one i.e. $c_i=0$ or 1 for each i.

Then $c_i^2=c_i$ for each $i=1,..., k$.

Now $T^2=(c_1E_1+...+c_kE_k)(c_1E_1+...+c_kE_k)$

$=c_1^2E_1+...+c_k^2E_k$

$=c_1E_1+...+c_kE_k$

$=T$.

$\therefore$ T is idempotent.

Example 2. *If $\sum_{i=1}^{n} c_iE_i$ is the spectral resolution of a normal operator T on a finite-dimensional complex inner product space V, then a necessary and sufficient condition that a linear operator S commutes with T is that it commutes with each E_i.*

Solution. **The condition is necessary.** Suppose S commutes with T. Then S commutes with every polynomial in T. Since each E_i is a polynomial in T, therefore S will commute with each E_i. Hence the condition is necessary.

The condition is sufficient. Suppose S commutes with each E_i i.e. $SE_i=E_iS$ for each i. Since $T=\sum_{i=1}^{k} c_iE_i$, therefore obviously $ST=TS$. Hence the condition is sufficient.

Note. A similar result can be proved if T is a self-adjoint linear operator on a finite-dimensional inner product space.

Example 3. *If T is a normal operator on a finite-dimensional complex inner product space V and S is an arbitrary linear operator that commutes with T, then S commutes with T^*.*

Solution. Let $T=c_1E_1+...+c_kE_k=\sum_{i=1}^{k} c_iE_i$ be the spectral resolution of T.

Then $T^*=(c_1E_1+...+c_kE_k)^*$

$=\bar{c}_1E_1^*+...+\bar{c}_kE_k^*$

$=\bar{c}_1E_1+...+\bar{c}_kE_k$ $[\because$ each E_i is self-adjoint$]$

Now for each i the operator E_i is a polynomial in T. Therefore T^* is also a polynomial in T. Since S commutes with T, therefore S will also commute with every polynomial in T. In particular S commutes with T^* which is also a polynomial in T.

Example 4. *Let T be a linear operator an a finite-dimensional complex inner product space V. Then T is normal if and only if the adjoint T^* is a polynomial in T.* **(Meerut 1971, 76, 88)**

Solution. Suppose T^* is a polynomial in T. Let $T^*=f(T)$. Then obviously $T^*T=TT^*$. Therefore T is normal.

Conversely suppose that T is normal. Let $T=c_1E_1+...+c_kE_k$ be the spectral resolution of T. Then

$T^*=(c_1E_1+...+c_kE_k)^*$

$=\bar{c}_1E_1^*+...+\bar{c}_kE_k^*$

$=\bar{c}_1E_1+...+\bar{c}_kE_k$. $[\because$ each E_i is self-adjon$]$

But we know that each E_i, where $i=1,..., k$, is a polynomial in T. Therefore T^* is also a polynomial in T.

Example 5. *Let T be a normal operator on a finite-dimensional complex inner product space V. Then every subspace of V which is invariant under T is also invariant under T^*.* **(Meerut 1973, 74, 87)**

Solution. Let W be a subspace of V which is invariant under T. We have $T^2\alpha=TT\alpha$.

Since W is invariant under T, therefore $\alpha \in W \Rightarrow T\alpha \in W$. Consequently $TT\alpha$, i.e. $T^2\alpha$ is also in W. Thus

$\alpha \in W \Rightarrow T^2\alpha \in W$.

Therefore W is invariant under T^2.

Similarly W is invariant under T^n where n is any positive integer. Consequently W is invariant under $f(T)$ where f is any polynomial.

Scanned with CamScanner

Now T^* is also a polynomial in T. Therefore W is invariant under T^*.

Example 6. *If T is a normal operator on a finite-dimensional inner product space V and if W is a subspace of V invariant under T, then the restriction of T to W is also normal.*

Solution. If W is invariant under T, then W is also invariant under T^*. Let U be the restriction of T to W and S be the restriction of T^* to W. Then

$$U\alpha=T\alpha \text{ and } S\alpha=T^*\alpha \text{ for all } \alpha \text{ in } W.$$

Obviously $S=U^*$ *i.e.* S is the adjoint of U because for all α, β in W, we have

$$(U\alpha, \beta)=(T\alpha, \beta)=(\alpha, T^*\beta)=(\alpha, S\beta).$$

$\therefore$ S is the adjoint of U.

Now let α be any vector in W. Then

$$\begin{aligned} UU^*\alpha &= UT^*\alpha && [\because U^*\alpha=T^*\alpha] \\ &= TT^*\alpha \\ &= T^*T\alpha && [\because T \text{ is normal}] \\ &= T^*U\alpha \\ &= U^*U\alpha. \end{aligned}$$

$\therefore$ $UU^*=U^*U$.

Hence U is normal.

Example 7. *Let T be a linear operator on a finite-dimensional inner product space V. Let $E_1, \dots, E_k$ be linear operators on V such that*

(1) $T=c_1E_1+\dots+c_kE_k$;

(2) $E_i=E_i^*$ and $E_iE_j=\hat{0}$ for $i\neq j$.

Then T is normal.

Solution. We have

$$\begin{aligned} T^* &= (c_1E_1+\dots+c_kE_k)^* \\ &= \bar{c}_1E_1^*+\dots+\bar{c}_kE_k^* \\ &= \bar{c}_1E_1+\dots+\bar{c}_kE_k. \end{aligned}$$

Now
$$\begin{aligned} TT^* &= (c_1E_1+\dots+c_kE_k)(\bar{c}_1E_1+\dots+\bar{c}_kE_k) \\ &= c_1\bar{c}_1E_1^2+\dots+c_k\bar{c}_kE_k^2 && [\because E_iE_j=\hat{0} \text{ for } i\neq j] \\ &= |c_1|^2E_1^2+\dots+|c_k|^2E_k^2. \end{aligned}$$

Also
$$\begin{aligned} T^*T &= (\bar{c}_1E_1+\dots+\bar{c}_kE_k)(c_1E_1+\dots+c_kE_k) \\ &= \bar{c}_1c_1E_1^2+\dots+\bar{c}_kc_kE_k^2 \end{aligned}$$

$$=|c_1|^2E_1^2+\dots+|c_k|^2E_k^2.$$

Example 8. *Let $W_1, \dots, W_k$ be subspaces of an inner product space V, and let E_i be the perpendicular projection on W_i, $i=1,\dots,k$. Then the following two statements are equivalent :*

(i) $T=W_1\oplus\dots\oplus W_k$, and this is an orthogonal direct sum i.e. the subspaces W_i and W_j are orthogonal for $i\neq j$.

(ii) $I=E_1+\dots+E_k$ and $E_iE_j=\hat{0}$ for $i\neq j$.

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

For each i the subspace W_i is the range of the perpendicular projection E_i. Since W_i and W_j are orthogonal subspaces for $i\neq j$, therefore $E_iE_j=\hat{0}$ for $i\neq j$.

Let $\alpha \in V$. Since V is the direct sum of $W_1,\dots,W_k$, therefore we can write

$$\alpha=\alpha_1+\dots+\alpha_k \text{ where } \alpha_i \in W_i,\ i=1,\dots,k.$$

Now $\alpha_i \in W_i \Rightarrow E_i\alpha_i=\alpha_i$. Therefore we have

$$\begin{aligned} &\alpha=E_1\alpha_1+\dots+E_k\alpha_k \\ \Rightarrow\ &(E_1+\dots+E_k)\alpha=(E_1+\dots+E_k)(E_1\alpha_1+\dots+E_k\alpha_k) \\ \Rightarrow\ &(E_1+\dots+E_k)\alpha=E_1^2\alpha_1+\dots+E_k^2\alpha_k && [\because E_iE_j=\hat{0} \text{ for } i\neq j] \\ \Rightarrow\ &(E_1+\dots+E_k)\alpha=E_1\alpha_1+\dots+E_k\alpha_k && [\because E_i^2=E_i] \\ \Rightarrow\ &(E_1+\dots+E_k)\alpha=\alpha=I\alpha. \end{aligned}$$

$\therefore$ $(E_1+\dots+E_k)\alpha=I\alpha$ for all α in V.

$\therefore$ $E_1+\dots+E_k=I$.

(ii) $\Rightarrow$ (i).

Let $\alpha \in V$. Then

$$\begin{aligned} \alpha &= I\alpha=(E_1+\dots+E_k)\alpha \\ &= E_1\alpha+\dots+E_k\alpha. && \dots(1) \end{aligned}$$

For each i the vector $E_i\alpha$ is in the range of E_i *i.e.* in W_i. Therefore (1) is an expression for α as a sum of vectors, one from each subspace W_i. We shall show that this expression is unique.

Let $\alpha=\alpha_1+\dots+\alpha_k$ with α_i in W_i.

Then $E_i\alpha_i=\alpha_i$. Therefore

$$\begin{aligned} &\alpha=E_1\alpha_1+\dots+E_k\alpha_k \\ \Rightarrow\ &E_i\alpha=E_iE_1\alpha_1+\dots+E_iE_k\alpha_k \\ &\quad=E_i^2\alpha_i && [\because E_iE_j=\hat{0} \text{ for } i\neq j] \\ &\quad=E_i\alpha_i=\alpha_i. \end{aligned}$$

$\therefore$ $\alpha=E_1\alpha+\dots+E_k\alpha.$

Scanned with CamScanner

Thus the expression (1) for α as a sum of vectors from the subspaces W_i is unique. Hence $V=W_1\oplus...\oplus W_k$.

Now to show that this sum is an orthogonal direct sum. Let $\alpha\in W_i, \beta\in W_j$ and $i\neq j$.

Then $(\alpha, \beta)=(E_i\alpha, E_j\beta)=(\alpha, E_i{}^*E_j\beta)=(\alpha, E_iE_j\beta)$

$$=(\alpha, \hat{0}\beta)=(\alpha, 0)=0.$$

Therefore the subspaces W_i and W_j are orthogonal for $i\neq j$.

Hence $V=W_1\oplus...\oplus W_k$ is an orthogonal direct sum.

Exercises

1. Let T be a self-adjoint operator on a finite-dimensional inner-product space V. Then V has an orthonormal basis B consisting of characteristic vectors of T.
2. If U and T are normal operators which commute, prove that $U+T$ and UT are normal. **(Meerut 1972, 80, 84)**

[**Hint.** Take help of Ex. 3 page 393, and Ex. 2 page 368].

4
Bilinear Forms

Suppose U and V are two vector spaces over the same field F. Let $W=U\times V$ *i.e.* $W=\{(\alpha, \beta) : \alpha\in U, \beta\in V\}$.

If (α_1, β_1) 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*.

§ 1. 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*

$$f(a\alpha_1+b\alpha_2, \beta)=af(\alpha_1, \beta)+bf(\alpha_2, \beta)$$

and $$f(\alpha, a\beta_1+b\beta_2)=af(\alpha, \beta_1)+bf(\alpha, \beta_2). \qquad \textbf{(Meerut 1975)}$$

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 $W=V\oplus V$, we shall simply say that f is a bilinear form on V.

Scanned with CamScanner

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 (α, β) in F, and which satisfies

$$f(a\alpha_1+b\alpha_2, \beta)=af(\alpha_1, \beta)+bf(\alpha_2, \beta)$$

and $$f(\alpha, a\beta_1+b\beta_2)=af(\alpha, \beta_1)+bf(\alpha, \beta_2).$$

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×V into F defined as*

$$f(\alpha, \beta)=L_1(\alpha)\, L_2(\beta).$$

Then f is a bilinear form on V.

If $\alpha, \beta \in V$, then $L_1(\alpha)$, $L_2(\beta)$ are scalars. We have

$$\begin{aligned} f(a\alpha_1+b\alpha_2, \beta) &= L_1(a\alpha_1+b\alpha_2)\, L_2(\beta) \\ &= [aL_1(\alpha_1)+bL_1(\alpha_2)]\, L_2(\beta) \\ &= aL_1(\alpha_1)\, L_2(\beta)+bL_1(\alpha_2)\, L_2(\beta) \\ &= af(\alpha_1, \beta)+bf(\alpha_2, \beta). \end{aligned}$$

Also $$\begin{aligned} f(\alpha, a\beta_1+b\beta_2) &= L_1(\alpha)\, L_2(a\beta_1+b\beta_2) \\ &= L_1(\alpha)\, [aL_2(\beta_1)+bL_2(\beta_2)] \\ &= aL_1(\alpha)\, L_2(\beta_1)+bL_1(\alpha)\, L_2(\beta_2) \\ &= af(\alpha, \beta_1)+bf(\alpha, \beta_2). \end{aligned}$$

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×V into F defined as*

$$g(\alpha, \beta)=f(T\alpha, T\beta).$$

Then g is a bilinear form on V.

We have $$\begin{aligned} g(a\alpha_1+b\alpha_2, \beta) &= f(T(a\alpha_1+b\alpha_2), T\beta) \\ &= f(aT\alpha_1+bT\alpha_2, T\beta) \\ &= af(T\alpha_1, T\beta)+bf(T\alpha_2, T\beta) \\ &= ag(\alpha_1, \beta)+bg(\alpha_2, \beta). \end{aligned}$$

Also $$\begin{aligned} g(\alpha, a\beta_1+b\beta_2) &= f(T\alpha, T(a\beta_1+b\beta_2)) \\ &= f(T\alpha, aT\beta_1+bT\beta_2) \\ &= af(T\alpha, T\beta_1)+bf(T\alpha, T\beta_2) \\ &= ag(\alpha, \beta_1)+bg(\alpha, \beta_2). \end{aligned}$$

Hence g is a bilinear form on V.

Example 3. *Let $V=V_n(F)$ i.e., let V be the vector space of all ordered n-tuples over the field F. If $\alpha=(a_1, \ldots, a_n)$ and $\beta=(b_1, \ldots, b_n)$ be any two elements in V, let f be a function from V×V into F defined as*

$$f(\alpha, \beta)=a_1b_1+\ldots+a_nb_n.$$

Then it can be easily seen that f is a bilinear form on V.

Example 4. *Let U and V be two vector spaces over the field F. Let $W=U\oplus V$. If $\hat{0}$ is the zero function from W into F (i.e. $\hat{0}$ maps each element of W into the zero of F), then $\hat{0}$ is a bilinear form on W.*

We have $\hat{0}(\alpha, \beta)=0 \ \forall\ (\alpha, \beta) \in W$.

Now $$\begin{aligned} \hat{0}(a\alpha_1+b\alpha_2, \beta) &= 0=0+0 \\ &= a0+b0 \\ &= a\hat{0}(\alpha_1, \beta)+b\hat{0}(\alpha_2, \beta). \end{aligned}$$

Also $$\begin{aligned} \hat{0}(\alpha, a\beta_1+b\beta_2) &= 0=0+0 \\ &= a0+b0 \\ &= a\hat{0}(\alpha, \beta_1)+b\hat{0}(\alpha, \beta_2). \end{aligned}$$

Thus $\hat{0}$ is a bilinear form on W

Example 5. *Let U and V be two vector spaces over the field F. Let f be a bilinear form on U×V. Then the function −f from U×V into F defined as*

$$(-f)(\alpha, \beta)=-f(\alpha, \beta)$$

is a bilinear form on U×V.

Solved Examples

Example. *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*) $f(\alpha, \beta)=x_1y_2-x_2y_1$ **(Meerut 1972, 77, 79, 85, 90, 91, 93 P)**

(*ii*) $f(\alpha, \beta)=(x_1-y_1)^2+x_2y_2$. **(Meerut 1978, 85, 91, 93 P)**

Solution. Let $\alpha=(x_1, x_2)$,

$\beta=(y_1, y_2)$,

and $\gamma=(z_1, z_2)$

be any three vectors in $\mathbf{R}^2$. Let $a, b \in \mathbf{R}$. Then

$$\begin{aligned} a\alpha+b\beta &= a(x_1, x_2)+b(y_1, y_2) \\ &= (ax_1+by_1, ax_2+by_2). \end{aligned}$$

Scanned with CamScanner

(*i*) By definition of f, we have

$f(\alpha, \gamma)=f((x_1, x_2), (z_1, z_2))=x_1z_2-x_2z_1,$

$f(\beta, \gamma)=y_1z_2-y_2z_1, f(\gamma, \alpha)=z_1x_2-z_2x_1,$

and $f(\gamma, \beta)=z_1y_2-z_2y_1.$

Now

$$\begin{aligned} f(a\alpha+b\beta, \gamma) &= f((ax_1+by_1, ax_2+by_2), (z_1, z_2)) \\ &= (ax_1+by_1)z_2-(ax_2+by_2)z_1 \\ &= a(x_1z_2-x_2z_1)+b(y_1z_2-y_2z_1) \\ &= af(\alpha, \gamma)+bf(\beta, \gamma). \end{aligned}$$

Also

$$\begin{aligned} f(\gamma, a\alpha+b\beta) &= f((z_1, z_2), (ax_1+by_1, ax_2+by_2)) \\ &= z_1(ax_2+by_2)-z_2(ax_1+by_1) \\ &= a(z_1x_2-z_2x_1)+b(z_1y_2-z_2y_1) \\ &= af(\gamma, \alpha)+bf(\gamma, \beta). \end{aligned}$$

Hence f is a bilinear form on $\mathbf{R}^2$.

(*ii*) By definition of f, we have

$f(\alpha, \gamma)=(x_1-z_1)^2+x_2z_2,$

and $f(\beta, \gamma)=(y_1-z_1)^2+y_2z_2.$

Now

$$\begin{aligned} f(a\alpha+b\beta, \gamma) &= f((ax_1+by_1, ax_2+by_2), (z_1, z_2)) \\ &= (ax_1+by_1-z_1)^2+(ax_2+by_2)z_2 \end{aligned}$$

Also

$$\begin{aligned} af(\alpha, \gamma)+bf(\beta, \gamma) &= a(x_1-z_1)^2+ax_2z_2+b(y_1-z_1)^2+by_2z_2 \\ &= a(x_1-z_1)^2+b(y_1-z_1)^2+(ax_2+by_2)z_2. \end{aligned}$$

Obviously $f(a\alpha+b\beta, \gamma) \neq af(\alpha, \gamma)+bf(\beta, \gamma)$. Hence f is not a bilinear form on $\mathbf{R}^2$.

§ 2. **Bilinear forms 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

$$\begin{aligned} (f+g)(a\alpha_1+b\alpha_2, \beta) &= f(a\alpha_1+b\alpha_2, \beta)+g(a\alpha_1+b\alpha_2, \beta) \\ &= [af(\alpha_1, \beta)+bf(\alpha_2, \beta)]+[ag(\alpha_1, \beta)+bg(\alpha_2, \beta)] \\ &= a[f(\alpha_1, \beta)+g(\alpha_1, \beta)]+b[f(\alpha_2, \beta)+g(\alpha_2, \beta)] \\ &= a[(f+g)(\alpha_1, \beta)]+b[(f+g)(\alpha_2, \beta)]. \end{aligned}$$

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

$$\begin{aligned} (cf)(a\alpha_1+b\alpha_2, \beta) &= cf(a\alpha_1+b\alpha_2, \beta) \\ &= c[af(\alpha_1, \beta)+bf(\alpha_2, \beta)] \\ &= caf(\alpha_1, \beta)+cbf(\alpha_2, \beta) \\ &= a[cf(\alpha_1, \beta)]+b[cf(\alpha_2, \beta)] \\ &= a[(cf)(\alpha_1, \beta)]+b[(cf)(\alpha_2, \beta)]. \end{aligned}$$

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.*

(Meerut 1975, 78, 85)

The bilinear from $\hat{0}$ 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 $\{\alpha_1, ..., \alpha_n\}$, if V is an m-dimensional vector space with basis $\{\beta_1, ..., \beta_m\}$, and if $\{a_{ij}\}$ is any set of nm scalars ($i=1, ..., n$; $j=1, ..., m$) then there is one and only one bilinear form f on $U \oplus V$ such that*

$$f(\alpha_i, \beta_j)=a_{ij} \text{ for all } i \text{ and } j. \quad \textbf{(Meerut 1969, 78, 88, 91, 93)}$$

Proof. Let $\alpha=\sum_{i=1}^{n} x_i\alpha_i \in U$ and

$$\beta=\sum_{j=1}^{m} y_j\beta_j \in V. \quad \text{Let us define a function}$$

f from $U \times V$ into F such that

$$f(\alpha, \beta) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j a_{ij}. \quad \ldots(1)$$

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 $\quad \alpha_i = 0\alpha_1 + \ldots + 0\alpha_{i-1} + 1\alpha_i + 0\alpha_{i+1} + \ldots + 0\alpha_n$

and $\quad \beta_j = 0\beta_1 + \ldots + 0\beta_{j-1} + 1\beta_j + 0\beta_{j+1} + \ldots + 0\beta_m$.

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

$$g(\alpha_i, \beta_j) = a_{ij}. \quad \ldots(2)$$

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

$$g(\alpha, \beta) = g\left(\sum_{i=1}^{n} x_i \alpha_i, \sum_{j=1}^{m} y_j \beta_j \right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j\, g(\alpha_i, \beta_j) \quad [\because \ g \text{ is a bilinear form}]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i y_j a_{ij} \quad \text{[from (2)]}$$

$$= f(\alpha, \beta). \quad \text{[from (1)]}$$

$\therefore$ By the equality of two functions, we have

$$g = f.$$

Thus f is unique.

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 $A = [a_{ij}]_{n \times n}$ such that*

$$f(\alpha_i, \alpha_j) = a_{ij},\ i = 1, \ldots, n;\ j = 1, \ldots, n.$$

We shall denote this matrix A by $[f]_B$.

Rank of a bilinear form. 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

$$f(\alpha, \beta) = f\left(\sum_{i=1}^{n} x_i \alpha_i, \sum_{j=1}^{n} y_j \alpha_j \right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j\, f(\alpha_i, \alpha_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j\, a_{ij}$$

$$= X' A Y,$$

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

$$f(\alpha, \beta) = [\alpha]'_B\, A\, [\beta]_B$$

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 $A = [a_{ij}]_{n \times n}$ over the field F such that

Scanned with CamScanner

$A=[f]_B$.

Conversely, if $A=[a_{ij}]_{n\times n}$ 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

$$f(\alpha, \beta)=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j a_{ij}=X' A Y, \qquad \ldots(2)$$

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).

Example. *Let f be the bilinear form on $V_2(\mathbf{R})$ defined by*

$$f((x_1, y_1), (x_2, y_2))=x_1y_1+x_2y_2.$$

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

$$\therefore \quad [f]_B=\begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Theorem 2. *Let V be a finite-dimensional vector space over the field F. For each ordered basis B of V, the function ψ which associates with each bilinear form on V its matrix in the ordered basis B is an isomorphism of the space $L(V, V, F)$ onto the space of $n\times n$ matrices over the field F.*

Proof. Let M denote the vector space of all $n\times n$ matrices over the field F. Let ψ be a function from $L(V, V, F)$ into M such that

$$\psi(f)=[f]_B \ \forall f\in L(V, V, F). \qquad \ldots(1)$$

ψ is a linear transformation.

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

Let $a, b\in F$ and $f, g\in L(V, V, F)$. Then

$$\psi(af+bg)=[af+bg]_B.$$

Now
$$(af+bg)(\alpha_i, \alpha_j)=(af)(\alpha_i, \alpha_j) [illegible] (\alpha_i, \alpha_j)$$
$$=af(\alpha_i, \alpha_j) [illegible]. \qquad \ldots(2)$$

The result (2) is true for each $i=1, \ldots, n$, $j=1, \ldots, n$.

Therefore $\quad [af+bg]_B=a[f]_B+b[g]_B$
$$\Rightarrow \psi(af+bg)=a\psi(f)+b\psi(g).$$

Therefore ψ is a linear transformation.

ψ is one-one.

Let $f, g\in L(V, V, F)$. Then
$$\psi(f)=\psi(g)$$
$$\Rightarrow [f]_B=[g]_B$$
$$\Rightarrow f=g. \qquad \text{[by theorem 1]}$$

$\therefore$ ψ is one-one.

ψ is onto.

Let $A=[a_{ij}]_{n\times n}$ be an element of M. Then from theorem (1), there exists a bilinear form f on V such that $[f]_B=A$

$\therefore$ ψ is onto. $\qquad \Rightarrow \psi(f)=A.$

Hence ψ is an isomorphism of $L(V, V, F)$ onto M.

Corollary. *If V is an n-dimensional vector space over the field F, then $\dim L(V, V, F)=n^2$*

Proof. If M is the vector space of all $n\times n$ matrices over the field F, then

$$L(V, V, F)\cong M.$$

Since $\dim M=n^2$, therefore $\dim L(V, V, F)=n^2$.

Theorem 3. *If U is an n-dimensional vector space with basis $\{\alpha_1, \ldots, \alpha_n\}$, and if V is an m-dimensional vector space with basis $\{\beta_1, \ldots, \beta_m\}$, then there is a uniquely determined basis*

$$\{f_{pq}\} \ (p=1, \ldots, n\ ;\ q=1, \ldots, m)$$

in the vector space of all bilinear forms on $U\oplus V$ with the property that $f_{pq}(\alpha_i, \beta_j)=\delta_{ip}\delta_{jq}$. Consequently dimension of the vector space $L(U, V, F)$ of bilinear forms on $U\oplus V$ is the product of the dimensions of U and V. **(Meerut 1992)**

Proof. By theorem 1 for each fixed p and q we determine a unique bilinear form f_{pq} on $U\oplus V$ such that

$$f_{pq}(\alpha_i, \beta_j)=\delta_{ip}\,\delta_{jq},\ i=1, \ldots, n;\ j=1, \ldots, m. \qquad \ldots(1)$$

Let B denote the set of these nm uniquely determined bilinear forms on $U\oplus V$.

First we shall show that $L(U, V, F)$ is generated by B. Let f be any bilinear form on $U\times V$ i.e. $f\in L(U, V, F)$.

Scanned with CamScanner

If $f(\alpha_i, \beta_j)=a_{ij}$, then we shall prove that

$$f=\sum_{p=1}^{n}\sum_{q=1}^{m} a_{pq} f_{pq}.$$

Let α be any element of U and β be any element of V. Let

$$\alpha=\sum_{i=1}^{n} x_i\alpha_i,\ \beta=\sum_{j=1}^{m} y_j\beta_j.$$

Then $f_{pq}(\alpha, \beta)=f_{pq}\left(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{m} y_j\beta_j\right)$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j f_{pq}(\alpha_i, \beta_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j \delta_{ip}\delta_{jq}$$

$$=\sum_{i=1}^{n} x_i y_q \delta_{ip} \qquad \text{[Summing with respect to } j]$$

$$=x_p y_q. \qquad \ldots(2)$$

Now $f(\alpha, \beta)=f\left(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{m} y_j\beta_j\right)$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j f(\alpha_i, \beta_j)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j a_{ij}$$

$$=\sum_{p=1}^{n}\sum_{q=1}^{m} a_{pq} x_p y_q$$

$$=\sum_{p=1}^{n}\sum_{q=1}^{m} a_{pq}[f_{pq}(\alpha, \beta)] \qquad \text{[From (2)]}$$

$$=\left[\sum_{p=1}^{n}\sum_{q=1}^{m} a_{pq} f_{pq}\right](\alpha, \beta).$$

$$\therefore\ f=\sum_{p=1}^{n}\sum_{q=1}^{m} a_{pq} f_{pq}.$$

Thus every element f in $L(U, V, F)$ is a linear combination of elements of B.

Now we shall prove that B is linearly independent. Let

$$\sum_{p=1}^{n}\sum_{q=1}^{m} b_{pq} f_{pq}=\hat{0}$$

$$\Rightarrow \left(\sum_{p=1}^{n}\sum_{q=1}^{m} b_{pq} f_{pq}\right)(\alpha, \beta)=\hat{0}(\alpha, \beta)$$

$\forall\ \alpha$ in U and $\forall\ \beta$ in V

$$\Rightarrow \sum_{p=1}^{n}\sum_{q=1}^{m} b_{pq} f_{pq}(\alpha, \beta)=0$$

$\forall\ \alpha$ in U and $\forall\ \beta$ in V

$$\Rightarrow \sum_{p=1}^{n}\sum_{q=1}^{m} b_{pq} f_{pq}(\alpha_i, \beta_j)=0,\ i=1, \ldots, n,\ j=1, \ldots, m$$

$$\Rightarrow \sum_{p=1}^{n}\sum_{q=1}^{m} b_{pq}\delta_{ip}\delta_{jq}=0,\ i=1, \ldots, n;\ j=1, \ldots, m$$

$$\Rightarrow \sum_{p=1}^{n} b_{pj}\delta_{ip}=0,\ i=1, \ldots, n;\ j=1,\ldots, m$$

$\Rightarrow$ $b_{ij}=0$, $i=1,\ldots, n;\ j=1,\ldots, m$
$\Rightarrow$ B is linearly independent.

$\therefore$ B is a basis for $L(U, V, F)$.

Dim $L(U, V, F)$=number of vectors in B

$=nm$.

Theorem 4. *Let V be a finite-dimensional vector space over the field F. If $B=\{\alpha_1, \ldots, \alpha_n\}$ is an ordered basis for V, and $B'=\{L_1, \ldots, L_n\}$ is the dual basis for V', then the n^2 bilinear forms*

$$f_{ij}(\alpha, \beta)=L_i(\alpha)\, L_j(\beta),\ 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n.$$

form a basis for the space $L(V, V, F)$. In particular, the dimension of $L(V, V, F)$ is n^2.

Proof. Since B' is the dual basis of B, therefore $L_i(\alpha_p)=\delta_{ip}$ and $L_j(\alpha_q)=\delta_{jq}$ for all values of i, j, p, q from 1 to n.

Scanned with CamScanner

Thus $f_{ij}(\alpha_p, \alpha_q)=L_i(\alpha_p)\, L_j(\alpha_q)=\delta_{ip}\, \delta_{jq}$.

Now proceed as in theorem 3. Simply replace U by V and $\{\beta_1, \ldots, \beta_m\}$ by $\{\alpha_1, \ldots, \alpha_n\}$.

Theorem 5. *Let V be a finite-dimensional vector space over the field F, and let*

$$B=\{\alpha_1, \ldots, \alpha_n\} \text{ and } B'=\{\beta_1, \ldots, \beta_n\}$$

be ordered bases for V. Suppose f is a bilinear form on V. Then there exists an invertible $n\times n$ matrix P over the field F such that

$$[f]_{B'}=P'\,[f]_B\, P$$

where P' denotes the transpose of the matrix P.

The matrix P is called the **transition matrix** *from the ordered basis B to the ordered basis B'.*

Proof. Let T be the linear operator on V defined as

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

Since T maps a basis onto a basis, therefore T is invertible. Let P denote the matrix of T relative to the ordered basis B. Then P is an invertible matrix. We recall that if $P=[p_{ij}]_{n\times n}$ then

$$T(\alpha_j)=\beta_j=\sum_{i=1}^{n} p_{ij}\,\alpha_i.$$

Also if α is any vector in V, then

$$[\alpha]_B=P[\alpha]_{B'}$$

[See theorem 8 page 166] ...(1)

For any vectors α, β in V, we have

$$f(\alpha, \beta)=[\alpha]'_B\,[f]_B\,[\beta]_B$$
$$=(P\,[\alpha]_{B'})'\,[f]_B\,[\beta]_B$$
$$=[\alpha]'_{B'}\,P'[f]_B\,[\beta]_B \qquad \text{[from (1)]}$$
$$[\because \text{ if } P, Q \text{ are two matrices then } (PQ)'=Q'P']$$
$$=[\alpha]'_{B'}\,P'\,[f]_B\,P[\beta]_{B'}$$

[from (1), taking β in place of α]

By the definition and uniqueness of the matrix representing f in the ordered basis B', we must have

$$[f]_{B'}=P'[f]_B\,P.$$

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 $0\neq\alpha \in V$, there is a β in V such that $f(\alpha, \beta)\neq 0$ and (ii) for each $0\neq\beta \in V$, there is an α in V such that $f(\alpha, \beta)\neq 0$.

§ 3. Symmetric bilinear forms. 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 1. *If V is a finite-dimensional vector space, then a bilinear form f on V is symmetric if and only if its matrix A in some (or every) ordered basis is symmetric, i.e., $A'=A$.* **(Meerut 1990)**

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 A is the matrix of f in the ordered basis B, then

$$f(\alpha, \beta)=X'\,A\,Y,$$

and

$$f(\beta, \alpha)=Y'\,A\,X.$$

$\therefore$ 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. $\quad A'=A$

i.e. $\quad A$ is symmetric.

Hence the theorem.

Quadratic 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*

$$q(\alpha)=f(\alpha, \alpha) \text{ for all } \alpha \text{ in } V.$$

Scanned with CamScanner

Theorem 2. *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.*

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

$$\begin{aligned} q(\alpha+\beta) &= f(\alpha+\beta, \alpha+\beta) \\ &= f(\alpha, \alpha+\beta)+f(\beta, \alpha+\beta) \\ &= f(\alpha, \alpha)+f(\alpha, \beta)+f(\beta, \alpha)+f(\beta, \beta) \\ &= q(\alpha)+f(\alpha, \beta)+f(\alpha, \beta)+q(\beta) \\ &= q(\alpha)+(1+1) f(\alpha, \beta)+q(\beta). \end{aligned}$$

$$\therefore \quad (1+1) f(\alpha, \beta)=q(\alpha+\beta)-q(\alpha)-q(\beta). \qquad \ldots(1)$$

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 detemined by its associated quadratic form according to the polarization identity

$$f(\alpha, \beta)=\tfrac{1}{4} q(\alpha+\beta)-\tfrac{1}{4} q(\alpha-\beta).$$

As in theorem 2, we have

$$2f(\alpha, \beta)=q(\alpha+\beta)-q(\alpha)-q(\beta). \qquad \ldots(1)$$

Also
$$\begin{aligned} q(\alpha-\beta) &= f(\alpha-\beta, \alpha-\beta) \\ &= f(\alpha, \alpha-\beta)-f(\beta, \alpha-\beta) \\ &= f(\alpha, \alpha)-f(\alpha, \beta)-f(\beta, \alpha)+f(\beta, \beta) \\ &= q(\alpha)+q(\beta)-2f(\alpha, \beta). \end{aligned}$$

$$\therefore \quad 2f(\alpha, \beta)=q(\alpha)+q(\beta)-q(\alpha-\beta). \qquad \ldots(2)$$

Adding (1) and (2), we get

$$4f(\alpha, \beta)=q(\alpha+\beta)-q(\alpha-\beta)$$
$$\Rightarrow f(\alpha, \beta)=\tfrac{1}{4} q(\alpha+\beta)-\tfrac{1}{4} q(\alpha-\beta).$$

Theorem 3. *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.*

(Meerut 1981, 82, 83P, 84, 87, 89; G.N.D.U. Amritsar 90)

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 $f=\hat{0}$ or $n=1$, the theorem is obviously true. So let us suppose that $f\neq\hat{0}$ 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 $f=\hat{0}$ 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 vectors β in V such that $f(\alpha_1, \beta)=0$. Obviously W_2 is a subspace of V. Now we claim that $V=W_1\oplus W_2$. We shall first prove our claim.

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
$$\begin{aligned} \gamma\in W_2 &\Rightarrow f(\alpha_1, \gamma)=0 \\ &\Rightarrow f(\alpha_1, c\alpha_1)=0 \\ &\Rightarrow cf(\alpha_1, \alpha_1)=0 \\ &\Rightarrow c=0 \qquad [\because f(\alpha_1, \alpha_1)\neq 0] \\ &\Rightarrow \gamma=0\alpha_1=0. \end{aligned}$$

$\therefore$ 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
$$\begin{aligned} f(\alpha_1, \beta) &= f\left(\alpha_1, \gamma-\frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)}\alpha_1\right) \\ &= f(\alpha_1, \gamma)-\frac{f(\gamma, \alpha_1)}{f(\alpha_1, \alpha_1)} f(\alpha_1, \alpha_1) \\ &= f(\alpha_1, \gamma)-f(\gamma, \alpha_1) \\ &= f(\alpha_1, \gamma)-f(\alpha_1, \gamma) \qquad [\because f \text{ is symmetric}] \\ &= 0. \end{aligned}$$

Scanned with CamScanner

$\therefore$ $\beta \in W_2$ 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 $V = W_1 + W_2$.

$\therefore$ $V = W_1 \oplus W_2$.

So $\dim W_2 = \dim V - \dim W_1 = n - 1$.

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

$$g(\alpha_i, \alpha_j) = 0,\ i \neq j\ (i \geqslant 2, j \geqslant 2)$$
$$\Rightarrow f(\alpha_i, \alpha_j) = 0,\ i \neq j\ (i \geqslant 2, j \geqslant 2)$$

[$\because$ g is restriction of f]

Now by definition of W_2, we have

$$f(\alpha_1, \alpha_j) = 0 \text{ for } j = 2, 3, \ldots, n.$$

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.

§ 4. 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.*

Proof. Let f be a bilinear form on a vector space V. Let

$$g(\alpha, \beta) = \tfrac{1}{2} [f(\alpha, \beta) + f(\beta, \alpha)] \quad \ldots(1)$$
$$h(\alpha, \beta) = \tfrac{1}{2} [f(\alpha, \beta) - f(\beta, \alpha)] \quad \ldots(2)$$

for all α, β in V.

Then it can be easily seen that both g and h are bilinear forms on V. We have

$$g(\beta, \alpha) = \tfrac{1}{2} [f(\beta, \alpha) + f(\alpha, \beta)] = g(\alpha, \beta).$$

$\therefore$ g is symmetric.

Also $h(\beta, \alpha) = \tfrac{1}{2} [f(\beta, \alpha) - f(\alpha, \beta)] = -\tfrac{1}{2} [f(\alpha, \beta) - f(\beta, \alpha)]$
$= -h(\alpha, \beta)$.

$\therefore$ h is skew-symmetric.

Adding (1) and (2), we get

$$g(\alpha, \beta) + h(\alpha, \beta) = f(\alpha, \beta)$$
$$\Rightarrow (g+h)(\alpha, \beta) = f(\alpha, \beta) \text{ for all } \alpha, \beta \text{ in } V.$$
$$\therefore \quad g + h = f.$$

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 $f(\alpha, \beta) = f_1(\alpha, \beta) + f_2(\alpha, \beta)$. $\quad \ldots(3)$

Also $f(\beta, \alpha) = (f_1 + f_2)(\beta, \alpha)$

or $f(\beta, \alpha) = f_1(\beta, \alpha) + f_2(\beta, \alpha)$

or $f(\beta, \alpha) = f_1(\alpha, \beta) - f_2(\alpha, \beta)$. $\quad \ldots(4)$

[$\because$ f_1 is symmetric and f_2 is skew-symmetric]

Adding (3) and (4), we get

$$f(\alpha, \beta) + f(\beta, \alpha) = 2 f_1(\alpha, \beta)$$

i.e. $f_1(\alpha, \beta) = \tfrac{1}{2} [f(\alpha, \beta) + f(\beta, \alpha)]$
$= g(\alpha, \beta)$.

$\therefore$ $f_1 = g$.

Subtracting (4) from (3), we get

$$f(\alpha, \beta) - f(\beta, \alpha) = 2 f_2(\alpha, \beta)$$

i.e. $f_2(\alpha, \beta) = \tfrac{1}{2} [f(\alpha, \beta) - f(\beta, \alpha)] = h(\alpha, \beta)$.

$\therefore$ $f_2 = h$.

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 A in some (or every) ordered basis is skew-symmetric, i.e., $A' = -A$.*

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

Scanned with CamScanner

V and 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 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. $A'=-A$

i.e. A is skew-symmetric.

§ 5. Groups Preserving Bilinear forms.

Definition. *Let f be a bilinear form and T be a linear operator on a vector space V over the field F. We say that T* **preserves** *f if $f(T\alpha, T\beta)=f(\alpha, \beta)$ for all α, β in V.*

The identity operator I preserves every bilinear form. For if f is any bilinear form, then

$$f(I\alpha, I\beta)=f(\alpha, \beta) \text{ for all } \alpha, \beta \text{ in } V.$$

If S and T are linear operators which preserve f, then the product ST also preserves f. In fact for all α, β in V we have

$$f(ST\alpha, ST\beta)=f(T\alpha, T\beta) \quad [\because S \text{ preserves } f]$$
$$=f(\alpha, \beta) \quad [\because T \text{ preserves } f]$$

$\therefore$ ST preserves f.

Therefore if G is the set of all linear operators on V which preserve a given bilinear form f on V, then G is closed with respect to the operation of product of two linear operators.

Theorem. *Let f be a non-degenerate bilinear form on a finite-dimensional vector space V. The set G of all linear operators on V which preserve f is a group under the operation of composition.*

Proof. If S, T are elements of G *i.e.*, if S and T are linear operators on V preserving f, then ST is also a linear operator on V preserving f. Therefore ST is also an element of G. Thus G is closed with respect to the operation of product of two linear operators.

The product of linear operators is an associative operation because the product of functions is associative.

The identity operator I on V also preserves f. Therefore I is an element of G.

Now let $T \in G$. Then from the fact that f is non-degenerate, we shall prove that any operator T in G is invertible, and T^{-1} is also in G. Since $T \in G$, therefore T preserves f. Let α be a vector in the null space of T *i.e.*, $T\alpha=\mathbf{0}$. Then for any β in V, we have

$$f(\alpha, \beta)=f(T\alpha, T\beta)=f(\mathbf{0}, T\beta)$$
$$=f(0\mathbf{0}, T\beta)=0f(\mathbf{0}, T\beta)=0.$$

Since $f(\alpha, \beta)=0$ for all β in V and f is non-degenerate, therefore $\alpha=\mathbf{0}$.

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

$\therefore$ T is non-singular. Since V is finite-dimensional, therefore T is invertible. Clearly T^{-1} also preserves f; for

$$f(T^{-1}\alpha, T^{-1}\beta)=f(TT^{-1}\alpha, TT^{-1}\beta) \quad [\because T \text{ preserves } f]$$
$$=f(I\alpha, I\beta)=f(\alpha, \beta).$$

Thus $T^{-1} \in G$.

Hence G is a group with respect to the operation of product of linear operators.

Exercises

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) $f(\alpha, \beta)=x_1y_1+x_2y_2$;

   (ii) $f(\alpha, \beta)=1$;

   (iii) $f(\alpha, \beta)=x_1y_1+x_1y_2+x_2y_1+x_2y_2$. (Meerut 1979)

2. Let m and n be positive integers and F a field. Let V be the vector space of all $m\times n$ matrices over F. Let A be a fixed $m\times m$ matrix over F. Define

   $f(X, Y)=\text{tr}(X' AY)$, where tr stands for trace and X' denotes the transpose of the matrix X.

   Show that f is a bilinear form on V.

3. Let F be a field. Find all bilinear forms on the vector space F^2. (Meerut 1972)

4. Let f be the bilinear form on $\mathbf{R}^2$ defined by

   $$f((x_1, y_1), (x_2, y_2))=x_1y_1+x_2y_2.$$

   Find the matrix of f in each of the following bases.

   (i) $\{(1, 0), (0, 1)\}$; (Meerut 1976, 77)

   (ii) $\{(1, 2), (3, 4)\}$; (Meerut 1977)

   (iii) $\{(1, 1), (0, 1)\}$.

5. Let f be a bilinear form on $\mathbf{R}^2$ defined by

   $$f[(x_1, x_2), (y_1, y_2)]=2x_1y_1-3x_1y_2+x_2y_2.$$

   Find the matrix of f in the basis (Meerut 1976)

   $$\{\alpha_1=(1, 0), \alpha_2=(1, 1)\}.$$

Scanned with CamScanner

6. Let f be the bilinear form on $\mathbf{R}^2$ defined by

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

(i) Find the matrix of f in the standard ordered basis

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

(ii) Find the transition matrix from the basis B to the basis

$$B' = \{(1, -1), (1, 1)\}.$$

(iii) Find the matrix of f in the basis B'.

7. Describe explicitly all bilinear forms f on $\mathbf{R}^3$ with the property that $f(\alpha, \beta) = f(\beta, \alpha)$ for all α, β. (Meerut 1973, 84)

Or

Describe explicitly all symmetric bilinear forms on $\mathbf{R}^3$.

8. Find all skew-symmetric bilinear forms on $\mathbf{R}^3$. (Meerut 1980)

9. Let V be an n-dimensional vector space over the field of complex numbers and f be a skew-symmetric bilinear form on V. Prove that the rank of f is even. (Meerut 1977)

Answers

1. (i) Bilinear form; (ii) not a bilinear form; (iii) bilinear form.

3. Let A be any 2×2 matrix over F and B any ordered basis of F^2. Then the bilinear forms on F^2 are precisely those obtained by $f(\alpha, \beta) = X' AY$, where X, Y are the co-ordinate matrices of α and β in the ordered basis B.

4. (i) $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$; (ii) $\begin{bmatrix} 4 & 14 \\ 14 & 24 \end{bmatrix}$; (iii) $\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$.

5. $\begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}$.

6. (i) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$; (ii) $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$; (iii) $\begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$.

Scanned with CamScanner