மனோன்மணியம் சுந்தரனார் பல்கலைக்கழகம்

MANONMANIAM SUNDARANAR UNIVERSITY

TIRUNELVELI-627 012

தொலைநிலை தொடர் கல்வி இயக்ககம்

DIRECTORATE OF DISTANCE AND CONTINUING EDUCATION

B.Sc. MATHEMATICS

III YEAR

LINEAR ALGEBRA

Sub. Code: JMMA61

Prepared by

Dr. S. KALAISELVI

Assistant Professor

Department of Mathematics

Sarah Tucker College (Autonomous), Tirunelvei-7.

B.Sc. MATHEMATICS –III YEAR

JMMA61: LINEAR ALGEBRA

SYLLABUS

UNIT I

Vector Spaces: Definition and examples–Subspaces–Linear Transformations–Fundamental theorem of homomorphism.

Chapter 1: Sections -1.1 to 1.3

UNIT II

Span of a set–Linear Dependence and Independence–Basis and Dimension.

Chapter 2: Sections-2.1 to 2.3

UNIT III

Rank and Nullity of a transformation– Matrix of a linear transformation – Inner product space: Definition and examples– Orthogonality – Orthogonal complement.

Chapter 3: Sections- 3.1 to 3.5

UNIT IV

Matrices – Elementary transformation – Rank of a matrix – Simultaneouslinearequations–CharacteristicequationandCayley-Hamilton

Theorem. Chapter7: Sections -7.4 to 7.7

UNIT V

Eigen values and Eigen vectors–Properties and problems– Bilinear forms – Quadratic forms – Reduction of quadratic form to diagonal form.

Chapter7: Sections-7.8 and Chapter8: Sections-8.1, 8.2

Recommended Text

S. Arumugam and A. Thangapandi Isaac, Modern Algebra, Scitech, 2014.

JMMA61: LINEAR ALGEBRA

CONTENTS

Unit I		
1.1	Vector Spaces: Definition and examples	3
1.2	Subspaces	15
1.3	Linear Transformations	24
Unit II		
2.1	Span of a set	33
2.2	Linear Dependence and Independence	37
2.3	Basis and Dimension	45
Unit III		
3.1	Rank and Nullity of a transformation	61
3.2	Matrix of a linear transformation	65
3.3	Inner product space: Definition and examples	72
3.4	Orthogonality	82
3.5	Orthogonal complement	89
Unit IV		
4.1	Matrices	94
4.2	Elementary transformation	104
4.3	Rank of a matrix	110
4.4	Simultaneous linear equations	118
4.5	Characteristic equation and Cayley-Hamilton	
Unit V		
5.1	Eigen values and Eigen vectors–Properties and problem	129
5.2	Bilinear forms	151
5.3	Quadratic forms	157
5.4	Reduction of quadratic form to diagonal form	162

UNIT I

Vector Spaces: Definition and examples–Subspaces–Linear Transformations–Fundamental theorem of homomorphism.

Chapter 1: Sections -1.1 to 1.3

1.Vector Spaces

1.1. Definition and Examples

A non-empty set V is said to be a vector space over a field F if

(i) V is an abelian group under an operation called addition which we denote by + .

(ii) For every $\alpha \in F$ and $v \in V$, there is defined an element αv in V subject to the following conditions.

(a) $\alpha(u+v) = \alpha u + \alpha v$ for all $u, v \in V$ and $\alpha \in F$.

(b) $(\alpha+\beta)u = \alpha u + \beta u$ for all $u \in V$ and $\alpha, \beta \in F$.

(c) $\alpha(\beta u) = (\alpha\beta)u$ for all $u \in V$ and $\alpha, \beta \in F$.

(d) $lu = u$ for all $u \in V$.

Remarks.

1. The elements of F are called scalars and the elements of V are called vectors.
2. The rule which associates with each scalar $\alpha \in F$ and a vector $v \in V$, a vector αv is called the scalar multiplication. Thus, a scalar multiplication gives rise to a function from $F \times V \to V$ defined by $(\alpha, v) \to \alpha v$.

Example 1:

R $\times$ **R** is a vector space over **R** under addition and scalar multiplication defined by $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$

Proof:

Clearly the binary operation + is commutative and associative and $(0,0)$ is the zero element.

The inverse of (x_1, x_2) is $(-x_1, -x_2)$.

Hence ($\mathbf{R} \times \mathbf{R}, +$) is an abelian group.

Now, let $u = (x_1, x_2)$ and $v = (y_1, y_2)$ and let $\alpha, \beta \in \mathbf{R}$.

Then $\alpha(u + v) = \alpha[(x_1, x_2) + (y_1, y_2)]$

$$
\begin{aligned}
&= \alpha(x_1 + y_1, x_2 + y_2) \\
&= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2) \\
&= (\alpha x_1, \alpha x_2) + (\alpha y_1, \alpha y_2) \\
&= \alpha(x_1, x_2) + \alpha(y_1, y_2) \\
&= \alpha u + \alpha v.
\end{aligned}
$$

Now, $(\alpha + \beta)u = (\alpha + \beta)(x_1, x_2)$

$$
\begin{aligned}
&= ((\alpha + \beta)x_1, (\alpha + \beta)x_2) \\
&= (\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2) \\
&= (\alpha x_1, \alpha x_2) + (\beta x_1, \beta x_2) \\
&= \alpha(x_1, x_2) + \beta(x_1, x_2) \\
&= \alpha u + \beta u
\end{aligned}
$$

Also $\alpha(\beta u) = \alpha\big(\beta(x_1, x_2)\big)$

$$
\begin{aligned}
&= \alpha(\beta x_1, \beta x_2) \\
&= (\alpha\beta x_1, \alpha\beta x_2) \\
&= (\alpha\beta)(x_1, x_2) \\
&= (\alpha\beta)u
\end{aligned}
$$

Obviously $1u = u$.

$\therefore$ $\mathbf{R} \times \mathbf{R}$ is a vector space over $\mathbf{R}$.

Example 2:

$\mathbf{R}^n = \{(x_1, x_2, \dots, x_n)/x_i \in R, 1 \leq i \leq n\}$. Then $\mathbf{R}^n$ is a vector space over $\mathbf{R}$ under addition and scalar multiplication defined by

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$

And $\alpha(x_1, x_2, \ldots x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$.

Proof:

Clearly the binary operation + is commutative and associative.

$(0,0, \ldots, \ldots, 0)$ is the zero element.

The inverse of $(x_1, x_2, \ldots, x_n)$ is

$(-x_1, -x_2, \ldots, -x_n)$

Hence $(\mathbf{R}^n, +)$ is an abelian group.

Now, let $u = (x_1, x_2, \ldots, x_n)$ and

$v = (y_1, y_2, \ldots, y_n)$ and let $\alpha, \beta \in R$.

$$\begin{aligned}
\text{Then } \alpha(u+v) &= \alpha[(x_1, x_2, \ldots \ldots, x_n) \\
&\quad + (y_1, y_2, \ldots \ldots, y_n)] \\
&= \alpha(x_1 + y_1, x_2 + y_2, \ldots \ldots, x_n + y_n) \\
&= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, \ldots \ldots, \alpha x_n + \alpha y_n) \\
&= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n) + (\alpha y_1, \alpha y_2, \ldots, \alpha y_n) \\
&= \alpha(x_1, x_2, \ldots, x_n) + \alpha(y_1, y_2, \ldots, y_n) \\
&= \alpha u + \alpha v.
\end{aligned}$$

Similarly, $(\alpha + \beta)u = \alpha u + \beta u$ and $\alpha(\beta u) = (\alpha\beta)u$.

Obviously $1u = u$.

$\therefore$ $\mathbf{R}^n$ is vector space over $\mathbf{R}$.

Note. We denote this vector space by V_n(R).

Example 3:

Let F be any field. Let $F^n = \{(x_1, x_2, \ldots, x_n)/x_i \in F\}$.

In F^n we define addition and scalar multipli. cation as in Example 2.

Then F^n is a vector space over F and we denote this vector space by $V_n(F)$.

Note. In this example if we take $n = 1$ then we see that any field F is a vector space over itself. The addition and scalar multiplication in this vector space are simply the addition and multiplication of the field F.

Example 4:

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

Here addition is the usual addition in $\mathbf{C}$ and the scalar multiplication is the usual multiplication of a real number and a complex number.

(i.e) $$(x_1 + ix_2) + (y_1 + iy_2) = (x_1 + y_1) + i(x_2 + y_2)$$
$$\text{and } \alpha(x_1 + ix_2) = \alpha x_i + i\alpha x_2$$

Proof:

Clearly $(\mathbf{C}, +)$ is an abelian group. Also, the remaining axioms of a vector space are true since the scalars and vectors involved are complex numbers and further the operations are usual addition and multiplication Hence $\mathbf{C}$ is a vector space over $\mathbf{R}$.

Example 5:

Let $V = \{a + b\sqrt{2}/a, b \in \mathbf{Q}\}$. Then V is a vector space over $\mathbf{Q}$ under addition and multiplication.

Proof:

Obviously V is an abelian group under usual addition.

The remaining axioms of a vector space are true since the scalars and vectors are real numbers and the operations are usual addition and multiplication. Hence V is a vector space over $\mathbf{Q}$.

Example 6:

Let F be a field. Then $F[x]$, the set of all polynomials over F, is a vector space over F under the addition of polynomials and scalar multiplication defined by

$$\alpha(a_0 + a_1x + \cdots + a_nx^n)$$
$$= \alpha a_0 + \dot{\alpha} a_1 x + \cdots + \alpha a_n x^n$$

Example 7:

The set V of all polynomials of degree $\leq n$ including the zero polynomial in $F[x]$ is a vector space over the field F under the addition and scalar multiplication defined as in Example 6.

Proof:

Let $f, g \in V$. Then f and g are polynomials of degree $\leq n$.

$\therefore f + g$ and αf are of degree $\leq n$.

$\therefore f + g, \alpha f \in V$.

The other axioms of a vector space can easily be verified. Hence V is a vector space over F.

Example 8:

The set $M_2(\mathbf{R})$ of all 2×2 matrices is a vector space over $\mathbf{R}$ under matrix addition and scalar multiplication defined by $\alpha \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{pmatrix}$

Example 9:

Let V be the set of all functions from $\mathbf{R}$ to $\mathbf{R}$.

Let $f, g \in V$. We define

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

$(\alpha f)(x) = \alpha[f(x)]$.

V is a vector space over $\mathbf{R}$. (verify)

Example 10:

Let V denote the set of all solutions of the differential equation

$2\frac{d^2y}{dx^2} - 7\frac{dy}{dx} + 3y = 0$. Then V is a vector space over $\mathbf{R}$.

Proof:

Let $f, g \in V$ and $\alpha \in \mathbf{R}$.

Then $2\frac{d^2f}{dx^2} - 7\frac{df}{dx} + 3f = 0$ and

$$2\frac{d^2g}{dx^2} - 7\frac{dg}{dx} + 3g = 0$$

$$\therefore 2\left(\frac{d^2f}{dx^2} + \frac{d^2g}{dx^2}\right) - 7\left(\frac{df}{dx} + \frac{dg}{dx}\right) + 3(f + g) = 0$$

$$\therefore 2\frac{d^2}{dx^2}(f + g) - 7\frac{d}{dx}(f + g) + 3(f + g) = 0$$

Hence $f + g \in V$.

Also $2\frac{d^2}{dx^2}(\alpha f) - 7\frac{d}{dx}(\alpha f) + 3(\alpha f) = 0.$

Hence af $\in V$.

Since the operations are usual addition and scalar multiplication, the axioms of vector space are true.

Hence V is a vector space over **R**.

Example 11:

Any sequence of real numbers $a_1, a_2, \dots, a_n$, ... is usually denoted by the symbol (a_n). Let V denote the set of all sequences of real numbers. V is a vector space over the field of real numbers. The addition and scalar multiplication are defined by $(a_n) + (b_n) = (a_n + b_n)$ and

$$\alpha(a_n) = (\alpha a_n)$$

Example 12:

Let $V = \{0\}$. V is a vector space over any field F under the obvious operations of addition and scalar multiplication.

Example 13:

$\mathbf{R}$ is not a vector space over $\mathbf{C}$.

Clearly $(\mathbf{R}, +)$ is an abelian group.

But the scalar multiplications are not defined, for if $\alpha = a + ib \in \mathbf{C}$ and $u \in \mathbf{R}$, then $\alpha u = au + ibu \notin \mathbf{R}$.

$\therefore$ $\mathbf{R}$ is not vector space over $\mathbf{C}$.

Example 14:

Consider $\mathbf{R} \times \mathbf{R}$ with usual addition. We define scalar multiplication by $\alpha(x, y) = (\alpha x, \alpha^2 y)$. Then $\mathbf{R} \times \mathbf{R}$ is not a vector space over $\mathbf{R}$. Clearly $\mathbf{R} \times \mathbf{R}$ with usual addition is an abelian group.

$$(\alpha + \beta)(x, y) = ((\alpha + \beta)x, (\alpha + \beta)^2 y)$$
$$= (\alpha x + \beta x, \alpha^2 y + \beta^2 y + 2\alpha\beta y)$$

Also, $\alpha(x, y) + \beta(x, y) = (\alpha x, \alpha^2 y) + (\beta x, \beta^2 y)$

$$= (\alpha x + \beta x, \alpha^2 y + \beta^2 y).$$

Hence $(\alpha + \beta)(x, y) \neq \alpha(x, y) + \beta(x, y)$.

$\therefore$ $\mathbf{R} \times \mathbf{R}$ is not a vector space over $\mathbf{R}$.

Example 15:

Consider $\mathbf{R} \times \mathbf{R}$ with usual addition. Define the scalar multiplication as $\alpha(a, b) = (0,0)$. Clearly $\mathbf{R} \times \mathbf{R}$ is an abelian group. Also,

(i) $\alpha(u + v) = 0$ and $\alpha u + \alpha v = 0 + 0 = 0$; so that $\alpha(u + v) = \alpha u + \alpha v$.

(ii) Similarly $(\alpha + \beta)u = \alpha u + \beta u = 0$.

(iii) $\alpha(\beta u) = (\alpha\beta)u = 0$.

However, $1(a, b) = (0,0)$.

Hence it is not a vector space.

Note. In this example all the axioms except the axiom $1u = u$ are satisfied. Hence the axiom $1u = u$ cannot be derived from the other axioms of the vector space.

Thus, the axiom $lu = u$ is independent of the other axioms of the vector space. We say that the axiom $lu = u$ is irredundant.

16. Let V be the set of all ordered pairs of real numbers. Addition and multiplication are defined by $(x, y) + (x_1, y_1) = (x + x_1, y + y_1)$ and $\alpha(x, y) = (x, \alpha y)$ where x, y, x_1, y_1 and α are real numbers.

Then V is not a vector space over **R**.

Clearly V is an abelian group under the operation + defined above

Let $\alpha, \beta \in$ R and $(x, y) \in V$.

$$\text{Now } (\alpha + \beta)(x, y) = (x, (\alpha + \beta)y)$$
$$= (x, \alpha y + \beta y)$$

Also $\alpha(x, y) + \beta(x, y) = (x, \alpha y) + (x, \beta y) = (2x, \alpha y + \beta y)$

$\therefore\ (\alpha + \beta)(x, y) \neq \alpha(x, y) + \beta(x, y)$.

Hence V is not a vector space over **R**.

Example 17:

Let $\mathbf{R}^+$be the set of all positive real numbers, define addition and scalar multiplication as follows

$u + v = uv$ for all $u, v \in \mathbf{R}^+$; $\alpha u = u^a$ for all $u \in \mathbf{R}^+$and $\alpha \in \mathbf{R}$. Then $\mathbf{R}^+$is a real vector space.

Proof:

Clearly ($\mathbf{R}^+, +$) is an abelian group with identity 1. (verify)

Now,

$$
\begin{aligned}
\alpha(u+v) &= \alpha(uv)\\
&= (uv)^{\alpha}\\
&= u^{\alpha}v^{\alpha}\\
&= \alpha u + \alpha v\\
(\alpha+\beta)u &= u^{\alpha+\beta}\\
&= u^{\alpha}u^{\beta}\\
&= \alpha u + \beta u\\
\alpha(\beta u) &= \alpha u^{\beta}\\
&= \left(u^{\beta}\right)^{\alpha}\\
&= u^{\beta\alpha}\\
&= u^{\alpha\beta}\\
&= (\alpha\beta)u
\end{aligned}
$$

Also, $1u = u1 = u$

$\therefore$ $\mathbf{R}^{+}$is a vector space over **R**.

Exercises

1. Prove that the following are vector spaces over the stated fields under usual operations.

   (a) $V = \{a + b\sqrt{2} + c\sqrt{3} / a, b, c \in \mathbf{Q}\}$ over **Q**.

   (b) $V = \left\{\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} / a, b \in \mathbf{R}\right\}$ over **R**.

   (c) V = set of all constant functions from R to **R**.

   (d) V = set of all expressions of the form $ax + by + cz$ where $a, b, c \in \mathbf{R}$, under usual addition and scalar multiplication over **R**.

   (e) Let F be any field and X any non-empty set. Then F^{x} is a vector space over F.

   (f) $V = \{0,1,2, x+1, x+2, 2x+1, 2x+2, x, 2x\} \subset \mathbf{Z}_3[x]$ over $\mathbf{Z}_3$.

   (g) $V = \{(a, b, c) / a, b, c \in$ Q and $a + 2b = 3c\}$ over **Q**.

   (h) $V = \{f / f \in F[x]$ and $f(\alpha) = 0\}$ over F.

2. Show that the following are not vector spaces under usual addition and scalar multiplication.

   (a) Z over Q

   (b) Q$[x]$ over R

   (c) Z over Z_5

   (d) The set of all polynomials in $F[x]$ of degree n where n is a fixed positive integer over F.

   (e) The set of all polynomials in $F[x]$ of degree $\geq n$ where n is a fixed positive integer over F.

3. Show that the following are not vector spaces.

   (a) **R** $\times$ **R** with addition defined by $(a, b) + (c, d) = (a - c, b - d)$ and usual scalar multiplication.

   (b) **R** $\times$ **R** with addition defined by $(a, b) + (c, d) = (0, b + d)$ and usual scalar multiplication.

   (c) **R** $\times$ **R** with usual addition and scalar multiplication defined by $\alpha(a, b) = (0, \alpha b)$.

   (d) **R** $\times$ **R** with usual addition and scalar multiplication defined by $\alpha(a, b) = (|\alpha|a, |\alpha|b)$.

   (e) $V = \{0, x + 4, 2x + 3, 3x + 1\} \subset Z_5[x]$ w.r.t. addition modulo 5 and scalar multiplication modulo 5 .

   (f) **R** $\times$ **R** with addition defined by $(a, b) + (c, d) = (ac, bd)$ and usual scalar multiplication.

   (g) **R** $\times$ **R** with addition defined by $(a, b) + (c, d) = (a + c, b + d)$ and scalar multiplication defined by $k(a, b) = (k^2a, k^2b)$.

4. V is the set of polynomials in $\mathbf{R}[x]$ of degree ≤ 1. If $f = \alpha_0 + \alpha_1 x$ and $g = \beta_0 + \beta_1 x$, then we define

$$f + g = (\alpha_0 + \beta_0) + (\alpha_0\beta_1 + \alpha_1\beta_0)x$$
$$\text{and } \alpha f = \alpha\alpha_0 + \alpha\alpha_1 x.$$

Show that V is not a vector space over $\mathbf{R}$.

5. Let X be any set. Show that $\wp(X)$ is a vector space over Z_2 under the following operations.

$$A + B = A\Delta B, \alpha A = \begin{cases} \Phi \text{ if } \alpha = 0 \\ A \text{ if } \alpha = 1 \end{cases}$$

6.Let U and W be vector spaces over the same field $F. V = \{(u, v)/u \in U, v \in W\}$. Show that V is a vector space over F with addition and scalar multiplication defined by $(u, w) + (u', w') = (u + u', w + w')$ and $k(u, w) = (ku, kw)$ where $u, u' \in U; w, w' \in W$ and $k \in F$.

Remark.

Commutativity of addition in a vector space can be derived from the other axioms of the vector space (i.e), the axiom of commutativity of addition in a vector space is redundant, for,

$$\begin{aligned} (1+1)(u+v) &= 1(u+v) + 1(u+v) \\ &= 1u + 1v + 1u + 1v \\ &= u + v + u + v \end{aligned}$$

$$\begin{aligned} \text{Also, } (1+1)(u+v) &= (1+1)u + (1+1)v \\ &= u + u + v + v. \end{aligned}$$

$$\therefore\ u + v + u + v = u + u + v + v.$$
$$\therefore\ v + u = u + v.$$

Theorem 1:

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

(i) $\alpha\mathbf{0} = \mathbf{0}$ for all $\alpha \in F$.

(ii) $0v = 0$ for all $v \in V$.

(iii) $(-\alpha)v = \alpha(-v) = -(\alpha v)$ for all $\alpha \in F$ and $v \in V$.

(iv) $\alpha v = \mathbf{0} \Rightarrow \alpha = 0$ or $v = \mathbf{0}$.

Proof:

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

(ii) $0v = (0 + 0)v = 0v + 0v$. Hence $0v = 0$.

(iii) $\mathbf{0} = [\alpha + (-\alpha)]v = \alpha v + (-\alpha)v$.

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

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

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

(iv) Let $\alpha v = 0$. If $\alpha = 0$, there j, thing to prove.

$\therefore$ Let $\alpha \neq 0$. Then $\alpha^{-1} \in F$

Now, $v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v)$

$= \alpha^{-1}0 = 0$.

Exercises

1. Let V be any vector space over a field F. Show that

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

   (b) $\alpha u = \alpha v$ and $\alpha \neq 0 \Rightarrow u = v$.

   (c) $\alpha u = \beta u$ and $u \neq 0 \Rightarrow \alpha = \beta$.

2. Determine which of the following statements are true and which are false.

   (a) $\{(a, 0,0)/a \in \mathbf{R})\}$ is a vector space over **R**.

   (b) $\mathbf{R}[x]$ is a vector space over **C**.

   (c) The set of all polynomials of even degree in $\mathbf{R}[x]$ is a vector space over R.

   (d) The set of all solutions in $\mathbf{R} \times \mathbf{R}$ of the equation $x + 2y = 0$ is a vector space over **R**.

(e) The sum of two vectors is a vector.

(f) The product of a scalar and a vector is a vector.

Answers.

2.(a) T (b) F (c) F (d) T (e) T (f) T.

1.2. Subspaces:

Definition:

Let V be a vector space over a field F. A non-empty subset W of V is called a subspace of V if W itself is a vector space over F under the operations of V.

Theorem 2:

Let V be a vector space over F. A non-empty subset W of V is a subspace of V iff W is closed with respect to vector addition and scalar multiplication in V.

Proof:

Let W be a subspace of V.

Then W itself is a vector space and hence W is closed with respect to vector addition and scalar multiplication.

Conversely, let W be a non-empty subset of V such that

$u, v \in W \Rightarrow u + v \in W$ and $u \in W$ and $\alpha \in F \Rightarrow \alpha u \in W$.

We prove that W is a subspace of V.

Since W is non-empty, there exists an element $u \in W$.

$\therefore\ 0u = 0 \in W$.

Also $v \in W \Rightarrow (-1)v = -v \in W$.

Thus W contains 0 and the additive inverse of each of its element.

Hence W is an additive subgroup of V.

Also $u \in W$ and $\alpha \in F \Rightarrow \alpha u \in W$.

Since the elements of W are the elements of V the other axioms of a vector space are true in W.

Hence W is a subspace of V.

Theorem 3:

Let V be a vector space over a field F. A non-empty subset W of V is a subspace of V iff $u, v \in W$ and $\alpha, \beta \in F \Rightarrow \alpha u + \beta v \in W$.

Proof:

Let W be a subspace of V. Let $u, v \in W$ and $\alpha; \beta \in F$.

Then αu and $\beta v \in W$ and hence $\alpha u + \beta v \in W$.

Conversely,

let $u, v \in W$ and $\alpha, \beta \in F \Rightarrow \alpha u + \beta v \in W$.

Taking $\alpha = \beta = 1$, we get $u, v \in W \Rightarrow u + v \in W$.

Taking $\beta = 0$, we get

$\alpha \in F$ and $u \in W \Rightarrow \alpha u \in W$. W is a subspace of V.

Examples

1. $\{0\}$ and V are subspaces of any vector space V. They are called the trivial subspaces of V.
2. $W = \{(a, 0,0)/a \in \mathbf{R}\}$ is a subspace of $\mathbf{R}^3$. for, let $u = (a, 0,0), v = (b, 0,0) \in W$ and $\alpha, \beta \in \mathbf{R}$.

   Then $\alpha u + \beta v = \alpha(a, 0,0) + \beta(b, 0,0)$

   $= (\alpha a + \beta b, 0,0) \in W$.

   Hence W is a subspace of $\mathbf{R}^3$.

   Note. Geometrically W consists of all points on the x-axis in the Euclidean 3 space.

3. In $\mathbf{R}^3, W=\{(ka,kb,kc)/k\in\mathbf{R}\}$ is a subspace of $\mathbf{R}^3$.

For, if $u=(k_1a,k_1b,k_1c)$ and

$v=(k_2a,k_2b,k_2c)\in W$ and $\alpha,\beta\in\mathbf{R}$

then $\alpha u+\beta v=\alpha(k_1a,k_1b,k_1c)+\ \beta(k_2a,k_2b,k_2c)$

$=\big((\alpha k_1+\beta k_2)a,(\alpha k_1+\beta k_2)b,(\alpha k_1+\beta k_2)c\big)\in W$

Hence W is a subspace of $\mathbf{R}^3$.

Note:

Geometrically W consists of all points of the line $\frac{x}{a}=\frac{y}{b}=\frac{z}{c}$ provided a,b,c are not all zero. Thus, the set of all points on a line through the origin is a subspace of $\mathbf{R}^3$. However a line not passing through the origin is not a subspace of $\mathbf{R}^3$, since the additive identity (0,0,0) does not lie on the line.

4. $W=\{(a,b,0)/a,b\in\mathbf{R}\}$ is a subspace of $\mathbf{R}^3$. W consists of all points of the xy-plane.

5. Let W be the set of all points in $\mathbf{R}^3$ satisfying the equation $lx+my+nz=0$. W is a subspace of $\mathbf{R}^3$. For, let $u=(a_1,b_1,c_1)$ and $v=(a_2,b_2,c_2)\in W$ and $\alpha,\beta\in\mathbf{R}$.

Then we have

$la_1+mb_1+nc_1=0=la_2+mb_2+nc_2$.

Hence $\alpha(la_1+mb_1+nc_1)+\beta(la_2+mb_2+nc_2)=0$.

(i.e), $l(\alpha a_1+\beta a_2)+m(\alpha b_1+\beta b_2)+n(\alpha c_1+\beta c_2)=0$.

(i.e), $\alpha u+\beta v\in W$ so that W is a subspace of $\mathbf{R}^3$.

Note. Geometrically W consists of all points on the plane $lx+my+nz=0$, which passes through the origin.

6. Let $W=\{f/f\in F[x]$ and $f(a)=0\}$.

(i.e) W is the set of all polynomials in $F[x]$ having a as a root where $a\in F$. Then

W is a vector space over F.

We observe that $x - a \in W$ and hence W is non-empty.

Let $f, g \in F[x]$ and $\alpha, \beta \in F$.

To prove that $\alpha f + \beta g \in W$ we have to show that a is a root of $\alpha f + \beta g$.

Now, $(\alpha f + \beta g)(a) = \alpha f(a) + \beta g(a)$

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

Hence a is a root of $\alpha f + \beta g$.

$\therefore\ \alpha f + \beta g \in W$ and hence W is a subspace of $F[x]$.

7. $W = \left\{\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} / a, b \in \mathbf{R}\right\}$ is a subspace of $M_2(\mathbf{R})$

Solved problems

Problem 1:

Prove that the intersection of two subspaces of a vector space is a subspace.

Solution:

Let A and B be two subspaces of a vector space V over a field F.

We claim that $A \cap B$ is a subspace of V.

Clearly $0 \in A \cap B$ and hence $A \cap B$ is non-empty.

Now, let $u, v \in A \cap B$ and $\alpha, \beta \in F$.

Then $u, v \in A$ and $u, v \in B$.

$\therefore\ \alpha u + \beta v \in A$ and $\alpha u + \beta v \in B$

(since A and B are subspaces)

$\therefore\ \alpha u + \beta v \in A \cap B$.

Hence $A \cap B$ is a subspace of V.

Problem 2:

Prove that the union of two subspaces of a vector space need not be a subspace.

Solution:

Let $A = \{(a, 0,0)/a \in \mathbf{R}\}$

$B = \{(0, b, 0)/b \in \mathbf{R}\}$.

Clearly A and B are subspaces of $\mathbf{R}^3$ (example 2 of 1.2).

However, $A \cup B$ is not a subspace of $\mathbf{R}^3$.

For, (1,0,0) and $(0,1,0) \in A \cup B$.

But $(1,0,0) + (0,1,0) = (1,1,0) \notin A \cup B$.

Problem 3:

If A and B are subspaces of V prove that $A + B = \{v \in V/v = a + b, a \in A, b \in B\}$ is a subspace of V. Further show that $A + B$ is the smallest subspace containing A and B. (ie) If W is any any subspace of V containing A and B then W contains $A + B$.

Solution:

Let $v_1, v_2 \in A + B$ and $\alpha \in F$.

Then $v_1 = a_1 + b_1, v_2 = a_2 + b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$.

Now, $$\begin{aligned} v_1 + v_2 &= (a_1 + b_1) + (a_2 + b_2) \\ &= (a_1 + a_2) + (b_1 + b_2) \in A + B. \end{aligned}$$

Also $\alpha(a_1 + b_1) = \alpha a_1 + \alpha b_1 \in A + B$.

Hence $A + B$ is a subspace of V. Clearly $A \subseteq A + B$ and $B \subseteq A + B$.

Now, let W be any subspace of V containing A and B.

We shall prove that $A + B \subseteq W$.

Let $v \in A + B$. Then $v = a + b$ where $a \in A$ and $b \in B$.

Since $A \subseteq W, a \in W$. Similarly, $b \in W$.

$\therefore\ a + b = v \in W$.

$\therefore\ A + B \subseteq W$ so that $A + B$ is the smallest subspace of V containing A and B.

Problem 4:

Let A and B be subspace of a vector space V. Then $A \cap B = \{0\}$ iff every vector $v \in A + B$ can be uniquely expressed in the form $v = a + b$ where $a \in A$ and $b \in B$.

Solution:

Let $A \cap B = \{0\}$. Let $v \in A + B$.

Let $v = a_1 + b_1 = a_2 + b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$.

Then $a_1 - a_2 = b_2 - b_1$.

But $a_1 - a_2 \in A$ and $b_2 - b_1 \in B$.

Hence $a_1 - a_2, b_2 - b_1 \in A \cap B$.

Since $A \cap B = \{\mathbf{0}\}, a_1 - a_2 = \mathbf{0}$ and $b_2 - b_1 = 0$ so that $a_1 = a_2$ and $b_1 = b_2$.

Hence the expression of v in the form $a + b$ where $a \in A$ and $b \in B$ is unique.

Conversely suppose that any element in $A + B$ can be uniquely expressed in the form $a + b$ where $a \in A$ and $b \in B$.

We claim that $A \cap B = \{\mathbf{0}\}$.

If $A \cap B \neq \{\mathbf{0}\}$, let $v \in A \cap B$ and $v \neq \mathbf{0}$.

Then $\mathbf{0} = v - v = \mathbf{0} + \mathbf{0}$. Thus $\mathbf{0}$ has been expressed in the form $a + b$ in two different ways which is a contradiction. Hence $A \cap B = \{\mathbf{0}\}$.

Definition:

Let A and B be subspaces of a vector space V. Then V is called the direct sum of A and B if (i) $A + B = V$ (ii) $A \cap B = \{0\}$.

If V is the direct sum of A and B we write

$V = A \oplus B$.

Note. $V = A \oplus B$ iff every element of V can be uniquely expressed in the form $a + b$ where $a \in A$ and $b \in B$.

Examples

1. In $V_3(\mathbf{R})$ let $A = \{(a, b, 0)/a, b \in \mathbf{R}\}$ and $B = \{(0,0, c)/c \in \mathbf{R}\}$. Clearly A and B are subspaces of V and $A \cap B = \{0\}$. Also let $v = (a, b, c) \in V_3(\mathbf{R})$. Then $v = (a, b, 0) + (0,0, c)$ so that $A + B = V_3(\mathbf{R})$.

   Hence $V_3(\mathbf{R}) = A \oplus B$.

2. In $M_3(\mathbf{R})$, let A be the set of all matrices of the form $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ and B be the set of all matrices of the form $\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}$. Clearly A and B are subspaces of $M_2(\mathbf{R})$ and

   $A \cap B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $A + B = M_2(\mathbf{R})$.

   Hence $M_2(\mathbf{R}) = A \oplus R$.

3. Show that the following subsets of $\mathbf{R}^3$ are subspaces. Interpret them geometrically.

   (a) $\{(a, 0, c)/a, c, \in \mathbf{R}\}$

   (b) $\{(a, b, c)/a = b = c\}$

   (c) $\{(a, b, c)/a = b + c\}$

   (d) $\{(a, b, a + b)/a, b \in \mathbf{R}\}$

4. Show that the set of all polynomials in R$[x]$ having at least one rational root is not a subspace of $\mathbf{R}[x]$.

5. Let V be the set of all sequences in $\mathbf{R}$. V is a vector space over $\mathbf{R}$. Let W be the set of all convergent sequences in V and U be the set of all sequences converging to zero. Prove that (a) U is a subspace of W (b) U is a subspace of V. (c) W is a subspace of V.

6. Let W be a subspace of V and U be a subspace of W. Then show that U is a subspace of V.

7. Show that each of the following subsets of $V_3(\mathbf{R})$ is not a subspace.

   (a) $S = \{(x, y, z)/x^2 + y^2 + z^2 \leq 1\}$

   (b) $S = \{(x, y, z)/x + y + z = 1\}$

   (c) $S = \{(x, y, z)/x \geq y \geq z\}$.

Theorem 4:

Let V be a vector space over F and W a subspace of V. Let $V/W = \{W + v/v \in V\}$. Then V/W is a vector space over F under the following operations.

(i) $(W + v_1) + (W + v_2) = W + v_1 + v_2$.

(ii) $\alpha(W + v_1) = W + \alpha v_1$.

Proof:

Since W is a subspace of V it is a subgroup of $(V, +)$.

Since $(V, +)$ is abelian, W is a normal subgroup of $(V, +)$ so that

(i) is a well-defined operation.

Now we shall prove that (ii) is a well-defined operation.

$W + v_1 = W + v_2 \Rightarrow v_1 - v_2 \in W$

$\Rightarrow \alpha(v_1 - v_2) \in W$ (since W is a

$\Rightarrow \alpha v_1 - \alpha v_2 \in W$

$\Rightarrow \alpha v_1 \in W + \alpha v_2$

$\Rightarrow W + \alpha v_1 = W + \alpha v_2$

Hence (ii) is a well-defined operation.

Now, let $W + v_1, W + v_2, W + v_3 \in V/W$

Then $(W + v_1) + [(W + v_2) + (W + v_3)]$

$= (W + v_1) + (W + v_2 + v_3)$

$= W + v_1 + v_2 + v_3$

$= (W + v_1 + v_2) + (W + v_3)$

$= [(W + v_1) + (W + v_2)] + (W + v_3)$

Hence + is associative.

$W + 0 = W \in V/W$ is the additive identity element.

For $(W + v_1) + (W + 0) = W + v_1 = (W + 0) + (W + v_1)$ for all $v_1 \in V$

Also $W - v_1$ is the additive inverse of $W + v_1$.

Hence V/W is a group under + .

Further
$$\begin{aligned}(W+v_1) + (W + v_2) &= W + v_1 + v_2 \\ &= W + v_2 + v_1 \\ &= (W + v_2) + (W + v_1)\end{aligned}$$

Hence V/W is an abelian group.

Now, let $\alpha, \beta \in F$.

$$\begin{aligned}\alpha[(W + v_1) + (W + v_2)] &= \alpha(W + v_1 + v_2) \\ &= W + \alpha(v_1 + v_2) \\ &= W + \alpha v_1 + \alpha v_2 \\ &= (W + \alpha v_1) + (W + \alpha v_2) \\ &= \alpha(W + v_1) + \alpha(W - v_2) \\ (\alpha + \beta)(W + v_1) &= W + (\alpha + \beta)v_1 \\ &= W + \alpha v_1 + \beta v_1 \\ &= (W + \alpha v_1) + (W + \beta v_1) \\ &= \alpha(W + v_1) + \beta(W + v_1)\end{aligned}$$

$$\begin{aligned}\alpha[\beta(W + v_1)] &= \alpha(W + \beta v_1) \\ &= W + \alpha\beta v_1 \\ &= (\alpha\beta)(W + v_1) \\ 1(W + v_1) &= W + 1v_1 \\ &= W + v_1\end{aligned}$$

Hence V/W is a vector space.

The vector space V/W is called the quotient space of V by W

1.3. Linear Transformation

Definition:

Let V and W be vector spaces over a field F. A mapping $T: V \to W$ is called a homomorphism if

(a) $T(u+v) = T(u) + T(v)$ and

(b) $T(\alpha u) = \alpha T(u)$ where $\alpha \in F$ and $u, v \in V$.

A homomorphism T of vector spaces is also called a linear transformation.

(i) If T is 1-1 then T is called monomorphism.

(ii) If T is onto then T is called an epimorphism.

(iii) If T is 1-1 and onto T is called an isomorphism.

(iv) Two vector spaces V and W are said to be isomorphic if there exists an isomorphism T from V to W and we write $V \cong W$.

(v) A linear transformation $T: V \to F$ is called a linear functional.

Examples

1. $T: V \to W$ defined by $T(v) = 0$ for all $v \in V$ is a trivial linear transformation.
2. $T: V \to V$ defined by $T(v) = v$ for all $v \in V$ is the identity linear transformation.
3. Let V be a vector space over a field F and W a subspace of V. Then $T: V \to V/W$ defined by $T(v) = W + v$ is a linear transformation, for,

$$T(v_1 + v_2) = W + (v_1 + v_2)$$
$$= (W + v_1) + (W + v_2)$$
$$= T(v_1) + T(v_2)$$

Also $T(\alpha v_1) = W + \alpha v_1$
$$= \alpha(W + v_1) = \alpha T(v_1)$$

This is called the natural homomorphim from V to V/W. Clearly T is onto and hence. T is an epimorphism.

4. $T: V_3(\mathbf{R}) \to V_3(\mathbf{R})$ defined by

$T(a, b, c) = (a, 0,0)$ is a linear transforma. tion.

5. Let V be the set of all polynomials of degree $\leq n$ in $\mathbf{R}[x]$ including the zero polynomial. $T: V \to V$ defined by $T(f) = \frac{df}{dx}$ is a linear transformation.

$$\text{For, } T(f+g) = \frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx}$$
$$= T(f) + T(g)$$

Also $T(\alpha f) = \frac{d(\alpha f)}{dx} = \alpha \frac{df}{dx} = \alpha T(f)$.

6. Let V be as in Example 5.

Then $T: V \to V_{n+1}(\mathbf{R})$ defined by $T(a_0 + a_1 x + \cdots \ldots + a_0 x^n)$

$= (a_0, a_1, \ldots, a_n)$ is a linear transformation

For, let $f = a_0 + a_1 x + \cdots + a_n x^n$ and $g = b_0 + b_1 x + \cdots + b_n x^n$.

Then $f + g = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)r^n$

$$\therefore\ T(f+g) = \Big((a_0 + b_0), (a_1 + b_1) \ldots (a_n + b_n)\Big)\Big)$$

$$= (a_0, a_1, \ldots, a_n) + (b_0, b_1, \ldots, b_n)$$

$$= T(f) + T(g)$$

$$\text{Also } T(\alpha f) = (\alpha a_0, \alpha a_1, \ldots, \ldots, \alpha a_n)$$
$$= \alpha(a_0, a_1, \ldots, \ldots, a_n)$$
$$= \alpha T(f)$$

Clearly T is $1 - 1$ and onto and hence T is an isomorphism.

7. Let V denote the set of all sequences in R, $T: V \to V$ defined by

$T(a_1, a_2, \ldots, a_n, \ldots) = (0, a_1, a_2, \ldots a_n, \ldots)$ is a linear transformation.

8. $T: \mathbf{R}^2 \to \mathbf{R}^2$ defined by

$T(a,b) = (2a - 3b, a + 4b)$ is a linear transformation.

Let $u = (a,b)$ and $v = (c,d)$ and $\alpha \in \mathbf{R}$.

$\therefore\ T(u+v) = T((a,b) + (c,d))$

$$=T(a+c, b+d)$$
$$= (2(a+c) - 3(b+d), (a+c) + 4(b+d))$$

$$= (2a + 2c - 3b - 3d, a + c + 4b + 4d)$$

$$= (2a - 3b + 2c - 3d, a + 4b + c + 4d)$$

$$= (2a - 3b, a + 4b) + (2c - 3d, c + 4d)$$

$$=T(a,b) + (c,d)$$
$$=T(u) + T(v)$$

Also, $T(\alpha u) = T(\alpha(a,b))$

$$= T(\alpha a, \alpha b)$$
$$= (2\alpha a - 3\alpha b, \alpha a + 4\alpha b)$$
$$= \alpha(2a - 3b, a + 4b)$$
$$= \alpha T(a,b)$$
$$= \alpha T(u).$$

$\therefore\ T$ is a linear transformation.

Theorem 5:

Let $T: V \to W$ be a linear transformation. Then $T(V) = \{T(v)/v \in V\}$ is a subspace of W.

Proof:

Let w_1 and $w_2 \in T(V)$ and $\alpha \in F$. Then there exist $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$.

Hence $w_1 + w_2 = T(v_1) + T(v_2)$

$= T(v_1 + v_2) \in T(V)$.

Similarly, $\alpha w_1 = \alpha T(v_1) = T(\alpha v_1) \in T(V)$.

Hence $T(V)$ is subspace of W.

Definition:

Let V and W be vector spaces over a field F and $T: V \to W$ be a linear transformation. Then the kernel of T is defined to be

$\{v / v \in V \text{ and } T(v) = 0\}$ and is denoted by ker T.

Thus ker$T = \{v / v \in V \text{ and } T(v) = 0\}$.

For example, in example 1 , ker $T = V$.

In example 2, ker$T = \{\mathbf{0}\}$.

In example 5 , ker$T$ is the set of all constant polynomials.

Note. Let $T: V \to W$ be a linear transformation. Then T is a monomorphism iff ker$T = \{0\}$

Theorem 6: (Fundamental theorem of homomorphism)

Let V and W be vector spaces over a field F and $T: V \to W$ be an epimorphism.

Then (i) ker$T = V_1$ is a subspace of V and

(ii) $\frac{V}{V_1} \cong W$.

Proof:

(i) Given $V_1 = \text{ker}T = \{v / v \in V \text{ and } T(v) = 0\}$

Clearly $T(\mathbf{0}) = \mathbf{0}$. Hence $\mathbf{0} \in \text{ker}T = V_1$.

$\therefore$ V_1 is nonempty subset of V.

Let $u, v \in \text{ker}T$ and $\alpha, \beta \in F$.

$\therefore$ $T(u) = \mathbf{0}$ and $T(v) = \mathbf{0}$. Now

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

$\therefore\ \alpha u + \beta v \in \ker T$

$\therefore$ By theorem 3, ker T is a subspace of V.

(ii) We define a map $\varphi: \frac{V}{V_1} \to W$ by

$\varphi(V_1 + v) = T(v)$.

φ is well defined. Let $V_1 + v = V_1 + w$.

$\therefore\ v \in V_1 + w$
$\therefore\ v = v_1 + w$ where $v_1 \in V_1$
$\therefore\ T(v) = T(v_1 + w) = T(v_1) + T(u$
$= \mathbf{0} + T(w) = 7$
$\therefore\ \varphi(V_1 + v) = \varphi(V_1 + w)$

φ is $I - I$.

$\varphi(V_1 + v) = \varphi(V_1 + w)$
$\Rightarrow T(v) = T(w)$
$\Rightarrow T(v) - T(w) = \mathbf{0}$
$\Rightarrow T(v) + T(-w) = \mathbf{0}$
$\Rightarrow T(v - w) = \mathbf{0}$
$\Rightarrow v - w \in \ker T = V_1$
$\Rightarrow v \in V_1 + w$
$\Rightarrow V_1 + v = V_1 + w$.

φ is onto.

Let $w \in W$. Since T is onto there exists $v \in V$ such that $T(v) = w$.

$\therefore\ \varphi(V_1 + v) = w$.

$\therefore \varphi(V_1 + v) = w$.

φ is a homomorphism.

$$\varphi[(V_1+v)+(V_1+w)] = \varphi[V_1+(v+w)]$$
$$= T(v+w)$$
$$= T(v)+T(w)$$
$$= \varphi(V_1+v)+\varphi(V_1+w)$$

Also $\varphi[\alpha(V_1+v)] = \varphi(V_1+\alpha v)$
$$= T(\alpha v)$$
$$= \alpha T(v)$$
$$= \alpha T(V_1+v)$$

$\therefore$ φ is an isomorphism from $\frac{V}{V_1}$ onto W.

$\therefore \dfrac{V}{V_1} \cong W$.

Theorem 7:

Let V be a vector space over a field F.

Let A and B be subspaces of V.

Then $\frac{A+B}{A} \cong \frac{B}{A\cap B}$.

Proof:

We know that $A+B$ is a subspace of V containing A.

Hence $\frac{A+B}{A}$ is also a vector space over F,

An element of $\frac{A+B}{A}$ is of the form $A+(a+b)$ where $a \in A$ and $b \in B$. But $A+a = A$.

Hence an element of $\frac{A+B}{A}$ is of the form $A+b$.

Now, consider $f: B \to \frac{A+B}{A}$ defined by

$f(b) = A+b$.

Clearly f is onto.

$$
\text{Also} \quad \begin{aligned} f(b_1 + b_2) &= A + (b_1 + b_2) \\ &= (A + b_1) + (A + b_2) \\ &= f(b_1) + f(b_2) \\ \text{and } f(\alpha b_1) &= A + \alpha b_1 \\ &= \alpha(A + b_1) \\ &= \alpha f(b_1) \end{aligned}
$$

Hence f is a linear transformation.

Let K be the kernel of f.

Then $K = \{b / b \in B, A + b = A\}$.

Now, $A + b = A$ iff $b \in A$. Hence $K = A \cap B$.

Hence by theorem 6, $\frac{B}{A \cap B} \cong \frac{A+B}{A}$.

Theorem 8:

Let V and W be vector spaces over a field F. Let $L(V, W)$ represent the set of all linear transformations from V to W. Then $L(V, W)$ itself is a vector space over F under addition and scalar multip iication defined by

$(f + g)(v) = f(v) + g(v)$ and $(\alpha f)(v) = \alpha f(v)$.

Proof:

Let $f, g \in L(V, W)$ and $v_1, v_2 \in V$.

Then $(f + g)(v_1 + v_2) = f(v_1 + v_2) + g(v_1 + v_2)$

$$
\begin{aligned} &= f(v_1) + f(v_2) + g(v_1) + g(v_2) \\ &= f(v_1) + g(v_1) + f(v_2) + g(v_2) \\ &= (f + g)(v_1) + (f + g)(v_2) \end{aligned}
$$

$$
\text{Also} \quad \begin{aligned} (f + g)(\alpha v) &= f(\alpha v) + g(\alpha v) \\ &= \alpha f(v) + \alpha g(v) \\ &= \alpha[f(v) + g(v)] \\ &= \alpha(f + g)(v). \end{aligned}
$$

Hence $(f + g) \in L(V, W)$.

Now,
$$\begin{aligned}(\alpha f)(v_1+v_2) &= (\alpha f)(v_1)+(\alpha f)(v_2)\\ &= \alpha f(v_1)+\alpha f(v_2)\\ &= \alpha[f(v_1)+f(v_2)]\\ &= \alpha f(v_1+v_2)\end{aligned}$$

Also $(\alpha f)(\beta v) = \alpha[f(\beta v)] = \alpha[\beta f(v)]$
$$= \beta[\alpha f(v)] = \beta[(\alpha f)(v)].$$

Hence $\alpha f \in L(V,W)$.

Addition defined on $L(V,W)$ is obviously commutative and associative.

The function $f: V \to W$ defined by $f(v) = 0$ for all $v \in V$ is clearly a linear transformation and is the additive identity of $L(V,W)$.

Further $(-f): V \to W$ defined by

$(-f)(v) = -f(v)$ is the additive inverse of f.

Thus $L(V,W)$ is an abelian group under addition.

The remaining axioms for a vector space can be easily verified.

Hence $L(V,W)$ is a vector space over F.

Exercises

1. Let V and W be vector spaces over a field F. Show that a mapping $T: V \to W$ is a linear transformation iff

   $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$ for all $v_1, v_2 \in V$ and $\alpha.\beta \in F$.
2. For each of the linear transformations T given in examples of section 1.3 determine whether it is one-one and onto. Also find its kernel.
3. Find the kernel of the following linear transformations.

   (i) $T: V_4(R) \to V_4(R)$ defined by $T(x_1, x_2, x_3, x_4) = (x_1, 0, x_3, 0)$.

   (ii) $T: V_3(R) \to V_3(R)$ defined by $T(a, b, c) = (a, b, 0)$.

4. Let V and W be two vector spaces over a field F and $T: V \to W$ be a linear transformation. Show that if V_1 is a subspace of $V. T(V_1)$ is a subspace of W and if W_1 is a subspace of $W, T^{-1}(W_1)$ is a subspace of V.
5. Prove that $T: V_3(R) \to R$ defined by $T(x, y, z) = x^2 + y^2 + z^2$ is not a linear transformation.
6. Determine which of the following statements are true.

   (a) Any line passing through the origin is a subspace of $V_3(R)$

   (b) Any plane passing through the origin is a subspace of $V_3(R)$

   (c) If A and B are subspaces of V then

   (i) $A + B$ is a subspace of V. (ii) A is a subspace of $A + B$.

   (iii) B is a subspace of $A + B$.

   (iv) Every element of $A + B$ can be uniquely written in the form $a + b$ where $a \in A$ and $b \in B$. (v) (iv) is true iff $A \cap B = \{0\}$

Answers.

3. 1. Not 1-1, not onto; kernel V. 2.1-1 and onto; kernel $= \{0\}$. 3. Not 1-1, onto kernel $= W$ 4. Not 1-1, not onto;

   kernel $= \{(0, b, c)/b, c \in \mathbf{R}\}$ 5. Not 1-1 not onto; kernel = all constant polynomials 6. 1-1 and onto; kernel $= \{0\}$ 7. 1-1, not onto; kernel $= \{0\}$
2. (i) Not 1-1, not onto;

   kernel $\{(0, a, 0, b)/a, b \in \mathbf{R}\}$

   (ii) Not 1-1, not onto;

   kernel $\{(0,0, c)/c \in \mathbf{R}\}$
3. (a) T (b) T (c) [i] T [ii] T [iii] T [iv] F [v] T.

UNIT II

Span of a set–Linear Dependence and Independence–Basis and Dimension.

Chapter 2: Sections-2.1 to 2.3

2.1. Span of a set

Definition:

Let V be a vector space over a field F.

Let $v_1 v_2, \dots \dots \dots, v_n \in V$. Then an element of the form $\alpha_1 v_1 + \alpha_2 v_2 + \cdots \dots + \alpha_n v_n$ where $\alpha_i \in F$ is called a linear combination of the vectors $v, v_2, \dots \dots, v_n$.

Definition:

Let S be a non-empty subset of a vector space V. Then the set of all linear combinations of finite sets of elements of S is called the linear span of S and is denoted by $L(S)$.

Note. Any element of $L(S)$ is of the form $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ where $\alpha_1, \alpha_2, \dots, \alpha_n \in F^*$.

Theorem 1:

Let V be a vector space over a field F and S be a non-empty subset of V. Then

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

(ii) $S \subseteq L(S)$.

(iii) If W is any subspace of V such that $S \subseteq W$, then $L(S) \subseteq W$

(ie), $L(S)$ is the smallest subspace of V containing S.

Proof:

(i) Let $v, w \in L(S)$ and $\alpha, \beta \in F$.

Then $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots . + \alpha_n v_n$ where $v_i \in S$ and $\alpha_i \in F$.

Also, $w = \beta_1 w_1 + \beta_2 w_2 + \cdots \dots + \beta_m w_m$ where $w_j \in S$ and $\beta_j \in F$.

Now, $\alpha v + \beta w = \alpha(\alpha_1 v_1 + \alpha_2 v_2 + \cdots \ldots + \alpha_n v_n) + \beta(\beta_1 w_1 + \beta_2 w_2 + \cdots \ldots + \beta_m w_m)$.

$= (\alpha\alpha_1)v_1 + \cdots \ldots + (\alpha\alpha_n)v_n + (\beta\beta_1)w_1 + \cdots \ldots + (\beta\beta_m)w_m$.

$\therefore$ $\alpha v + \beta w$ is also a linear combination of a finite number of elements of S.

Hence $\alpha v + \beta w \in L(S)$.

$\therefore$ $L(S)$ is a subspace of V.

(ii) Let $u \in S$.

Then $u = 1u \in L(S)$. Hence $S \subseteq L(S)$.

(iii) Let W be any subspace of V such that $S \subseteq W$. We claim that $L(S) \subseteq W$.

Let $u \in L(S)$

Then $u = \alpha_1 u_1 + \alpha_2 u_2 + \cdots \ldots \ldots + \alpha_n u_n$ where $u_i \in S$ and $\alpha_i \in F$.

Since $S \subseteq W$ we have $u_1, u_2, \ldots \ldots, u_n \in W$.

$\therefore$ $a_1 u_1 + \alpha_2 u_2 + \cdots \ldots + \alpha_n u_n \in W$. (since u is a subspace of V.)

$\therefore$ $u \in W$. Hence $L(S) \subseteq W$.

Note. $L(S)$ is called the subspace spanned (gener. ated) by the set S.

Examples

1. In $V_3(R)$ let $e_1 = (1,0,0), e_2 = (0,1,0)$ and $e_3 = (0,0,1)$

   (a) Let $S = \{e_1, e_2\}$

Then $L(S) = \{\alpha e_1 + \beta e_2 / \alpha, \beta \in \mathrm{R}\} = \{(\alpha, \beta, 0)/\alpha, \beta \in \mathbf{R}\}$

(b) Let $S = \{e_1, e_2, e_3\}$. Then

$$\begin{aligned} L(S) &= \{\alpha e_1 + \beta e_2 + \gamma e_3 / \alpha, \beta, \gamma \in \mathbb{R}\} \\ &= \{(\alpha, \beta, \gamma)/\alpha, \beta, \gamma \in \mathbb{R}\} \\ &= V_3(\mathbb{R}) \end{aligned}$$

Thus $V_3(\mathbf{R})$ is spanned by $\{e_1, e_2, e_3\}$.

2. In $V_n(\mathbb{R})$, let $e_1 = (1,0, \ldots 0)$, $e_2 = (0,1,0, \ldots,0), \ldots, e_n = (0,0, \ldots,1)$ Let $S = \{e_1, e_2 \ldots \ldots, e_n\}$. Then

$$
\begin{aligned}
L(S) &= \{\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n / \alpha_i \in \mathbb{R}) \\
&= \{(\alpha_1, \alpha_2, \dots \dots, a_n) / \alpha_i \in \mathbb{R}\} \\
&= V_n(\mathbb{R})
\end{aligned}
$$

Thus $V_n(\mathbf{R})$ is spanned by $\{e_1, e_2, \dots, e_n\}$.

Theorem 2:

Let V be a vector space over a field F.

Let $S, T \subseteq V$. then

(a) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$.

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

(c) $L(S) = S$ iff S is a subspace of V.

Proof:

(a) Let $S \subseteq T$. Let $s \in L(S)$.

Then $s = \alpha_1 s_1 + \alpha_2 s_2 + \cdots .. + \alpha_n s_n$ where $s_i \in S$ and $\alpha_i \in F$.

Now, since $S \subseteq T, s_i \in T$.

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \cdots \dots + \alpha_n s_n \in L(T)$.

Thus $L(S) \subseteq L(T)$.

(b) Let $s \in L(S \cup T)$.

Then $s = \alpha_1 s_1 + \alpha_2 s_2 + \cdots . + \alpha_n s_n$ where $s_i \in S \cup T$ and $\alpha_i \in F$.

Without loss of generality we can assume that

$s_1, s_2, \dots, s_m \in S$

And $s_{m+1}, \dots \dots \dots, s_n \in T$.

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_m s_m \in L(S)$ and

$\alpha_{m+1} s_{m+1} + \cdots \dots \dots + \alpha_n s_n \in L(T)$.

$\therefore s = (\alpha_1 s_1 + \cdots .. + \alpha_m s_m)$

$+(\alpha_{m+1}, s_{m+1} + \cdots + \alpha_n s_n) \in L(S) + L(T)$.

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

Also by (a) $L(S) \subseteq L(S \cup T)$ and $L(T) \subseteq L(S \cup T)$.

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

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

(c) Let $L(S) = S$. By theorem $5.9 L(S) = S$ is a subspace of V.

Conversely, let S be a subspace V. Then the smallest subspace containing S is S itself.

Hence $L(S) = S$.

Corollary. $L[L(S)] = L(S)$.

Exercises

1. Find $L(S)$ in the following cases.

   (a), $S = \{(1,0), (0,1)\}$ in $V_2(\boldsymbol{R})$

   (b) $S = \{(1,0,0)(2,0,0), (3,0,0)\}$ in $V_3(\boldsymbol{R})$

   (c) $S = \{(1,2,3), (2,3,1), (3,1,2)\}$ in $V_3(\boldsymbol{R})$

   (d) $S = \{(2,0)\}$ in $V_2(\mathbf{R})$.

   (e) $S = \{1, x, x^2, \dots \dots, x^n\}$ in $\mathbf{R}[x]$

   (f) $S = \{(1,2,3)\}$ in $\mathbf{Z}_5 \times \mathbf{Z}_5 \times \mathbf{Z}_5$ over Z_5.

   (g) $S = \{(0,1,2), (1,2,0)\}$ in $\mathbf{Z}_3 \times \mathbf{Z}_3 \times \mathbf{Z}_3$ over $\mathbf{Z}_3$.

   (h) $S = \left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right\}$ in $M_2(\mathbf{R})$.

2. Let V be a vector space over a field F and $S = \{v_1, v_2, \dots, v_n\} \subseteq V$.

   Let $S_1 = \{\alpha_1 v_1, \alpha_2 v_2, \dots, \alpha_n v_n\}$ where $\alpha_i \in F - \{0\}$.

   Let $S_2 = \{v_1 + \alpha v_2, v_2, v_3, \dots, v_n\}$ where $\alpha \in F$.

   Show that $L(S) = L(S_1) = L(S_2)$.

3. Show that in $V_2(R)$, $(3,7) \in L(\{(1,2), (0,1)\})$.

Answers:

1.(a) V_2(R) (b) $\{(x, 0,0)/x \in \mathbf{R}\}$ (c) V_3(R) (d) $\{(x, 0)/x \in \mathbf{R}\}$

(e) The set of all polynomials of degree $\leq n$ and zero polynomial.

(f) $\{(0,0,0), (1,2,3), (2,4,1)(3,1,4), (4,3,2))\}$

(g) $\{(0,0,0), (0,1,2), (1,2,0), (1,0,2),$

$(0,2,1), (2,1,0), (2,0,1), (1,1,1), (2,2,2)\}$

(h) $\left\{\begin{bmatrix} a & 0 \\ b & 0 \end{bmatrix} \middle| \ a, b \in \mathbf{R}\right\}$

2.2. Linear Independence:

In $V_3(\mathbf{R})$, let $S = \{e_1, e_2, e_3\}$. We have seen that $L(S) = V_3(\mathbf{R})$.

Thus S is a subset of $V_3(\mathbf{R})$ which spans the whole space $V_3(\mathbf{R})$

Definition:

Let V be a vector space over a field F、 V is said to be finite dimensional if there exists a finite subset S of V such that $L(S) = V$.

Examples

1. $V_3(\mathbf{R})$ is a finite-dimensional vector space.
2. $V_n(\mathbf{R})$ is a finite dimensional vector space. since $S = \{e_1, e_2, \dots \dots, e_n\}$ is a finite subset of $V_n(\mathbf{R})$ such that $L(S) = V_n(\mathbf{R})$. In general if F is any field $V_n(F)$ is a finite dimensional vector space over F.
3. Let V be the set of all polynomials in $F[x]$ of degree $\leq n$. Let $S = \{1, x, x^2, \dots \dots, x^n\}$. Then $L(S) = V$ and hence V is finitedimensional.
4. $\mathbf{C}$ is a finite-dimensional vector space over $\mathbf{R}$, since $L(\{1, i\}) = \mathbf{C}$.
5. In $M_2(\mathbf{R})$ consider the set S consisting of the matrices

$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

$C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Then

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aA + bB + cC + dD$

Hence $L(S) = M_2(\mathbf{R})$ so that $M_2(\mathbf{R})$ is finitedimensional.

Note:

All the vector spaces, we have considered above are finite dimensional. However, there are vector spaces which cannot be spanned by a finite number of vectors. For example, consider $\mathbf{R}[x]$. Let S be any finite subset of $\mathbf{R}[x]$. Let f be a polynomial of maximum degree in S. Let $\deg f = n$. Then any element of $L(S)$ is a polynomial of degree $\leq n$ and hence $L(S) \neq \mathbf{R}[x]$ Thus $\mathbf{R}[x]$ is not finite-dimensional.

Throughout the rest of this chapter all the vector spaces we consider are finite dimensional."

Although we have defined what is meant by a finite dimensional space, we have not yet defined what is meant by the dimension of a vector space. We now proceed to introduce the concepts necessary to define the dimension of a finite dimensional vector space.

Consider the vectors $e_1 = (1,0,0)$,

$$e_2 = (0,1,0), e_3 = (0,0,1) \text{ in } V_3(\mathbf{R}).$$

Suppose that $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$.

Then $(\alpha_1, 0,0) + (0, \alpha_2, 0) + (0,0, \alpha_3) = (0,0,0)$.

$$\therefore \ (\alpha_1, \alpha_2, \alpha_3) = (0,0,0)$$

$$\therefore \ \alpha_1 = \alpha_2 = \alpha_3 = 0$$

(i.e.) $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$ iff $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Thus, a linear combination of the vectors, e_1, e_2 and e_3 will yield the zero vector iff all the coefficients are zero.

Definition:

Let V be a vector space over a field F, A finite set of vectors $v_1, v_2, \ldots \ldots, v_n$ in V is said to be linearly independent if

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots \ldots + \alpha_n v_n = 0 \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$

If $v_1, v_2, \ldots \ldots, v_n$ are not linearly independent, then they are said to be linearly dependent.

Note. If $v_1, v_2, \ldots, v_n$ are linearly dependent, then there exist scalars $\alpha_1, \alpha_2, \ldots \ldots, \alpha_n$ not all zero, such that $\alpha_1 v_1 + \cdots + \alpha_r v_n = \mathbf{0}$.

Examples

1. In $\dot{V}_n(F), \{e_1, e_2, \ldots., e_n\}$ is a linearly independent set of vectors, for,

$$\begin{aligned}
&\alpha_1 e_1 + \alpha_2 e_2 + \cdots \ldots + \alpha_n e_n = 0 \\
&\Rightarrow \alpha_1(1,0, \ldots \ldots, 0) + \alpha_2(0,1, \ldots \ldots, 0) \\
&+ \cdots \ldots + \alpha_n(0,0, \ldots \ldots, 1) \\
&= (0,0, \ldots \ldots, 0) \\
&\Rightarrow (\alpha_1 \alpha_2, \ldots \ldots, \alpha_n) = (0,0, \ldots \ldots, 0) \\
&\Rightarrow \alpha_1 = \alpha_2 = \cdots \ldots = \alpha_n = 0
\end{aligned}$$

2. In $V_3(\mathbf{R})$ the vectors $(1,2,1), (2,1,0)$ and $(1,-1,2)$ are linearly independent. For, let

$$\alpha_1(1,2,1) + \alpha_2(2,1,0) + \alpha_3(1,-1,2) = (0,0,0)$$

$$\therefore\ (\alpha_1 + 2\alpha_2 + \alpha_3, 2\alpha_1 + \alpha_2 - \alpha_3, \alpha_1 + 2\alpha_3) = (0,0,0)$$

$$\begin{aligned}
\therefore\ \alpha_1 + 2\alpha_2 + \alpha_3 &= 0 \ldots \ldots (1) \\
2\alpha_1 + \alpha_2 - \alpha_3 &= 0 \ldots \ldots (2) \\
\alpha_1 + 2\alpha_3 &= 0 \ldots \ldots (3)
\end{aligned}$$

Solving equations (1), (2) and (3) we get $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

$\therefore$ The given vectors are linearly independent

3. In $V_3(\mathrm{R})$ the vectors $(1,4,-2), (-2,1,3)$ and $(-4,11,5)$ are linearly dependent. For, let

$\alpha_1(1,4,-2)+\alpha_2(-2,1,3)+\alpha_3(-4,11,5)=(0,0,0)$

$\therefore\ \alpha_1-2\alpha_2-4\alpha_3=0\ \ldots\ldots$ (1)
$4\alpha_1+\alpha_2+11\alpha_3=0\ \ldots\ldots$ (2)
$-2\alpha_1+3\alpha_2+5\alpha_3=0\ \ldots\ldots$ (3)

From (1) and (2),

$\frac{\alpha_1}{-18}=\frac{\alpha_2}{-27}=\frac{\alpha_3}{9}=k$ (say)

$\therefore\ \alpha_1=-18k, \alpha_2=-27k, \alpha_3=9k.$

These values of α_1,α_2 and α_3, for any k satisfy (3) also.

Taking $k=1$ we get

$\alpha_1=-18, \alpha_2=-27, \alpha_3=9$ as a non-trivial solution.

Hence the three vectors are linearly dependent.

4. Let V be a vector space over a field F. Then any subset S of V containing the zero vector is linearly dependent.

Proof. Let $S=\{0, v_1, \ldots\ldots, v_n\}$

Clearly $\alpha 0+0v_1+0v_2+\cdots\ldots+0v_n=0$ where is any element of F. Hence for any $\alpha\neq 0$, we get non-trivial linear combination of vectors in S giving a zero vector. Hence S is linearly dependent.

Exercises:

1. Determine whether the following sets of vector are linearly independent or linearly dependent in $V_3(\mathbf{R})$.
   (a) $\{(1,0,0),(0,1,0),(1,1,0)\}$.
   (b) $\{(1,2,3),(2,3,1)\}$.
   (c) $\{(1,2,3),(4,1,5),(-4,6,2)\}$.
   (d) $\{(0,0,0),(2,5,3),(-1,0,6)\}$.
   (e) $\{(1,0,0),(1,1,0),(1,1,1),(0,1,0)\}$.

2. Determine whether the following sets of vectors are linearly independent or not.

   (a) $\{(1,1,0,0), (0,0,1,1)(1,0,0,4). (0,0,0,2)\}$ in $V_4(R)$.

   (b) $\{(2i, 1,0), (2, -i, 1), (0,1, +i, -i)\}$ in V_3(C).

   (c) $\{(\pi, 0,0), (0, e, 0), (0,0, \sqrt{5})\}$ in V_3(R).

   (d) $V =$ the set of all polynomials of degree $\leq n$ in R$[x]$ and

   $S = \{1, x, x^2, \dots, \dots, x^n\}$

   (e) $\left\{\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}\right\}$ in $M_2(R)$

3. In $V_3(Z_5)$ determine whether the following sets of vectors are linearly dependent.

(a) $\{(1,3,2), (2,1,3)\}$

(b) $\{(1,1,2), (2,1,0), (0,4,1)\}$.

4. In V_2(R) prove that the vectors (a, b) and (c, d) are linearly dependent iff $ad - bc = 0$.

5. Let $\{v_1, v_2, v_3\}$ be a linearly independent set of vectors in V_3(R).

Show that

(a) $\{v_1 + v_2, v_2 + v_3, v_3 + v_1\}$ is linearly independent.

(b) $\{2v_1 + v_2, v_1 + v_2, v_1 - v_3\}$ is linearly independent.

6. If the vectors $(0,1, a), (1, a, 1)$ and $(a, 1,0)$ of V_3(R) are linearly dependent then find the value of a.

Answers.

1. (b) is linearly independent.
2. (a), (b), (c), (d) and (e) are linearly independent.
3. (a) is linearly independent 6. $a = 0, \pm\sqrt{2}$.

Theorem 3:

Any subset of a linearly independent set is linearly independent.

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

Let $S = \{v_1, v_2, \dots, v_n\}$ be a linearly independent set.

Let S' be a subset of S. Without loss of generality we take $S' = \{v_1, v_2, \dots \dots, v_k\}$ where $k \leq n$.

Suppose S' is a linearly dependent set. Then there exist $\alpha_1, \alpha_2, \dots \dots \dots, a_k$ in F not all zero, such that

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots \dots \dots + \alpha_k v_k = 0.$

Hence $\alpha_1 v_1 + \alpha_2 v_2 + \cdots \dots + \alpha_k v_k + 0v_{k+1} + \cdots \dots + 0v_n = 0$ is a non-trivial linear combination giving the zero vector.

Here S is a linearly dependent set which is a contradiction.

Hence S' is linearly independent.

Theorem 4:

Any set containing a linearly dependent set is also linearly dependent.

Proof:

Let V be a vector space. Let S be a linearly dependent set. Let $S' \supset S$.

If S' is linearly independent S is also linearly independent (by theorem 5.11) which is a contradiction. Hence S' is linearly dependent.

Theorem 5:

Let $S = \{v_1, v_2, \dots., v_n\}$ be a linearly independent set of vectors in a vector space V over a field F. Then every element of $L(S)$ can be uniquely written in the form $\alpha_1 v_1 + \alpha_2 v_2 + \cdots \dots + \alpha_n v_n$, where $\alpha_i \in F$.

Proof:

By definition every element of $L(S)$ is of the form

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots \ldots + \alpha_n v_n$.

Now, let $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 +$

Hence $(\alpha_i - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots \ldots + (\alpha_n - \beta_n)v_n = 0$.

Since S is a linearly independent set, $\widehat{\alpha_\iota} - \beta_i = 0$ for all i.

$\therefore$ $\alpha_i = \beta_i$ for all i. Hence the theorem?

Theorem 6:

$S = \{v_1, v_2, \ldots, v_n\}$ is a linearly dependent set of vectors in V iff there exists a vector $v_k \in S$ such that v_k is a linear combination of the preceding vectors $v_1, v_2, \ldots \ldots, v_{k-1}$.

Proof:

Suppose $v_1, v_2, \ldots, v_n$ are linearly dependent.

Then there exist $\alpha_1, \alpha_2, \ldots \ldots, \alpha_n \in F$, not all zero, such that

$\alpha_1 v_1 + \alpha_2 v_2 + \cdots .. + \alpha_n v_n = 0$.

Let k be the largest integer for which $\alpha_k \neq 0$.

Then $\alpha_1 v_1 + \cdots \ldots . + \alpha_k v_k = 0$.

$\therefore \alpha_k v_k = -\alpha_1 v_1 - \alpha_2 v_2 \ldots \ldots \ldots - \alpha_{k-1} v_{k-1}$

$\therefore v_k = (-\alpha_k^{-1}\alpha_1)v_1 + \cdots \ldots + (-\alpha_k^{-1}\alpha_{k-1})v_{k-1}$

$\therefore$ v_k is a linear combination of the preceding vectors,

Conversely, suppose there exists a vector v_k such that

. $v_k = \alpha_1 v_1 + \cdots \ldots + \alpha_{k-1} v_{k-1}$.

Hence $-\alpha_1 v_1 - \cdots - \alpha_{k-1} v_{k-1} + v_k + 0 v_{k+1} +$

$$\cdots + 0 v_n = 0$$

Since the coefficient of $v_k = 1$, we have

$S = \{v_1, \ldots \ldots, v_n\}$ is linearly dependent.

Example.

In $V_3(\mathbf{R})$, let $S = \{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$

Here $(1,1,1) = (1,0,0) + (0,1,0) + (0,0,1)$.

Thus $(1,1,1)$ is a linear combination of the preceding vectors. Hence S is a linearly dependent set.

Theorem 7:

Let V be a vector space over F. Le $S = \{v_1, v_2, \dots \dots, v_n\}$ and $L(S) = W$. Then there exists a linearly independent subset S' of S such that $L(S') = W$.

Proof:

Let $S = \{v_1, v_2, \dots \dots, v_n\}$.

If S is linearly independent there is nothing to prove

If not, let v_k be the first vector in S which is a linear combination of the preceding vectors.

Let $S_1 = \{v_1, v_2, \dots \dots, v_{k-1}, v_{k+1}, \dots \dots, v_n\}$.

(i.e) S_1 is obtained by deleting the vector v_k from S.

We claim that $L(S_1) = L(S) = W$.

Since $S_1 \subseteq S, L(S_1) \subseteq L(S)$. (refer theorem 5.10).

Not, let $v \in L(S)$.

Then $v = \alpha_1 v_1 + \cdots .. + \alpha_k v_k + \cdots + \alpha_n v_r$.

Now, v_k is a linear combination of the preceding vectors.

Let $v_k = \beta_1 v_1 + \cdots \dots + \beta_{k-1} v_{k-1}$.

Hence $v = \alpha_1 v_1 + \cdots \dots + \alpha_{k-1} v_{k-1} + \alpha_k (\beta_1 v^{1-} \dots .. + \beta_{k-1} v_{k-1}) + \alpha_{k+1} v_{k+1} + \cdots \dots + \alpha_n v_n$.

v can be expressed as a linear combination of the vectors of S_1 so that $v \in L(S_1)$.

Hence $L(S) \subseteq L(S_1)$

Thus $L(S) = L(S_1) = W$.

Now, if S_1 is linearly independent, the proof is complete.

If not, we continue the above process of removing a vector from S_1, which is a linear combination of the preceeding vectors until we arrive at a linearly independent subset S' of S such that $L(S') = W$.

2.3. Basis and Dimension:

Definition:

A linearly independent subset S of a vector space V which spans the whole space V is called a basis of the vector space.

Theorem 8:

Any finite-dimensional vector space V contains a finite number of linearly independent vectors which span V. (ie) A finite dimensional vector space has a basis consisting of a finite number of vectors.

Proof:

Since V is finite dimensional there exists a finite subset S of V such that $L(S) = V$.

By theorem 5.15 this set S contains a linearly independent subset $S' = \{v_1, v_2, \dots \dots, v_n\}$ such that

$L(S') = L(S) = V$.

Hence S' is a basis for V.

Theorem 9:

Let V be a vector space over a field F. Then $S = \{v_1, v_2, \dots, v_n\}$ is a basis for V iff every element of V can be uniquely expressed as a linear combination of element of S.

Proof:

Let S be a basis for V.

Then by definition S is linearly independent and $L(S) = V$. Hence by theorem 5 every element of V can be uniquely expressed as a linear combination of elements of S.

Conversely, suppose every element of V can be uniquely expressed as a linear combination of elements of S.

Clearly $L(S) = V$.

Now, let $\alpha_1 v_1 + \alpha_2 v_2 + \cdots .. + \alpha_n v_n = 0$.

Also, $0v_1 + 0v_2 + \cdots + 0v_n = 0$.

Thus we have expressed 0 as a linear combination of vectors of S in two ways.

∴ By hypothesis $\alpha_1 = \alpha_2 = \cdots .. = \alpha_n = 0$.

Hence S is linearly independent. Hence S is a basis.

Examples

$S = \{(1,0,0), (0,1,0), (0,0,1)\}$ is a basis for $V_3(\mathbf{R})$ for, $(a, b, c) = a(1,0,0) + b(0,1,0) + c(0,0,1)$.

∴ Any vector (a, b, c) of $V_3(\mathbf{R})$ has been expressed uniquely as a linear combination of the elements of S and hence S is a basis for $V_3(\mathbf{R})$

2. $S = \{e_1, e_2, \dots \dots, e_n\}$ is a basis for $V_n(F)$. This is known as the standard basis for $V_n(F)$.

3. $S = \{(1,0,0), (0,1,0), (1,1,1)\}$ is a basis for $V_3(\mathbf{R})$.

Proof:

We shall show that any element (a, b, c) of $V_3(\mathbf{R})$ can be uniquely expressed as a linear combination of the vectors of S.

Let $(a, b, c) = \alpha(1,0,0) + \beta(0,1,0) + \gamma(1,1,1)$

Then $\alpha + \gamma = a, \beta + \gamma = b, \gamma = c$.

Hence $\alpha = a - c$ and $\beta = b - c$.

Thus $(a, b, c) = (a - c)(1,0,0)$

$$+(b - c)(0,1,0) + c(1,1,1).$$

$\therefore$ S is a basis for $V_3(\mathbf{R})$

4. $S = \{1\}$ is a basis for the vector space $\mathbf{R}$ over R.

5. $S = \left\{\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix},\right.$

$\left.\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$ is a basis for $M_2(\mathbf{R})$, since any matrix $\begin{pmatrix}a & b\\c & d\end{pmatrix}$ can be uniquely written as

$$\begin{pmatrix}a & b\\c & d\end{pmatrix} = a\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix} + b\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}$$
$$+c\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix} + d\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$$

6. $\{1, i\}$ is a basis for the vector space $\mathbf{C}$ over $\mathbf{R}$.

7.Let V be the set of all polynomials of degree $\leq n$ in $\mathbf{R}[x]$. Then

$\{1, x, x^2, \dots\dots, x^n\}$ is a basis for V.

8. $\{(1,0), (i, 0), (0,1), (0, i)\}$ is a basis, for the vector space $\mathbf{C} \times \mathbf{C}$ over $\mathbf{R}$, for

$$(a + ib, c + id) =$$
$$a(1,0) + b(i, 0) + c(0,1) + d(0, i)$$

9. $S = \{(1,0,0), (0,1,0), (1,1,1), (1,1,0)\}$ spans the vector space V_3(R) but is not a basis.

Proof. Let $S' = \{(1,0,0), (0,1,0), (1,1,1)\}$

Then $L(S') = V_3(\mathbf{R})$ (refer example 3).

Now, since $S \subseteq S'$, we get $L(S) = V_3(\mathbf{R})$.

Thus S spans $V_3(\mathbf{R})$.

But S is linearly dependent since

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

Hence S is not a basis.

10. $S = \{(1,0,0), (1,1,0)\}$ is linearly independent but not a basis of $V_3(\mathbf{R})$.

Proof. Let $\alpha(1,0,0) + \beta(1,1,0) = (0,0,0)$. Then $\alpha + \beta = 0$ and $\beta = 0$.

$\therefore \alpha = \beta = 0$. Hence S is linearly independent.

Also $L(S) = \{(a, b, 0)/a, b \in \mathbf{R}\} \neq V_3(\mathbf{R})$.

$\therefore$ S is not a basis.

Exercises

1. Show that the following three vectors from a basis for $v_3(\mathrm{R})$.

   (a) $(1,2,-3), (2,5,1), (-1,1,4)$

   (b) $(1,1,0), (0,1,1), (1,0,1)$

   (c) $(2,-3,1), (0,1,2), (1,1,2)$.

2. Show that the following sets of vectors do not ^{f} orm a basis for $V_3(\mathbf{R})$.

   (a) $\{(1,0,0), (1,1,0)\}$

   (b) $\{(1,2,1), (1,3,5), (-1,0,1), (1,-1,2)\}$

   (c) $I(0,0,0), (1,0,0), (0,1,0), (0,0,1)$

   (d) $\{(3,2,1), (3,1,5), (3,4,-7)\}$

   (e) $\{(1,2,3), (2,3,4), (3,4,5)\}$

3. Show that $(1, i, 0), (2i, 1,1), (0,1 + i, 1 - i)$ form a basis for $V_3(\mathrm{C})$.

4. Find a basis for the vector space consisting of all matrices of the form

   (a) $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$

   (b) $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$

5. If $\{v_1, v_2, v_3\}$ is a basis for $V_3(\mathbf{R})$, show that $\{v_1 + v_2, v_2 + v_3, v_3 + v_1\}$ is also a basis. Is this true in (a) $V_3(\mathbf{Z}_2)$ (b) $V_3(\mathbf{Z}_3)$?

Answers.

4.(a) $\left\{\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix},\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right\}$

(b) $\left\{\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix},\begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right\}$

Theorem 10:

Let V be a vector space over a field F. Let $S = \{v_1, v_2, \dots, v_n\}$ span V. Let $S = \{w_1, w_2, \dots, w_m\}$ be a linearly independent set of vectors in V. Then $m \le n$.

Proof:

Since $L(S) = V$, every vector in V and in particular w_1, is a linear combination of $v_1, v_2, \dots, v_n$.

Hence $S_1 = \{w_1, v_1, v_2, \dots, v_n\}$ is a linearly dependent set of vectors. Hence there exists a vector $v_k \ne w_1$ in S_1 which is a linear combination of the preceding vectors.

Let $S_2 = \{w_1, v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$.

Clearly, $L(S_2) = V$.

Hence w_2 is a linear combination of the vectors in S_2.

Hence $S_3 = \{w_2, w_1, v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$ is linearly dependent. Hence there exists a vector in S_3 which is a linear combination of the preceding vectors. Since the w_i 's are linearly independent, this vector cannot be w_2 or w_1 and hence must be some v_j where $j \ne k$ (say, with $j > k$). Deletion of v_j from the set S_3 gives the set

$S_4 = \big(w_2, w_1, v_1, v_2, \dots, v_{l-1}, v_{l+1}, \dots v_{j-1}$, vyth..., in of n vectors spanning V.

In this process, at each step we insert one vector from $\{w_1, w_2, \dots, w_m\}$ and delete one vector from $[v_1, v_2, \dots, v_n)$.

If $m > n$ after repeating this process n times, we anive at the set $(w_n, w_{n-1} \dots, w_1)$ which spans V.

Hence w_{n+1} is a linear combination of $w_1, w_2 \dots . w_n$. Hence $\{w_1, w_2, \dots, w_n, w_{n+1}, \dots, w_m\}$ is linearly dependent which is a contradiction.

Hence $m \leq n$.

Theorem 11:

Any two bases of a finite dimensional vector space V have the same number of elements.

Proof:

Since V is finite dimensional, it has a basis say $S = \{v_1, v_2, \dots, v_n\}$.

Let $S' = \{w_1, w_2, \dots, w_m\}$ be any other basis for V.

Now, $L(S) = V$ and S' is a set of m linearly independent vectors. Hence by Theorem 10, $m \leq n$.

Also, since $L(S') = V$ and S is a set of n linearly independent vectors, $n \leq m$.

Hence $m = n$.

Definition:

Let V be a finite dimensional vector space over a field F. The number of elements in any basis of V is called the dimension of V and is denoted by $\dim V$.

Examples

1. $\dim V_n(\mathbf{R}) = n$, since $\{e_1, e_2, \dots, e_n\}$ is a basis of $V_n(\mathbf{R})$.
2. $M_2(\mathbf{R})$ is a vector space of dimension 4 over $\mathbf{R}$ since $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ is a basis for $M_2(\mathbf{R})$.
3. $\mathbf{C}$ is a vector space of dimension 2 over $\mathbf{R}$ since $\{1, i\}$ is a basis for $\mathbf{C}$.

4. Let V be the set of all polynomials of degree $\leq n$ in $\mathbf{R}|x|, V$ is a vector space over $\mathbf{R}$ having dimension $n+1$, since $(1, x, x^2, \dots, x^n)$ is a basis for V.

Theorem 12:

Let V be a vector space of dimension n. Then

(i) any set of m vectors where $m > n$ is linearly dependent.

(ii) any set of m vectors where $m < n$ cannotspan V.

Proof:

(i) Let $S = \{v_1, v_2, \dots, v_n\}$ be a basis for V. Hence $L(S) = V$.

Let S' be any set consisting of m vectors where $m > n$. Suppose S' is linearly independent. Since S spans V by Theorem 10, $m \leq n$ which is a contradiction. Hence S' is linearly dependent.

(ii) Let S' be a set consisting of m vectors where $m < n$.

Suppose $L(S') = V$.

Now, $S = \{v_1, v_2, \dots, v_n\}$ is a basis for V and hence linearly independent. Hence by Theorem 10, $n \leq m$ which is a contradiction. Hence S' cannot span V.

Theorem 13:

Let V be a finite dimensional vector space over a field F. Any linearly independent set of vectors in V is part of a basis.

Proof:

Let $S = \{v_1, v_2, \dots, v_r\}$ be a linearly independent set of vectors.

If $L(S) = V$ then S itself is a basis.

If $L(S) \neq V$, choose an element $v_{r+1} \in V - L(S)$.

Now, consider $S_1 = \{v_1, v_2, \dots, v_r, v_{r+1}\}$.

We shall prove that S_1 is linearly independent by showing that no vector in S_1 is a linear combination of the preceeding vectors. (refer theorem 6).

Since $\{v_1, v_2, \dots, \dots, v_r\}$ is linearly independent, v_i where $1 \leq i \leq r$ is not a linear combination of the preceeding vectors.

Also $v_{r+1} \notin L(S)$ and hence v_{r+1} is not a linear combination of $v_1, v_2, \dots \dots, v_r$.

Hence S_1 is linearly independent.

If $L(S_1) = V$, then S_1 is a basis for V. If not, we take an element $v_{r+2} \in V - L(S_1)$ and proceed as before. Since the dimension of V is finite, this process must stop at a certain stage giving the required basis containing S.

Theorem 14:

Let V be a finite dimensional vector space over a field F. Let A be a subspace of V. Then there exists a subspace B of V such that $V = A \oplus B$.

Proof:

Let $S = \{v_1, v_2, \dots \dots, v_r\}$ be a basis of A.

By theorem 13, we can find $w_1, w_2, w_i \in V$ such that such that $S' = (v_1, v_2, \dots \dots, v_r, w_1, u_2, \dots \dots, w_s)$ is a basis of V.

Now, let $B = L(\{w_1, w_2, \dots \dots \dots, w_t\})$

We claim that $A \cap B = \{0\}$ and $V = A + B$.

Now, let $v \in A \cap B$. Then $v \in A$ and $v \in B$.

$$\text{Hence } v = \alpha_1 v_1 + \cdots \dots \alpha_r v_r$$
$$= \beta_1 w_1 + \cdots \dots + \beta_s w_s$$
$$\therefore \ \alpha_1 v_1 + \cdots + \alpha_r v_r - \beta_1 w_1 - \cdots \dots - \beta_s w_s = 0$$

Now, since S' is linearly independent, $\alpha_i = 0 = \beta_j$ for all i and j.

Hence $v = 0$. Thus $A \cap B = \{0\}$. Now, let $v \in V$.

Then $v = (\alpha_1 v_1 + \cdots . + \alpha_r v_r) + (\beta_1 w_1 + \cdots \dots + \beta_s w_s) \in A + B$.

Hence $A + B = V$ so that $V = A \oplus B$.

Exercises

1. Let V be a finite-dimensional vector space. Let A and B be subspaces of V such that $V = A \oplus B$. Then show that $\dim V = \dim A + \dim B$.
2. Construct 3 subspaces W_1, W_2, W_3 of a vector space V such that $V = W_1 \oplus W_2 = W_1 \oplus W_3$ but $W_2 \neq W_3$.
3. For each of the following subspaces A of $V_3(\mathbf{R})$ find another subspace B such that $A \oplus B = V_3(\mathbf{R})$

   [i] $A = L((1,1,0),(0,1,1)\}$.

   [ii] $A = L((1,1,1))$.

   (iii) $A = L(\{e_1, e_2, e_3\})$.

Definition:

Let V be a vector apace and

$S = \{v_1, v_2, \dots \dots, v_n\}$ be a set of independent vectors in V. Then S is called a maximal linearly independent set if for every $v \in V - S$, the sa

$(v, v_1, v_2, \dots \dots, v_n)$ is linearly dependent.

Definition:

Let $S = \{v_1, v_2, \dots \dots, v_n\}$ be sen of vectors in V and let $L(S) = V$. Then S is called r minimal generating set if for any $v_i \in S$,

$L(S - \{v_i\}) \neq V$.

Theorem 15:

Let V be a vector space over a field P, Let $S = \{v_1, v_2, \dots \dots, v_n\} \subseteq V$. Then the following are equivalent.

(i) S is a basis for V.

(ii) S is a maximal linearly independent set.

(iii) S is a minimal generating set.

Proof:

(i) ⇒ (ii) Let $S = \{v_1, v_2, \dots\dots, v_n\}$ be a basis for V. Then by theorem 5.20 any $n + 1$ vectors in V are linearly dependent and hence S is a maximal linearly independent set.

(ii) ⇒ (i) Let $S = \{v_1, v_2, \dots, \dots, v_n\}$ be a maximal linearly independent set. Now to prove that S is a basis for V we shall show that $L(S) = V$.

Obviously $L(S) \subseteq V$.

Now, let $v \in V$.

If $v \in S$, then $v \in L(S)$. (since $S \subseteq L(S)$)

If $v \notin S, S' = \{v_1, v_2, \dots\dots, v_n, v\}$ is a linearly dependent set (since S is a maximal linearly independent set)

∴ There exists a vector in S' which is a lined combination of the preceeding vectors.

Since $v_1, v_2, \dots, v_n$ are linearly independent, this vector must be v. Thus, v is a linear combination of $v_1, v_2, \dots, \dots, v_B$. Therefore $v \in L(S)$.

Hence $V \subseteq L(S)$. Thus $V = L(S)$.

(i) ⇒ (iii) Let $S = \{v_1, v_2, \dots\dots\dots, v_n\}$ be a basis. Then $L(S) = V$.

If S is not minimal, there exists $v_f \in S$ such that $L(S - \{v_i\}) = V$.

Since S is linearly independent, $S - \{v_i\}$ is also linearly independent. Thus $S - \{v_i\}$ is a basis consisting of $n - 1$ elements which is a contradiction.

Hence S is a minimal generating set.

(iii) ⇒ (i) Let $S = \{v_1, v_2, \dots, v_n\}$ be a minimal generating set. To prove that S is a basis, we have to show that S is linearly independent.

If S is linearly dependent, there exists a vector v_k which is a linear combination of the preceeding vectors.

Clearly $L(S-\{v_k\}) = V$ contradicting the minimality of S.

Thus S is linearly independent and since

$L(S) = V, S$ is a basis for V.

Theorem 16:

Any vector space of dimension n over a field F is isomorphic to $V_n(F)$.

Proof:

Let V be a vector space of dimension n. Let $\{v_1', v_2, \dots, \dots, v_n\}$ be a basis for V.

Then we know that if $v \in V, v$ can be written uniquely as $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots . + \alpha_n v_n$, where $\alpha_i \in F$.

Now, consider the map $f: V \to V_n(F)$ given by

$f(\alpha_1 v_1 + \cdots + \alpha_n v_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)$.

Clearly f is $1-1$ and onto.

Let $v, w \in V$.

Then $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ and

$w = \beta_1 v_1 + \cdots \dots + \beta_n v_n$.

$$\begin{aligned} f(v+w) &= f[(\alpha_1+\beta_1)v_1 + \cdots + (\alpha_n+\beta_n)v_n] \\ &= \big((\alpha_1+\beta_1), (\alpha_2+\beta_2), \dots, (\alpha_n+\beta_n)\big) \end{aligned}$$

Also $f(\alpha v) = f(\alpha\alpha_1 v_1 + \cdots + \alpha\alpha_n v_n)$

$= (\alpha\alpha_1, \alpha\alpha_2, \dots \dots, \alpha\alpha_n)$

$= \alpha(\alpha_1, \alpha_2, \dots \dots, \alpha_n)$

$= \alpha f(v)$.

Hence f is an isomorphism of V to $V_n(F)$.

Corollary. Any two vector spaces of the same dimension over a field F are isomorphic, for, if the vector spaces are of dimension n, each is isomorphic to

$V_n(F)$ and hence they are isomorphic.

Theorem 17:

Let V and W be vector spaces over a field F. Let $T: V \to W$ be an isomorphism. Then T maps a basis of V onto a basis of W.

Proof:

Let $\{v_1, v_2, \dots, v_n\}$ be a basis for V.

We shall prove that $T(v_1), T(v_2), \dots \dots, T(v_n)$ are linearly independent and that they span W.

Now, $\alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots . + \alpha_n T(v_n) = 0.$

$\Rightarrow T(\alpha_1 v_1) + T(\alpha_2 v_2) + \cdots .. + T(\alpha_n v_n) = 0.$

$\Rightarrow T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots .. + \alpha_n v_n) = 0.$

$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$ (since T is $1 - 1$)

$\Rightarrow \alpha_1 = \alpha_2 = \cdots \dots = \alpha_n = 0$

(since $v_1, v_2, \dots, v_n$ are linearly independent).

$\therefore T(v_1), T(v_2), \dots ., T(v_n)$ are linearly independent.

Now, let $w \in W$. Then since T is onto, there exists a vector $v \in v$ such that $T(v) = w$

Let $v = \alpha_1 v_1 + \cdots \dots + \alpha_n v_n$.

$$\begin{aligned} &w = T(v) \\ \text{Then} \quad &= T(\alpha_1 v_1 + \cdots \dots + \alpha_n v_n) \\ &= \alpha_1 T(v_1) + \cdots \dots + \alpha_n T(v_n) \end{aligned}$$

Thus w is a linear combination of the vectors

$T(v_1), \dots \dots \dots, T(v_n)$.

$\therefore$ $T(v_1), \dots \dots, T(v_n)$ span W and hence is a has for W.

Corollary. Two finite dimensional vector spaces and W over a field F are isomorphic iff they have un same dimension.

Theorem 18:

Let V and W be finite dimensional vector spaces over a field F. Let $\{v_1, v_2, \dots \dots, v_n\}$ be a basis for V and let $w_1, w_2, \dots \dots, w_n$ be any n vectors in W (not necessarily distinct) Then there exists a unique linear transformation $T: V \to W$ such that $T(v_i) = w_i, i = 1,2, \dots \dots n$.

Proof:

Let $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots .. + \alpha_n v_n \in V$.

We define $T(v) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n$.

Now, let $x, y \in V$.

Let $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$ and

$y = \beta_1 v_1 + \cdots + \beta_n v_n$.

$$\therefore\ x + y = (\alpha_1 + \beta_1) v_1 + \cdots + (\alpha_n + \beta_n) v_n$$
$$\begin{aligned} \therefore T(x+y) &= (\alpha_1 + \beta_1) w_1 + \cdots + (\alpha_n + \beta_n) w_n \\ &= (\alpha_1 w_1 + \cdots + \alpha_n w_n) + \\ &= T(x) + T(y). \end{aligned}$$

Similarly, $T(\alpha x) = \alpha T(x)$. Hence T is a linear transformation.

Also $v_1 = 1v_1 + 0v_2 + \cdots + 0v_n$. Hence $T(v_1) = 1w_1 + 0w_2 + 0w_n = w_1$.

Similarly $T(v_i) = w_i$ for all $i = 1,2, \dots \dots, n$

Now, to prove the uniqueness, let $T': V \to W$ be any other linear transformation such that $T'(v_i) = w_i$.

Let $v = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V$

$$\begin{aligned} T'(v) &= \alpha_1 T'(v_1) + \cdots + \alpha_n T'(v_n) \\ &= \alpha_1 w_1 + \cdots + \alpha_n w_n = T(v). \end{aligned}$$

Hence $T = T'$.

Remark:

The above theorem shows that a linear transformation is completely determined by its values on the elements of a basis.

Theorem 19:

Let V be a finite dimensional vector space over a field F. Let W be a subspace of V. Then

(i) $\dim W \leq \dim V$.

(ii) $\dim \frac{V}{W} = \dim V - \dim W$.

Proof:

(i) Let $S = \{w_1, w_2, \ldots \ldots, w_m\}$ be a basis for W. Since W is a subspace of V, S is a part of a basis for V.

Hence $\dim W \leq \dim V$.

(ii) Let $\dim V = n$ and $\dim W = m$.

Let $S = \{w_1, w_2, \ldots, w_m\}$ be a basis for w. Clearly S is a linearly independent set of vectors in V.

Hence S is a part of a basis, in V. La $\{w_1, w_2, \ldots w_m, v_1, v_2, \ldots, v_r\}$ be a basij for V.

Then $m + r = n$.

Now, we claim $S' = \{W + v_1, W + v_2, \ldots, W + v_r\}$ is a basis for $\frac{V}{W}$.

$$\alpha_1(W + v_1) + \alpha_2(W + v_2) + \cdots + \alpha_r(W + v_r) = W + 0$$
$$\Rightarrow (W + \alpha_1 v_1) + (W + \alpha_2 v_2) + \cdots + (W + \alpha_r v_r) = W$$
$$\Rightarrow W + \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r = W$$
$$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r \in W$$

Now, since $\{w_1, w_2, \ldots, w_n\}$ is a basis for W

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = \beta_1 w_1 + \cdots + \beta_m w_m$$
$$\therefore \alpha_1 v_1 + \cdots + \alpha_r v_r - \beta_1 w_1 - \cdots - \beta_m w_m = 0$$
$$\therefore \alpha_1 = \alpha_2 = \alpha_r = \beta_1 = \beta_2 = \cdots = \beta_m = 0$$

$\therefore$ S' is a linearly independent set.

Now, let $W + v \in \frac{V}{W}$.

Let $v = \alpha_1 v_1 \dot{+} \ldots + \alpha_r v_r + \beta_1 w_1 + \cdots + \beta_m w_m$. Then $W + \dot{v} = W +$

$(\alpha_1 v_1 + \cdots + \alpha_r v_r$

$+\beta_1 w_1 + \cdots + \beta_m w_m)$

$= W + (\alpha_1 v_1 + \cdots + \alpha_r v_r)$

(since $\beta_1 w_1 + \cdots + \beta_m w_m \in W$)

$= (W + \alpha_1 v_1) + \cdots + (W + \alpha_r v_r)$

$= \alpha_1 (W + v_1) + \cdots + \alpha_r (W + v_r)$.

Hence S' spans $\frac{V}{W}$ so that S' is a basis for $\frac{V}{W}$.

$$\therefore \ \dim \frac{V}{W} = r = n - m$$

$$= \dim V - \dim W.$$

Theorem 20:

Let V be a finite-dimensional vector space over a field F. Let A and B be subspaces of V. Then $\dim(A + B) = \dim A + \dim B - \dim(A \cap B)$

Proof:

A and B are subspaces of V. Hence $A \cap B$ is subspace of V.

Let $\dim(A \cap B) = r$.

Let $S \doteq \{v_1, v_2, \ldots, v_r\}$ be a basis for $A \cap B$

Since $A \cap B$ is a subspace of A and B, S is a part of a basis for A and B.

Let $\{v_1, v_2, \ldots, v_r, u_1, u_2, \ldots u_s\}$ be a basis for A and $\{v_1, v_2, \ldots, v_r, w_1, w_2, \ldots, w_t\}$ be a basis for B.

We shall prove that $S' = \{v_1, \ldots v_r, u_1, \ldots \ldots u_s, w_1, \ldots, w_t]$ is a basis for $A + B$.

Let$\alpha_1 v_1 + \cdots + \alpha_r v_r + \beta_1 u_1 + \cdots + \beta_s u_s + \gamma_1 w_1 + \cdots + \gamma_t v_t = 0$

Then $\beta_1 u_1 + \cdots + \beta_s u_s = -(\gamma_1 w_1 + \cdots + \gamma_t w_t) - (\alpha_1 v_1 + \cdots + \alpha_r v_r) \in B$

Hence $\beta_1 u_1 + \cdots + \beta_s u_s \in B$.

Also $\beta_1 u_1 + \cdots \ldots + \beta_s u_s \in A$.

Hence $\beta_1 u_1 + \cdots + \beta_s u_s \in A \cap B$.

$\therefore\ \beta_1 u_1 + \cdots + \beta_s u_s = \delta_1 v_1 + \cdots + \delta_r v_r$.

$\therefore\ \beta_1 u_1 + \cdots + \beta_s u_s - \delta_1 v_1 - \cdots - \delta_r v_r = \mathbf{0}$.

$\therefore\ \beta_1 = \cdots = \beta_s = \delta_1 = \cdots = \delta_r = 0$

(since $\{u_1, \ldots, u_s, v_1, \ldots, v_r\}$ is linearly independent)

Similarly we can prove $\gamma_1 = \gamma_2 = \cdots = \gamma_t = 0$.

$\therefore\ \alpha_i = \beta_j = \gamma_k = 0$ for $1 \le i \le r$,

$1 \le j \le s; 1 \le k \le t$

Thus S' is a linearly independent set.

Clearly S' spans $A + B$.

$\therefore\ S'$ is a basis for $A + B$.

Hence $\dim(A + B) = r + s + t$.

Also $\dim A = r + s$; $\mathrm{aim} B = r + t$ and

$\dim(A \cap B) = r$.

$$\therefore\ \dim A + \dim B - \dim \Lambda \cap B = (r + s) + (r + t) - r$$
$$= r + s + t$$
$$= \dim(A + B).$$

Altier. By theorem 20, $\frac{A+B}{A} = \frac{B}{A \cap B}$.

Hence $\dim\left(\frac{A+B}{A}\right) = \dim\left(\frac{B}{A \cap B}\right)$.

$\therefore \dim(A + B) - \dim A = \dim B - \dim(A \cap B)$.

$\therefore \dim(A + B) = \dim A + \dim B - \dim(A \cap B)$.

Corollary, If $V = A \oplus B$, $\dim V = \dim A + \dim B$.

Proof. $V = A \oplus B \Rightarrow A + B = V$ and $A \cap B = \{0\}$

$\therefore \dim(A \cap B) = 0$.

Hence $\dim V^* = \dim A + \dim B$.

Exercises

1. Find the dimension of the subspace spanned by the following vectors in $V_3(\mathbf{R})$.

   (a) $(1,1,1), (-1,-1,-1)$.

   (b) $(1,0,2), (2,0,1), (1,0,1)$

   (c) $(1,2,-3), (0,0,1), (-1,2,1)$.

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

2. Find the dimension of the subspace spanned by the following vectors in $V_4(\mathbf{R})$

   (a) $(1, 0_1, e_2, e_3, e_4$

   (b) e_1, e_2

   (c) e_1, e_2, e_3

   (d) e_1

3. In $V_3(\mathbf{R})$, find $\dim(A + B)$ and $\dim(A \cap B)$ where

   (a) A is the subspace spanned by $(1,1,1)$ and B is the subspace spanned by $(-1,-1,-1)$

   (b) A is the subspace spanned by $(1,1,1)$ and B is the subspace spanned by (1,2,1).

   (c) A is the subspace spanned by $(1,1,1)$ and $(1,2,1)$ and B is the subspace spanned by $(0,0,1)$.

(d) A is the subspace spanned by $(1,1,1)$ and $(1,2,1)$ and B is the subspace spanned by $(1,-1,1)$ and $(-1,1,-1)$

4. Let V_1 and V_2 be subspaces of V such that $V_1 \cap V_2$ is the zero space. Prove that $\dim V_1 + \dim V_2 \leq \dim V$.
5. Let V_1 and V_2 be subspaces of V such that every vector $v \in V$ can be represented as $v = v_1 + v_2$ where $v_1 \in V_1$ and $v_2 \in V_2$. Prove that $\dim V_1 + \dim V_2 \geq \dim V$.
6. If A and B are finite dimensional subspaces of V such that $A \subseteq B$ and $\dim A = \dim B$ then show that $A = B$.
7. Let S be a subspace of a finite-dimensional vector space V. If $\dim V = \dim S$ then prove that $S = V$.
8. Let W_1 and W_2 be two subspaces of a finitedimensional vector space V.
   If $\dim V = \dim W_1 + \dim W_2$ and
   $W_1 \cap W_2 = \{\mathbf{0}\}$ prove that, $V = W_1 \oplus W_2$.

Answers.

1. (a) 1 (b) 2 (c) 3 (d) 2
2. (a) 4 (b) 2 (c) 3 (d) 1
3. (a) 1; 1 (b) 2; 0 (c) 3; 0 (d) 3; 0

UNIT III

Rank and Nullity of a transformation– Matrix of a linear transformation – Inner product space: Definition and examples– Orthogonality – Orthogonal complement.

Chapter 3: Sections- 3.1 to 3.5

3.1. Rank and Nullity:

Definition:

Let $T: V \to W$ be a linear transformation. Then the dimension of $T(V)$ is called the rank of T. The dimension of ker T is called the nullity of T.

Theorem 21:

Let $T: V \to W$ be a linear transformation. Then $\dim V = \text{rank} T + \text{nullity } T$.

Proof:

We know that $V/\ker T = T(V)$.

$\therefore\ \dim V - \dim(\ker T) = \dim(T(V))$

$\therefore\ \dim V - \text{nullity } T = \text{rank} T$

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

Note. $\ker T$ is also called null space of T.

Example:

Let V denote the set of all polynomials of degree $\leq n$ in R$[x]$. Let $T: V \to V$ be defined by $T(f) = \frac{df}{dx}$. We know that T is a linear transformation. Since $\frac{df}{dx} = 0 \Leftrightarrow$ f is constant, $\ker T$ consists of all constant polynomials. The dimension of this subspace of V is 1. Hence nullity T_* is 1.

Since $\dim V = n + 1$, rank $T = n$.

Exercises

1. Find the rank and nullity of the linear transformations given in section 1.3.

2. Let V be a finite-dimensional vector space over a field F. Let $T: V \to V$ be a linear transformation such that rank $T =$ nullity T. Show that dim$V$ is even. Give an example of such a transformation.

Answers.

1. nullity $T = \dim V$; rank$T = 0$.
2. nullity $T = 0$; rank$T = \dim V$:
3. nullity $T = \dim W$;

   rank$T = \dim V - \dim W$.
4. nullity $T = 2$; rank$T = 1$;
5. nullity $T = 1$; rank$T = n$.
6. nullity $T = 0$; rank$T = n + 1$.

Definition:

A linear transformation $T: V \to W$ is called non-singular if T is 1-1; otherwise T is called singular.

Exercises

1. Let V and W be finite dimensional vector spaces over a field F and $\dim V > \dim W$. Then show that any linear transformation $T: V \to W$ is singular.
2. Let V be a finite-dimensional vector space over a field F. Then any non-singular linear transformation $T: V \to V$ is onto.
3. Let $T: V \to W$ be a linear transformation. Show that T is a non-singular iff rank$T = \dim V$.
4. Let $T_1: V \to V$ and $T_2: V \to V$ be linear transformations. Prove that

   (a) $\text{rank}(T_2T_1) \leq \text{rank}T_2$.

(b) nullity $(T_2T_1) \geq$ nullity T_1.

(c) rank$(T_2T_1) =$ rankT_2 iff T_1 is nonsingular.

5. Let $T: V \to W$ be a linear transformation which is both 1-1 and onto. Show that $T^{-1} : W \to V$ is a linear transformation.

Determine which of the following statements are true and which are false.

(a) If $T: V \to W$ is a linear transformation then

(i) rank$T \leq$ dim$V$

(ii) nullity $T \leq$ dim$V$.

(iii) rank$T \leq$ dim$W$.

(iv) If T is onto rank$T =$ dim$W$.

(v) If T is non-singular rank$T =$ dim$V$

(vi) rank$T =$ dim$V \Rightarrow$ nullity $T = 0$.

(b) Every linear transformation $T: V_4(\mathbf{R}) \to V_3(\mathbf{R})$ is singular.

(c) If $T: V \to W$ is non-singular and $\{v_1, \ldots, v_n\}$ is a basis then $\{T(v_1), \ldots \ldots, T(v_n)\}$ is a basis for W.

Answers: 6. (a) (i) T (ii) T (iii) T (iv) T (v) T (vi) T (b) T (c) F.

3.2. Matrix of a Linear Transformation:

Let V and W be finite dimensional vector spaces over a field F. Let dim$V = m$ and dim$W = n$. Fix an ordered basis $\{v_1, v_2, \ldots, v_m\}$ for V and an ordered basis $\{w_1, w_2, \ldots, w_n\}$ for W.

Let $T: V \to W$ be a linear transformation. We have seen that T is completely specified by the elements $T(v_1), T(v_2), \ldots, T(v_m)$. Now, let

$T(v_1) = a_{11}w_1 + a_{12}w_2 + \cdots \dots + a_{1n}w_n$
$T(v_2) = a_{21}w_1 + a_{22}w_2 + \cdots \dots + a_{2n}w_n$ … … … …(1)
… … … … … … … … . … . … . .
$T(v_m) = a_{m1}w_1 + a_{m2}w_2 + \cdots \dots .$

Hence $T(v_1), T(v_2), \dots \dots, T(v_m)$ are completely specified by the mn elements a_{ij} of the field F. These a_{ij} can be conveniently arranged in the form of m rows and n columns as follows.

$$\begin{pmatrix} a_{11} & a_{12} & \dots \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots \dots & \dots \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Such an array of mn elements of F arranged in m rows and n columns is known as $m \times n$ matrix over the field F and is denoted by (a_{ij}). Thus, to every linear transformation T there is associated with it an $m \times n$ matrix over F. Conversely any $m \times n$ matrix over F defines a linear transformation $T: V \to W$ given by the formula (1).

Note:

The $m \times n$ matrix which we have associated with a linear transformation $T: V \to W$ depends on the choice of the basis for V and W.

For example, consider the linear transformation $T: V_2(\mathbf{R}) \to V_2(\mathbf{R})$ given by $T(a, b) = (a, a + b)$. Choose $\{e_1, e_2\}$ as a basis both for the domain and the range.

Then $T(e_1) = (1,1) = e_1 + e_2$

$T(e_2) = (0,1) = e_2$.

Hence the matrix representing T is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Now, we choose $\{e_1, e_2\}$ as a basis for the domain and $\{(1,1), (1, -1)\}$ as a basis for the range.

Let $w_1 = (1,1)$ and $w_2 = (1, -1)$.

Then $T(e_1) = (1,1) = w_1$,

and $T(e_2) = (0,1) = (1/2)w_1 - (1/2)w_2$.

Hence the matrix representing T is $\begin{pmatrix} 1 & 0 \\ 1/2 & -1/2 \end{pmatrix}$

Solved problems

Problem 1:

Obtain the matrix representing the linear transformation $T: V_3(\mathbf{R}) \to V_3(\mathbf{R})$ given by $T(a, b, c) = (3a, a - b, 2a + b + c)$ wrat. the standard basis (c_1, c_2, c_3).

Solution:

$$T(e_1) = T(1,0,0) = (3,1,2) = 3e_1 + e_2 + 2e_3$$
$$T(e_2) = T(0,1,0) = (0,-1,1) = -e_2 + e_3$$
$$T(e_3) = T(0,0,1) = (0,0,1) = e_3$$

Thus, the matrix representing T is $\begin{pmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

Problem 2:

Find the linear transformation

$T: V_3(\mathbf{R}) \to V_3(\mathbf{R})$ determined by the matrix

$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{pmatrix}$ w.r.t. the standard basis $\{e_1, e_2, e_3\}$.

Solution:

$$T(e_1) = e_1 + 2e_2 + e_3 = (1,2,1)$$
$$T(e_2) = 0e_1 + e_2 + e_3 = (0,1,1)$$
$$T(e_3) = -e_1 + 3e_2 + 4e_3 = (-1,3,4)$$

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

$$= ae_1 + be_2 + ce_3.$$

$$\therefore\ T(a,b,c) = T(ae_1 + be_2 + ce_3)$$
$$= aT(e_1) + bT(e_2) + cT(e_3)$$
$$= a(1,2,1) + b(0,1,1) + c(-1,3,4).$$

$$\therefore T(a,b,c) = (a - c, 2a + b + 3c, a + b + 4c)$$

This is the required linear transformation.

Exercises

1. Obtain the matrices for the following linear transformations.

(a) $T: V_2(\mathbf{R}) \to V_2(\mathbf{R})$ given by $T(a,b) = (-b,a)$ w.r.t.

(i) standard basis

(ii) the basis $\{(1,2),(1,-1)\}$ for both domain and range.

(b) $T: V_3(\mathbf{R}) \to V_2(\mathbf{R})$ given by $T(a,b,c) = (a+b, 2c-a)$ w.r.t.

(i) standard basis

(ii) $\{(1,0,-1),(1,1,1),(1,0,0)\}$ as a basis for V_3(R) and $\{(0,1),(1,0)\}$ for $V_2(R)$.

(c) $T: V_3$(R) $\to V_3$(R) given by

$T(a,b,c) = (3a + c, -2a + b, a + 2b + 4c)$ w.r.t.

(i) the standard basis

(ii) the basis $\{(1,0,1),(-1,2,1),(2,1,1)\}$ for both domain and range.

(d) Let V be the set of all polynomials of degree $\le n$ in $\mathbf{R}[x]$.

$T: V \to V$ defined by $T(f) = \frac{df}{dx}$ w.r.t. the basis $\{1, x, x^2, \dots, x^n\}$.

2. Obtain the linear transformation determined by the following matrices

(a) $T: V_2$(R) $\to V_2$(R) given by $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ w.r.t. the standard basis.

(b) $T: V_3(\mathbf{R}) \to V_3(\mathbf{R})$ given by $\begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix}$ w.r.t. the standard basis.

(c) $T: V_2(\mathbf{R}) \to V_3(\mathbf{R})$ given by $\begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$ w.r.t. the standard basis.

Answers:

(a) (i) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (ii) $\begin{pmatrix} -1/3 & -5/3 \\ 2/3 & 1/3 \end{pmatrix}$ (b) (i) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} -3 & 1 \\ 1 & 2 \\ -1 & 1 \end{pmatrix}$

(ii) $\begin{pmatrix} 17/4 & -3/4 & -1/2 \\ 35/4 & 15/4 & -7/2 \\ 17/2 & -3/2 & 0 \end{pmatrix}$ (c) (i) $\begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 0 & 0 & . & . & . & 0 & 0 \\ 1 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & 2 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & 3 & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & n & 0 \end{pmatrix}$

2. (a) $T(a, b) = (a\cos\theta + b\sin\theta, -a\sin\theta + .\, b\cos\theta)$

(b) $T(x, y, x) = (ax + by + cz, bx + cy + az, cx + ay + bz)$

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

Definition:

Let $A = (a_{ij})$ and $B = (b_{ij})$ be two $m \times n$ matrices. We define the sum of these two matrices by $A + b = (a_{ij} + b_{ij})$.

Note that we have defined addition only for two matrices having the same number of rows and the same number of columns.

Definition:

Let $A = (a_{ij})$ be an arbitrary matrix over I field F. Let $\alpha \in F$. We define $\alpha A = (\alpha a_{ij})$.

Theorem 1:

The set $M_{m\times n}(F)$ of all $m \times n$ matrices wer the field F is a vector space of dimension mn ver F under matrix addition and scalar multiplication efined above.

Proof:

Let $A = \left(a_{ij}\right)$ and $B = \left(b_{ij}\right)$ be two $m \times n$ matrices over the field F. The addition of $m \times n$ matrices is a binary operation which is both commutative and associative. The $m \times n$ matrix whose entries are 0 is the identity matrix and ($-a_{ij}$) is the inverse matrix of (a_{ij}). Thus, the set of all $m \times n$ matrices over the field F is an abelian group with respect to addition. The verification of the following axioms is straight forward.

(a) $\alpha(A + B) = \alpha A + \alpha B$

(b) $(\alpha + \beta)A = \alpha A + \beta A$

(c) $(\alpha\beta)A = \alpha(\beta A)$

(d) $1A = A$.

Hence the set of all $m \times n$ over F is a vector space over F.

Now, we shall prove that the dimension of this vector space is mn. Let E_{ij} be the matrix with entry 1 in the $(i,j)^{\text{th}}$ place and 0 in the other places. We have mn matrices of this form. Also, any matrix $A = \left(a_{ij}\right)$ can be written as $A = \sum \ a_{ij}E_{ij}$. Hence A is a linear combiation of the matrices E_{ij}. Further these mn matrices E_{ij} are linearly independent. Hence these mn matrices form a basis for the space of all $m \times n$ matrices. Therefore, the dimension of the vector space is mn.

Theorem 2:

Let V and W be two finite dimensional vector spaces over a field F. Let $\dim V = m$ and $\dim W = n$. Then $L(V, W)$ is a vector space of dimension mn over F.

Proof:

By theorem 8 sec 1.3, $L(V, W)$ is a vector space over F. Now, we shall prove that the vector space $L(V, W)$ is isomorphic to the vector space $M_{m \times n}(F)$. Since $M_{m \times n}(F)$ is of dimension mn, it follows that $L(V, W)$ is also of dimension mn.

Fix a basis $\{v_1, v_2, \dots, v_m\}$ for V and a basis $\{w_1, w_2, \dots \dots, w_n\}$ for W.

We know that any linear transformation

$T \in L(V, W)$ can be represented by an $m \times n$ matrix over F.

, et T be represented by $M(T)$. This function

$M: L(V, W) \to M_{m \times n}(F)$ is clearly 1-1 and onto.

Let $T_1, T_2 \in L(V, W)$ and $M(T_1) = (a_{ij})$ and $M(T_2) = (b_{ij})$.

$$M(T_1) = (a_{ij}) \Rightarrow T_1(v_i) = \sum_{j=1}^{n} a_{ij} w_j$$

$$M(T_2) = (b_{ij}) \Rightarrow T_2(v_i) = \sum_{j=1}^{n} b_{ij} w_j$$

$$\therefore\ (T_1 + T_2)(v_i) = \sum_{j=1}^{n} (a_{ij} + b_{ij}) w_j$$

$$\begin{aligned} \therefore\ M(T_1 + T_2) &= (a_{ij} + b_{ij}) \\ &= (a_{ij}) + (b_{ij}) \\ &= M(T_1) + M(T_2) \end{aligned}$$

Similarly, $M(\alpha T_1) = \alpha M(T_1)$.

Hence M is the required isomorphism from $L(V, W)$ to $M_{m \times n}(F)$.

Exercises:

Determine which of the following statements are true and which are false.

1. Any vector space is an abelian group with respect to vector addition.
2. Any vector space contains an infinite number of elements.
3. Given any field F, there exists a vector space of dimension n over F.

4. Any two vector spaces over a field F are isomorphic iff they have the same dimension,
5. Any two bases of a finite dimensional vector space have the same number of elements.
6. Any vector spaces of dimension $n > 1$ has non-trivial subspaces.
7. If V is a vector space of dimension n and $m < n$, then there exists a subspace of V of dimension m.
8. Any linear transformation from a finite dimensional vector space V to a finite dimensional vector space W can be represented by a matrix.

Answers.

1.T 2. F 3.T 4.T 5.T 6.T 7.T 8.T

3.3. Inner Product Spaces:

Introduction:

Up to this point we have dealt with the algebraic properties of a vector space and these properties are consequences of the basic operations, namely, vecto addition and scalar multiplication defined in the vector space. We know that in the usual three-dimensional vector space $V_3(\mathbf{R})$ it is possible to talk about the length of a vector and angle between two vectors. These concepts of length and angle can be defined in terms of the usual "dot product" or "scalar product" of two vectors. The dot product of $u = (a_1, b_1, c_1)$ and

$v = (a_2, b_2, c_2)$ is defined by

$u \cdot v = a_1a_2 + b_1b_2 + c_1c_2$

We note that the length of u is given by $\sqrt{u \cdot u}$ and the angle θ between u and v is determined by $\cos\theta = \frac{u \cdot v}{\sqrt{u \cdot u}\sqrt{v \cdot v}}$. Hence u and v are perpendicular or orthogonal iff $u.v = 0$.

An inner product on a vector space is a generalization of the dot product and in terms of such an inner product we can define the length of a vector and angle between two vectors. Our study about angle will be restricted to the concept of perpendicularity of two vectors.

Throughout this section we shall deal only with vector spaces over the field F of real or complex numbers.

Definition and Examples:

Definition:

Let V be a vector space over F. An inner product on V is a function which assings to each ordered pair of vectors u, v in V a scalar in F denoted by $\langle u, v\rangle$ satisfying the following conditions.

(i) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$

(ii) $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$

(iii) $\langle u, v\rangle = \overline{\langle v, u\rangle}$, where $\overline{\langle v, u\rangle}$ is the complex conjugate of $\langle u, v\rangle$.

(iv) $\langle u, u\rangle \geq 0$ and $\langle u, u\rangle = 0$ iff $u = 0$.

A vector space with an inner product defined on it is called an inner product space. An inner product space is called a Euclidean space or unitary space according as F is the field of real numbers or complex numbers.

Note 1: If F is the field of real numbers then condition (iii) takes the form $\langle u, v\rangle = \langle v, u\rangle$. Further (iii) asserts that $\langle u, u\rangle$ is always real and hence (iv) is meaningful whether F is the field of real or complex numbers.

Note 2:

$\langle u, \alpha v\rangle = \bar{\alpha}\langle u, v\rangle$.

For,
$$\begin{aligned}\langle u, \alpha v\rangle &= \overline{\langle \alpha v, u\rangle}\\ &= \overline{\alpha\langle v, u\rangle}\\ &= \bar{\alpha}\overline{\langle v, u\rangle}\\ &= \bar{\alpha}\langle u, v\rangle.\end{aligned}$$

Note 3:

$\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$

For,
$$\begin{aligned}\langle u, v + w\rangle &= \overline{\langle v + w, u\rangle}\\ &= \overline{\langle v, u\rangle + \langle w, u\rangle}\\ &= \overline{\langle v, u\rangle} + \overline{\langle w, u\rangle}\\ &= \langle u, v\rangle + \langle u, w\rangle.\end{aligned}$$

Note 4:

$\langle u, \mathbf{0}\rangle = \langle \mathbf{0}, v\rangle = 0$.

For, $\langle u, \mathbf{0}\rangle = \langle u, 00\rangle = 0\langle u, \mathbf{0}\rangle = 0$.

Similarly $\langle 0, v\rangle = 0$.

Exercises

Show that in an inner product space

(i) $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$

(ii) $\langle u, \alpha v + \beta w\rangle = \bar{\alpha}\langle u, v\rangle + \bar{\beta}\langle u, w\rangle$

(iii) $\langle \alpha u + \beta v, \gamma w + \delta z\rangle = \alpha\bar{\gamma}\langle u, w\rangle + \alpha\bar{\delta}\langle u, z\rangle + \beta\bar{\gamma}\langle v, w\rangle + \beta\bar{\delta}\langle v, z\rangle$ where $\alpha, \beta, \gamma, \delta \in F$ and $u, v, w, z \in V$

Examples

1. $V_n(\mathbf{R})$ is a real inner product space with inner product defined by

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n \text{ where}$$
$$x = (x_1, x_2, \dots, x_n) \text{ and}$$
$$y = (y_1, y_2, \dots, y_n)$$

This is called the standard inner product on $V_n(\mathbf{R})$.

Proof:

Let $x, y, z \in V_n(\mathbf{R})$ and $\alpha \in \mathbf{R}$.

(i) $\langle x + y, z\rangle = (x_1 + y_1)z_1 + (x_2 + y_2)z_2$

$$+ \cdots + (x_n + y_n)z_n$$
$$= (x_1z_1 + x_2z_2 + \cdots + x_nz_n)$$
$$+ (y_1z_1 + y_2z_2 + \cdots + y_nz_n)$$
$$= \langle x, z\rangle + \langle y, z\rangle$$

(ii) $\langle \alpha x, y\rangle = \alpha x_1y_1 + \alpha x_2y_2 + \cdots + \alpha x_ny_n$

$$= \alpha(x_1y_1 + x_2y_2 + \cdots + x_ny_n)$$
$$= \alpha\langle x, y\rangle$$

(iii) $\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots x_ny_n.$

$$= y_1x_1 + y_2x_2 + \cdots + y_nx_n$$
$$= \langle y, x\rangle$$

(iv) $\langle x, x\rangle = x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0$ and $\langle x, x\rangle = 0$ iff $x_1 = x_2 = \cdots = x_n = 0$

$\therefore\ \langle x, x\rangle = 0$ iff $x = \mathbf{0}$

2. $V_n(\mathrm{C})$ is a complex inner product space with inner product defined by

$\langle x, y\rangle = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$ where $x = (x_1, x_2, \dots, x_n)$ and $y =$

$(y_1, y_2, \dots, y_n)$.

Proof. Let $x, y, z \in V_n(\mathbf{C})$ and $\alpha \in \mathbf{C}$.

(i) $\langle x + y, z\rangle = (x_1 + y_1)\overline{z_1} + (x_2 + y_2)\overline{z_2}$

$$+ \dots + (x_n + y_n)\overline{z_{11}}$$
$$= (x_1\overline{z_1} + x_2\overline{z_2} + \cdots + x_n\overline{z_n})$$
$$+ (y_1\overline{z_1} + y_2\overline{z_2} + \cdots + y_n\overline{z_n})$$
$$= \langle x, z\rangle + \langle y, z\rangle$$

(ii) $\langle \alpha x, y\rangle = \alpha x_1\overline{y_1} + \alpha x_2\overline{y_2} + \cdots + \alpha x_n\bar{y}_n$

$$= \alpha(x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n})$$
$$= \alpha\langle x, y\rangle$$

(iii) $\overline{\langle y, x\rangle} = \overline{y_1\overline{x_1} + y_2\overline{x_2} + \cdots + y_n\overline{x_n}}$

$$= \overline{y_1}x_1 + \overline{y_2}x_2 + \cdots + \overline{y_n}x_n$$
$$= x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$$
$$= \langle x, y\rangle.$$

(iv) $\langle x, x\rangle = x_1\overline{x_1} + \cdots + x_n\overline{x_n}$

$$= |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \geq 0$$

and $\langle x, x\rangle = 0$ iff $x = \mathbf{0}$.

3. Let V be the set of all continuous real valued functions defined on the closed interval $[0,1]$. V is a real inner product space with inner product defined by

$$\langle f, g\rangle = \int_0^1 f(t)g(t)dt$$

Proof. Let $f, g, h \in V$ and $\alpha \in \mathbf{R}$.

(i) $\langle f + g, h\rangle = \int_0^1 [f(t) + g(t)h(t)dt$

$$= \int_0^1 f(t)h(t)dt + \int_0^1 g(t)h(t)dt$$
$$= \langle f, h\rangle + \langle g, h\rangle$$

(ii) $\langle \alpha f, g\rangle = \int_0^1 \alpha f(t)g(t)dt$

$$= \alpha\int_0^1 f(t)g(t)dt$$
$$= \alpha\langle f, g\rangle$$

(iii) $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$

$$= \int_0^1 g(t)f(t)dt$$
$$= \langle g, f \rangle$$

(iv) $\langle f, f \rangle = \int_0^1 [f(t)]^2 dt \geq 0$ and $\langle f, f \rangle = 0$ iff $f = 0$

1.Show that $V_2(\mathbf{R})$ is an inner product space with inner product defined by

$\langle x, y \rangle = x_1y_1 + x_2y_1 - x_1y_2 + 4x_2y_2$

where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

2. Show that $V_2(\mathbf{C})$ is an inner product space with inner product defined by

$\langle x, y \rangle = 2x_1\overline{y_1} + x_1\overline{y_2} + y_2\overline{y_1} + x_2\overline{y_2}$

where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

3. Let V be the set of all continuous complex valued functions defined on the closed interval $[0,1]$. Show that V is a complex inner product space with inner product detined by $\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}dt$.

4. Which of the following are inner products on $V_2(\mathbf{R})$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

(a) $\langle x, y \rangle = x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$.

(b) $\langle x, y \rangle = x_1^2 - 2x_1y_2 - 2x_2y_1 + y_1^2$.

(c) $\langle x, y \rangle = 6x_1y_1 + 7x_2y_2$.

(d) $\langle x, y \rangle = x_1y_1 - 2x_2y_1 - 2x_1y_2 + 4x_2y_2$.

Answers. 4. (b) and (d) are not inner products.

Definition:

Let V be an inner product space and let $x \in V$. The norm or length of x, denoted by $\|x\|$, is defined by $\|x\| = \sqrt{\langle x, x \rangle}$.

x is called a unit vector if $\|x\| = 1$.

Solved Problems

Problem 1:

Let V be the vector space of polynomials with inner product given by $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$. Let $f(t) = t + 2$ and $g(t) = t^2 - 2t - 3$.

Find (i) $\langle f, g\rangle$ (ii) $\|f\|$.

Solution:

(i) $\langle f, g\rangle = \int_0^1 f(t)g(t)dt$

$= \int_0^1 (t+2)(t^2-2t-3)dt$

$= \int_0^1 (t^3 - 7t - 6)dt$

$= \left[\frac{t^4}{4} - \frac{7t^2}{2} - 6t\right]_0^1$

$= \frac{1}{4} - \frac{7}{2} - 6$

$= -\frac{37}{4}$

(ii) $\|f\|^2 = \langle f, f\rangle$

$= \int_0^1 [f(t)]^2 dt$

$= \int_0^1 (t+2)^2 dt$

$= \int_0^1 (t^2 + 4t + 4)dt$

$= \left[\frac{t^3}{3} + 2t^2 + 4t\right]_0^1$

$= \frac{1}{3} + 2 + 4$

$$= \frac{19}{3}$$

$$\therefore \ \|f\| = \frac{\sqrt{19}}{\sqrt{3}}$$

Exercises

1. Find the norm of the following vectors in $V_3(\mathbf{R})$ with standard inner product.

   (a) $(1,1,1)$

   (b) $(1,2,3)$

   (c) $(3,-4,0)$

   (d) $4x + 5y$ where $x = (1,-1,0)$ and $y = (1,2,3)$.

2. Find the set of all unit vectors in $V_3(\mathbf{R})$ with standard norm.

Answers.

1. (a) $\sqrt{3}$ (b) $\sqrt{14}$ (c) 5 (d) $3\sqrt{38}$

2. All points on the unit sphere with center $(0,0,0)$.

Theorem 3:

The norm defined in an inner product. space V has the following properties.

(i) $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$.

(ii) $\|\alpha x\| = |\alpha|\|x\|$.

(iii) $|\langle x, y\rangle| \leq \|x\|\|y\|$ (Schwartz's inequality).

(iv) $\|x + y\| \leq \|x\| + \|y\|$ (Triangle inequality).

Proof:

(i) $\|x\| = \sqrt{\langle x, x\rangle} \geq 0$ and $\|x\| = 0$ iff $x = 0$.

(ii) $\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle$

$= \alpha\langle x, \alpha x\rangle$

$= \alpha\bar{\alpha}\langle x, x\rangle$

$= |\alpha|^2\|x\|^2$

Hence $\|\alpha x\| = |\alpha|\|x\|$.

(iii) The inequality is trivially true when $x = 0$ or $y = 0$. Hence let $x \neq 0$ and $y \neq 0$.

Consider $z = y - \frac{\langle y,x\rangle}{\|x\|^2}x$.

Then $0 \leq \langle z, z\rangle$

$$= \left\langle y - \frac{\langle y,x\rangle}{\|x\|^2}x, y - \frac{\langle y,x\rangle}{\|x\|^2}x\right\rangle$$

$$= \langle y, y\rangle - \frac{\overline{\langle y,x\rangle}}{\|x\|^2}\langle y, x\rangle - \frac{\langle y,x\rangle}{\|x\|^2}\langle x, y)$$

$$+ \frac{\langle y,x\rangle\overline{\langle y,x\rangle}}{\|x\|^2\|x\|^2}\langle x, \mathrm{x}\rangle$$

$$= \|y^2\| - \frac{\overline{\langle y,x\rangle}\langle y,x\rangle}{\|x\|^2} - \frac{\langle y,x\rangle\langle x,y\rangle}{\|x\|^2} + \frac{\langle y,x\rangle\overline{\langle y,x\rangle}}{\|x\|^2}$$

$$= \|y\|^2 - \frac{\overline{\langle x,y\rangle}\langle x,y\rangle}{\|x\|^2}$$

$\therefore 0 \leq \|x\|^2\|y\|^2 - |\,(x, y)\|^2$

$\therefore\ |\langle x, y\rangle| \leq \|x\|\|y\|$.

(iv) $\|x + y\|^2 = \langle x + y, x + y\rangle$

$= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle$
$= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2$
$= \|x\|^2 + 2\text{Re}\langle x, y \rangle + \|y\|^2$
$\leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2$
$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$ (by (iii))
$\leq (\|x\| + \|y\|)^2$
$\therefore \|x + y\| \leq \|x\| + \|y\|$.

Exercises

1.Applying Schwartz's inequality to $V_n(\mathrm{C})$ wid standard inner product, show that

$|\sum_{i=1}^{n} x_i\overline{y_i}| \leq [\sum_{i=1}^{n} |x_i|^2]^{1/2}(\sum_{i=1}^{n} |y_i|^2)^{1/2}$

2. Show that in any inner product space V

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2$$

3.Show that in any inner product space

$$\|\alpha x + \beta y\|^2 = |\alpha|^2\|x\|^2 + \alpha\bar{\beta}\langle x, y \rangle + \bar{\alpha}\beta\langle y, x \rangle + |\beta|^2\|y\|^2$$

4.Show that if equality is valid in Schwartz's inequality or triangle inequality then x and y are linearly dependent. Is the converse true?

5.(a) Show that in a real inner product space, if $\langle x, y \rangle = 0$, then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

(b) Show that in a real inner product space, if

$\|x\|^2 + \|y\|^2 = \|x + y\|^2$,

then $\langle x, y \rangle = 0$.

6. In an inner product space, we define the distance between any two vectors x and y by $d(x, y) = \|x - y\|$. Show that

(a) $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = y$.

(b) $d(x, y) = d(y, x)$.

(c) $d(x, y) \leq d(x, z) + d(z, y)$.

3.4. Orthogonality:

Definition:

Let V be an-inner product space and let $y \in V. x$ is said to be orthogonal to y if $\langle x, y \rangle = 0$.

Note 1. x is orthogonal to $y \Rightarrow \langle x, y \rangle = 0$

$$\Rightarrow \overline{\langle x, y \rangle} = \overline{0}$$
$$\Rightarrow \langle y, x \rangle = 0$$

$\Rightarrow y$ is orthogonal to x.

Thus x and y are orthogonal iff $\langle x, y \rangle = 0$.

Note 2. x is orthogonal to $y \Rightarrow \alpha x$ is orthogonal to y.

Note 3. x_1 and x_2 are orthogonal to $y \Rightarrow x_1 + x_2$ is thogonal to y.

Note 4. 0 is orthogonal to every vector in Y and is the only vector with this property.

Definition:

Let V be an inner product space. A set S of vectors in V is said to be an orthogonal set if any two distinct vectors in S are orthogonal.

Definition:

S is said to be an orthonormal set if S is orthogonal and $||x|| = 1$ for all $x \in S$.

Example:

The standard basis $\{e_1, e_2, \dots, e_n\}$ in R^n or C^{n} is an orthogonal set with respect to the standard inner product.

Theorem 4:

Let $S = \{v_1, v_2, \dots, v_n\}$ be an orthogonal set of non-zero vectors in an inner product space V. Then S is linearly independent.

Proof:

Let $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$

Then $\langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, v_1 \rangle = \langle \mathbf{0}, v_1 \rangle = 0$.

$\therefore\ \alpha_1 \langle v_1, v_1 \rangle + \alpha_2 \langle v_2, v_1 \rangle + \cdots + \alpha_n \langle v_n, v_1 \rangle = 0$.

$\therefore\ \alpha_1 \langle v_1, v_1 \rangle = 0$ (since S is orthogonal)

$\therefore\ \alpha_1 = 0$ (since $v_1 \neq 0$)

Similarly $\alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$.

Hence S is linearly independent.

Theorem 5:

Let $S = \{v_1, v_2, \dots, v_n\}$ be an orthogonal set of non-zero vectors in V. Let $v \in V$ and $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$. Then $\alpha_k = \frac{\langle v, v_k \rangle}{\|v_k\|^2}$.

Proof:

$\langle v, v_k \rangle = \langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, v_k \rangle$

$\quad = \alpha_1 \langle v_1, v_k \rangle + \alpha_2 \langle v_2, v_k \rangle + \cdots$
$\quad + \alpha_k \langle v_k, v_k \rangle + \cdots + \alpha_m \langle v_n, v_k \rangle$

$= \alpha_k \langle v_k, v_k \rangle$ (since S is orthogonal)

$= \alpha_k \|v_k\|^2$

$\therefore\ a_k = \frac{\langle v, v_k \rangle}{\|v_k\|^2}$.

Theorem 6:

Every finite dimensional inner product space has an orthonormal basis.

Proof:

Let V be a finite dimensional inner product space. Let $\{v_1, v_2, \dots v_n\}$ be a basis for V. From

this basis we shall construct an orthonormal basis $\{w_1, w_2, \dots, w_n\}$ by means of a construction known as Gram-Schmidt orthogonalisation process.

First, we take $w_1 = v_1$.

Let $w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1$.

We claim that $w_2 \neq 0$. For, if $w_2 = 0$ then v_2 is a scalar multiple of w_1 and hence of v_1 which is a contradiction since v_1, v_2 are linearly independent.

Also, $\langle w_2, w_1\rangle = \left\langle v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1, w_1 \right\rangle$

$$= \left\langle v_2 - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2} v_1, v_1 \right\rangle (\because w_1 = v_1)$$

$$= \langle v_2, v_1\rangle - \frac{\langle v_2, v_1\rangle}{\|v_1\|^2} \langle v_1, v_1\rangle$$

$$= \langle v_2, v_1\rangle - \langle v_2, v_1\rangle$$

$$= 0$$

Now, suppose that we have constructed non-zero orthogonal vectors $w_1, w_2, \dots, w_k$. Then put

$$w_{k+1} = v_{k+1} - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j\rangle}{\|w_j\|^2} w_j$$

We claim that $w_{k+1} \neq 0$. For, if $w_{k+1} = 0$, then v_{k+1} is a linear combination of $w_1, w_2, \dots, w_k$ and hence is a linear combination of $v_1, v_2, \dots, v_k$ which is a contradiction since $v_1, v_2, \dots, v_{k+1}$ are linearly independent.

Also

$$\langle w_{k+1}, w_i\rangle = \langle v_{k+1}, w_i\rangle - \sum_{j=1}^{k} \frac{\langle v_{k+1}, w_j\rangle}{\|w_j\|^2}\langle w_j, w_i\rangle$$
$$= \langle v_{k+1}, w_i\rangle - \frac{\langle v_{k+1}, w_i\rangle}{\|w_i\|^2}\langle w_i, w_i\rangle$$
$$= \langle v_{k+1}, w_i\rangle - \langle v_{k+1}, w_i\rangle$$
$$= 0$$

Thus, continuing in this way we ultimately obtain a non-zero orthogonal set $\{w_1, w_2, \dots, w_n\}$.

By Theorem 3 this set is linearly independent ar. hence a basis.

To obtain an orthonormal basis we replace each by $\frac{w_i}{\|w_i\|}$.

Solved problems

Problem 1:

Apply Gram-Schmidt process to construed an orthonormal basis for V_3(R) with the standard inne product for the basis $\{v_1, v_2, v_3\}$ where $v_1 = (1,0,1) v_2 = (1,3,1)$ and $v_3 = (3,2,1)$.

Solution:

Take $w_1 = v_1 = (1,0,1)$.

Then $\|w_1\|^2 = \langle w_1, w_1\rangle = 1^2 + 0^2 + 1^2 = 2$. and $\langle w_1, v_2\rangle = 1 + 0 + 1 = 2$

Put
$$w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1$$
$$= (1,3,1) - (1,0,1)$$
$$= (0,3,0)$$

$\therefore \|w_2\|^2 = 9$.

Also, $\langle w_2, v_3\rangle = 0 + 6 + 0 = 6$ and

$\langle w_1, v_3\rangle = 3 + 0 + 1 = 4$

$$w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2\rangle}{\|w_2\|^2} w_2$$

Now $$= (3,2,1) - \frac{4}{2}(1,0,1) - \frac{6}{9}(0,3,0)$$

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

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

$$\therefore\ \|w_3\|^2 = 2.$$

$\therefore$ The orthogonal basis is

$$\{(1,0,1), (0,3,0), (1,0,-1)\}$$

Hence the orthonormal basis is

$$\left\{\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), (0,1,0), \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}\right)\right\}.$$

Problem 2:

Let V be the set of all polynomials of degree ≤ 2 together with the zero polynomial. V is areal inner product space with inner product defined b) $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)dx$. Starting with the basis [1, x, x^2], obtain an orthonormal basis for V.

Solution:

Let $v_1 = 1; v_2 = x$ and $v_3 = x^2$.

Let $w_1 = v_1$.

Then $\|w_1\|^2 = \langle w_1, w_1\rangle = \int_{-1}^{1} 1dx = 2.$

Hence $\|w_1\| = \sqrt{2}.$

$$w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\|w_1\|^2} w_1$$
$$= x - \frac{1}{2}\int_{-1}^{1} x dx$$
$$= x$$

$\therefore\ \|w_2\|^2 = \langle w_2, w_2\rangle = \int_{-1}^{1} x^2 dx = \frac{2}{3}.$

Now, $w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2\rangle}{\|w_2\|^2} w_2$

$$= x^2 - \frac{1}{2}\int_{-1}^{1} x^2 dx - \left(\frac{3x}{2}\right)\int_{-1}^{1} x^3 dx$$
$$= x^2 - \frac{1}{3}$$

$$\|w_3\|^2 = \langle w_3, w_3\rangle = \int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45}$$

Hence the orthogonal basis is $\left\{1, x, x^2 - \frac{1}{3}\right\}$.

$\therefore$ The required orthonormal basis is

$$\left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{2}x, \frac{\sqrt{10}}{4}(3x^2 - 1)\right\}.$$

Problem 3:

Find a vector of unit length which is orthogonal to $(1,3,4)$ in $V_3(\mathbf{R})$ with standard inner roduct.

Solution:

Let $x = (x_1, x_2, x_3)$ be any vector orthogonal to $(1,3,4)$.

Then $x_1 + 3x_2 + 4x_3 = 0$. Any solution f this equation gives a vector orthogonal to $(1,3,4)$. or example $x = (1,1,-1)$ is orthogonal to $(1,3,4)$.

Also $\|x\| = \sqrt{3}$. Hence a unit vector orthogonal to $(1,2,3)$ is given by $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$.

Note: The set of all vectors orthogonal to $(1,3,4)$ are the points lying on the plane $x + 3y + 4z = 0$, which is a two-dimensional subspace of $V_3(\mathbf{R})$.

Problem 4:

Find an orthogonal basis containing the vector $(1,3,4)$ for $V_3(\mathrm{R})$ with the standard inner product.

Solution:

$(1,1,-1)$ is a vector orthogonal to $(1,3,4)$ (refer problem 3 above).

Now, let $y = (y_1, y_2, y_3)$ be a vector orthogonal to both (1,3,4) and ($1,1,-1$). Then $y_1 + 3y_2 + 4y_3 = 0$, $y_1 + y_2 - y_3 = 0$.

Any solution of this system of equations gives a vector orthogonal to $(1,3,4)$ and $(1,1,-1)$.

For example, $(7,-5,2)$ is one such vector. (by cross multiplication method).

Hence $\{(1,3,4), (1,1,-1), (7,-5,2)\}$ is an orthogonal basis containing (1,3,4).

Exercises

1.Applying Gram-Schmidt process find the orthonormal basis of $V_3(\mathbf{R})$ with the standard inner product starting with the following bases.

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

(b) $(2,-1,0), (4,-1,0), (4,0,-1)$

(c) $(1,0,1), (1,0,-1), (0,3,4)$

2.Let V bet the set of all polynomials of degree ≤ 2 over **R** with inner product defined by $\langle f, g\rangle = \int_0^1 f(x)g(x)dx$. Starting with the basis $\{1, x, x^2\}$ obtain an orthonormal basis for V.

3.Obtain an orthogonal basis for $V_3(\mathbf{R})$ with standard inner product containing the vectors.

(a) $(1,1,-1)$ and $(1,0,1)$ (b) $(7,-1,1)$.

Answers.

1. (a) $\{(\sqrt{2}/2, -\sqrt{2}/2, 0), (\sqrt{2}/6, \sqrt{2}/6, -2\sqrt{2}/3), (-2/3, -2/3, -1/3)\}$

(b) $\{(2\sqrt{5}/5, -\sqrt{5}/5, 0), (\sqrt{5}/5, 2\sqrt{5}/5, 0), (0,0,-1)\}$

(c) $(1/\sqrt{2}, 0, 1/\sqrt{2}), (1/\sqrt{2}, 0, -1/\sqrt{2}), (0,1,0)\}$

3. $\{1, (2x-1)\sqrt{3}, (5x^2 - 6x + 1)\sqrt{5}\}$.

3.5. Orthogonal Complement

Definition:

Let V be an inner product space. Let S be a subset of V. The orthogonal complement of S, denoted by $S^\perp$, is the set of all vectors in V which are orthogonal to every vector of S.

(i.e) $S^\perp = \{x / x \in V \text{ and } \langle x, u\rangle = 0 \text{ for all } u \in S\}$.

Examples

1. $V^\perp = \{0\}$ and $\{0\}^\perp = V$ since 0 is the only vector which is orthogonal to every vector.

2.Let $S = \{(x, 0,0)/x \in \mathbf{R}\} \subseteq V_3(\mathbf{R})$ with standard inner product. Then

$$S^{\perp} = \{(0, y, z)/y, z \in \mathbf{R}\}.$$

(i.e) The orthogonal complement of the x-axis is the yz plane.

Theorem 7:

If S is any subset of V then $S^{\perp}$ is a subspace of V.

Proof:

Clearly $0 \in S^{\perp}$ and hence $S^{\perp} \neq \Phi$.

Now, let $x, y \in S^{\perp}$ and $\alpha, \beta \in F$.

Then $\langle x, u\rangle = \langle y, u\rangle = 0$ for all $u \in S$.

$\therefore \langle \alpha x + \beta y, u\rangle = \alpha\langle x, u\rangle + \beta\langle y, u\rangle = 0$ for all $u \in S$.

$\therefore$ $\alpha x + \beta y \in S^{\perp}$. Hence $S^{\perp}$ is a subspace of V.

Theorem 8:

Let V be a finite dimensional into, product space. Let W be a subspace of V. Thenv is the direct sum of W and $W^{\perp}$ (i.e.) $V = W_{\oplus} W_1$

Proof:

We shall prove that

(i) $W \cap W^{\perp} = \{0\}$, and

(ii) $W + W^{\perp} = V$.

(i) Let $v \in W \cap W^{\perp}$. Then $v \in W$ and $v \in W \perp$

Now, $v \in W^{\perp} \Rightarrow v$ is orthogonal to evers) vector in W.

In particular, v is orthogonal to itself.

$\therefore$ $\langle v, v\rangle = 0$ and hence $v = 0$.

Hence $W \cap W^{\perp} = \{0\}$.

(ii) Let $\{v_1, v_2, \dots \dots, v_r\}$ be an orthonormal basis for W. Let $n \in V$.

Consider

$v_0 = v - \langle v, v_1\rangle v_1 - \langle v, v_2\rangle v_2 \ldots, -\langle v, v_r|_v,$

$\therefore\ \langle v_0, v_i\rangle = \langle v, v_i\rangle - \langle v, v_i\rangle\langle v_i, v_i\rangle$

(since $\langle v_i, v_j\rangle = 0$ if $i \neq j$)

$= \langle v, v_i\rangle - \langle v, v_i\rangle$

(since $\langle v_i, v_i\rangle = 1$)$= 0$

$\therefore\ v_0$ is orthogonal to each of v_1, v_2, ..., and hence is orthogonal to every vector in W Hence $v_0 \in \mathbb{W}^{\perp}$ and

$$v = [(v, v_1)v_1 + \langle v, v_2\rangle v_2 + \cdots \ldots + \langle v, v_r\rangle v_r] + v_0 \in W + W^{\perp}$$

$\therefore\ V = W \oplus W^{\perp}.$

Hence the theorem.

Corollary. $\dim V = \dim W + \dim W^{\perp}$.

Proof. $\dim V = \dim(W \oplus W^{\perp}) = \dim W + \dim W^{\perp}$

Theorem 9:

Let V be a finite dimensional inner product space. Let W be a subspace of V. Ther $(W^{\perp})^{\perp} = W$.

prof. Let $w \in W$. Then for any $u \in W^{\perp}, \langle w, u\rangle = 0$ Hence $w \in (W^{\perp})^{\perp}$). Thus $W \subseteq (W^{\perp})^{\perp}$

Now by Theorem 8, $V = W \oplus W^{\perp}$.

Also $V = W^{\perp} \oplus (W^{\perp})^{\perp}$.

Hence $\dim W = \dim(W^{\perp})^{\perp}$

From (1) and (2) we get $W = (W^{\perp})^{\perp}$.

Solved Problems

Problem 1: Let V be an inner product space and let S_1 and S_2 be subsets of V.

Then $S_1 \subseteq S_2 \Rightarrow S_2^{\perp} \subseteq S_1^{\perp}$.

Solution:

Let $u \in S_2^{\perp}$

Then $\langle u, v\rangle = 0$ for all $v \in S_2$.

But $S_1 \subseteq S_2$. Hence $\langle u, v\rangle = 0$ for all $v \in S_1$.

Hence $u \in S_1^{\perp}$. Thus $S_2^{\perp} \subseteq S_1^{\perp}$.

Problem 2:

Let W_1 and W_2 be subspaces of a finite dimensional inner product space. Then

(i) $(W_1 + W_2)^{\perp} = W_1^{\perp} \cap W_2^{\perp}$.

(ii) $(W_1 \cap W_2)^{\perp} = W_1^{\perp} + W_2^{\perp}$.

Solution:

(i) We know that $W_1 \subseteq W_1 + W_2$.

$\therefore$ $(W_1 + W_2)^{\perp} \subseteq W_1^{\perp}$ (by the problem 1).

Similarly, $(W_1 + W_2)^{\perp} \subseteq W_2^{\perp}$

Hence $(W_1 + W_2)^{\perp} \subseteq W_1^{\perp} \cap W_2^{\perp}$. ……… (1)

Now, let $w \in W_1^{\perp} \cap W_2^{\perp}$.

Then $w \in W_1^{\perp}$ and $w \in W_2^{\perp}$.

$\therefore$ $\langle w, u\rangle = 0$ for all $u \in W_1$ and W_2.

Now, let $v \in W_1 + W_2$.

Then $v = v_1 + v_2$ where $v_1 \in W_1$ and $v_2 \in W_2$.

$$\begin{aligned}\therefore \langle w, v\rangle &= \langle w, v_1 + v_2\rangle \\ &= \langle w, v_1\rangle + \langle w, v_2\rangle \\ &= 0 + 0 (\text{ since } v_1 \in W_1 \text{ and } v_2 \in W_2) \\ &= 0\end{aligned}$$

Hence $w \in (W_1 + W_2)^{\perp}$.

$\therefore$ $W_1^{\perp} \cap W_2^{\perp} \subseteq (W_1 + W_2)^{\perp}$ ……….. (2)

From (1) and (2) we get

$(W_1 + W_2)^1 = W_1^1 \cap W_2^1$.

(ii) Proof is similar to that of (i).

Exercises

1.Let V be an inner product space. Let S be any subspace of V. Then show that $S^{\perp} = [L(S)]^{\perp}$

2.Find a basis for the orthogonal complement of the subspace spanned by $(2,1,-2)$ in $V_3(\mathbf{R})$.

UNIT IV

Matrices – Elementary transformation – Rank of a matrix – Simultaneous linear equations–Characteristic equation and Cayley-Hamilton Theorem.

Chapter 4: Sections -4.1 to 4.4

4.1. Elementary Transformations:

Definition:

Let A be an $m \times n$ matrix over a field F. An elementary row-operation on A is of any one of the following three types.

1. The interchange of any two rows.
2. Multiplication of a row by a non-zero element c in E
3. Addition of any multiple of one row with any other row.

Similarly, we define an elementary column operation on A as any one of the following three types.

1. The interchange of any two columns.
2. Multiplication of a column by a non-zero element c in F.
3. Addition of any multiple of one column with any other column.

Example:

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & -1 \end{pmatrix}, A_1 = \begin{pmatrix} 3 & -1 \\ 2 & 1 \\ 1 & 2 \end{pmatrix}$,

$A_2 = \begin{pmatrix} 2 & 2 \\ 4 & 1 \\ 6 & -1 \end{pmatrix}$ $A_3 = \begin{pmatrix} 1 & 2 \\ 5 & 7 \\ 3 & -1 \end{pmatrix}$. A_1 is obtained from A by interchanging the first and third rows.

A_2 is obtained from A by multiplying the first column of A by 2 .

A_3 is obtained from A by adding to the second row the multiple by 3 of the first row.

Notation.

We shall employ the following notations for elementary transformations.

(i) Interchange of i^{th} and j^{th} rows will be denoted by $R_i \leftrightarrow R_j$.

(ii) Multiplication of i^{th} row by a non-zero element $c \in F$ will be denoted by

$R_i \to cR_i$.

(iii) Addition of k times the j^{th} row to the i^{th} row will be denoted by

$R_i \to R_i + kR_j$.

The corresponding column operations will be denoted by writing C in the place of R.

Definition:

An $m \times n$ matrix B is said to be row equivalent (column equivalent) to an $m \times n$ matrix A if B can be obtained from A by a finite succession of elementary row operations (column operations).

A and B are said to be equivalent if B can be obtained from A by a finite succession of elementary row or column operations.

If A and B are equivalent. We write $A \sim B$.

Exercise:

Prove that row equivalence, column equivalence and equivalence are equivalence relations in the set of all $m \times n$ matrices.

Definition:

A matrix obtained from the identity matrix by applying a single elementary row or column operation is called an elementary matrix.

For example, $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$ are elementary matrices obtained from the identity matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

by applying the elementary operations $R_1 \leftrightarrow R_2$,

$$R_1 \to 4R_1, R_3 \to R_3 + 2R_2 \text{ respectively.}$$

Exercise. Give examples of elementary matrices of order 4.

Theorem 1:

Any elementary matrix is nonsingular

Proof:

The determinant of the identity matrix of order is 1. Hence the determinant of an elementary matrix obtained by interchanging any two rows is the determinant of an elementary matrix obtained multiplying any row by $k \neq 0$ is k. The determinant an elementary matrix obtained by adding a multiple one row with another row is 1. Hence any element matrix is non-singular.

Theorem 2:

Let A be an $m \times n$ matrix and B an $n \times p$ matrix. Then every elementary row (colum) operation of the product AB can be obtained by su jecting the matrix A (matrix B) to the same element row (column) operation.

Proof:

Let $R_1, R_2, \ldots, R_m$ denote the rows the matrix A and $C_1, C_2, \ldots \ldots, C_p$ denote columns of B. By the definition of matrix multip cation

$$AB = \begin{pmatrix} R_1C_1 & R_1C_2 & \dots\dots & R_1C_p \\ R_2C_1 & R_2C_2 & \dots\dots & R_2C_p \\ . & . & \dots\dots & . \\ . & . & \dots\dots & . \\ . & . & \dots\dots & . \\ R_mC_1 & R_mC_2 & \dots\dots & R_mC_p \end{pmatrix}$$

It is obvious from the above representation of AB th if we apply any elementary row operation on A ti matrix AB is also subjected to the same elementad row operation. Also, if we apply any elementary co umn operation on B the matrix AB is also subjects to the same elementary column operation.

Theorem 3:

Each elementary row operation on as $m \times n$ matrix A is equivalent to pre-multiplying the matrix A by the corresponding elementary $m \times$ matrix.

Proof:

Since A is an $m \times n$ matrix we can write $A = IA$ where I is the identity matrix of ord m. By theorem 2 an elementary row operation I is equivalent to the same row operation on I. But elementary row operation on I gives an elementat matrix. Hence by pre-multiplying A by the cont sponding elementary matrix we get the required of operation on A.

Note:

Similarly, each elementary column operation of an $m \times n$ matrix A is equivalent to post-multiplying the matrix A by the corresponding elementary $n \times n$ matrix.

Corollary 1:

If two $m \times n$ matrices A and B are row equivalent then $A = PB$ where P is a non-singular $m \times m$ matrix.

Proof:

Since A is row equivalent to B, A can be obtained from B by applying successive elementary row operations. Hence $A = E_1 E_2 \dots E_n B$ where each E_i is an elementary matrix. Since each E_i is nonsingular, $A = PB$ where $P = E_1 E_2 \dots E_n$ and P is non-singular.

Corollary 2:

If two matrices A and B are column equivalent then $A = BQ$ where Q is a non-singular matrix.

Corollary 3:

If two $m \times n$ matrices A and B are equivalent then $A = PBQ$ where P is a non-singular $m \times m$ matrix and Q is a non- singular $n \times n$ matrix.

Corollary 4: The inverse of an elementary matrix is again an elementary matrix.

Proof:

Let E be an elementary matrix obtained from I by applying some elementary operations. If we apply the reverse operation on E, then E is carried back to 1. Let E^* be the elementary matrix corresponding to the reverse operation.

Then $E^*E = EE^* = I$. Hence $E^* = E^{-1}$.

Hence E^{-1} is also an elementary matrix.

Canonical form of a matrix.

We now use elementary row and column operations to reduce any matrix to a simple form, called the canonical form of a matrix.

Theorem 4:

By successive applications of elementary row and column operations, any non-zero $m \times n$ matrix A can be reduced to a diagonal matrix D in which the diagonal entries are either 0 or 1 and all the l's proceeding all the zeros on the diagonal. In other words, any non-zero $m \times n$ matrix is equivalent to a matrix of the form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ where I_r is the $r \times r$ identity matrix and O is the zero matrix.

Proof:

We shall prove the theorem by induction on the number of rows of A. Suppose A has just one row. Let $A = (a_{11}a_{12} \dots \dots a_{1n})$.

Since $A \neq 0$, by interchanging columns, if necessary, we can bring a non-zero entry c to the position a_{11}.

Multiplying A by c^{-1} we get 1 as the first entry. Other entries in A can be made zero by adding suitable multiples of 1.

Thus, the result is true when $m = 1$.

Now, suppose that the result is true for any non-zero matrix with $m - 1$ rows.

Let A be a non-zero $m \times n$ matrix. By permuting rows and columns, we can bring some non-zero entry c to the position a_{11}.

Multiplying the first row by c^{-1} we get 1 as the first entry.

All other entries in the first column can be made zero by adding suitable multiples of the first row to each other row.

Similarly, all the other entries in the first row can be made zero.

This reduces A to a matrix of the form

$B = \begin{bmatrix} I_1 & O \\ O & C \end{bmatrix}$ where C is an $(m-1) \times (n-1)$ matrix.

Now by induction hypothesis C can be reduced to the desired form by elementary row and column operations.

Hence A is equivalent to a matrix of the required form.

Corollary 1: If A is an $m \times n$ matrix there exist nonsingular square matrices P and Q of orders m and n respectively such that $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

The result follows from corollary 3 of theorem 7.25.

Corollary 2:

Any non-singular square matrix A of order n is equivalent to the identity matrix.

Proof:

By corollary 1, $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.Since P, A, Q are all non-singular $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ is non-singular. This is possible iff $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = I_n$.

Corollary 3:

Any non-singular matrix A can be expressed as a product of elementary matrices.

Proof:

By corollary 2, $PAQ = I_n$. Hence

$A = P^{-1}Q^{-1}$. P^{-1} and Q^{-1} are products of elementary matrices. Hence A is a product of elementary matrices.

Note:

The inverse of a non-singular matrix A can be computed by using elementary transformations. Let A be a non-singular matrix of order n.

Then $AA^{-1} = A^{-1}A = I$. Now, the non-singular matrix A^{-1} can be expressed as the product of elementary matrices.

Let $A^{-1} = E_1E_2 \dots \dots E_n$.Then $I = A^{-1}A = E_1E_2 \dots \dots E_nA$.

Thus every non-singular matrix A can be reduced to I by pre-multiplying A by elementary matrices. Hence A can be reduced to the identity matrix by applying successive elementary row operations. Now, $A = IA$. Reduce the matrix A in the left-hand side to I by applying successive elementary row operations and apply the same elementary row operations to the factor I in the right-hand side.

Then we get $I = BA$ so that $B = A^{-1}$.

Solved problems

Problem 1: Reduce the matrix $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{pmatrix}$ to the canonical form.

Solution: $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{pmatrix}$

$$\sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{matrix} C_2 \to C_2 - 2C_1 \\ C_3 \to C_3 + C_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad C_3 \to C_3 + 3C_2$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad R_2 \to -R_2$$

Problem 2: Find the inverse of the matrix $A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{pmatrix}$

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

$$\Rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 1 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} A \quad R_2 \to R_2 - 3R_1 \,,\, R_3 \to R_3 + 2R_1$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 0 & 14 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 5 & -1 & 1 \end{pmatrix} A$$

$R_3 \to R_3 - R_2$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{pmatrix} R_1 \to R_1 - \frac{1}{7}R_3 \; R_2 \to R_2 + \frac{1}{2}R_3 \; R_3 \to \frac{1}{14}R_3$$

$$\Rightarrow A^{-1} = \begin{pmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{pmatrix}$$

Exercises:

Now, let $A, B, C \in S$.

1. Write down six matrices (not necessarily square matrices) and reduce them to the canonical form.
2. Find the inverse of the following matrices by using elementary operations.

(a) $\begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{pmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Answers:

2. (a) $\begin{pmatrix} -9 & 8 & -5 \\ -8 & 7 & -4 \\ -2 & 2 & -1 \end{pmatrix}$ (b) $\frac{1}{2}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$

Definition:

Let A and B be two square matrices of order $n. B$ is said to be similar to A if there exists a $n \times n$ non-singular matrix P such that $B = P^{-1}AP$.

Solved problems

Problem 1:

Similarity of matrices is an equivalence relation in the set of all $n \times n$ matrices.

Proof:

Let S be the set of all $n \times n$ matrices.

Let $A \in S$.

Since $A = I^{-1}AI$ and I is non-singular, A is similar to A.

Hence similarity of matrices is reflexive.

Now, let $A, B \in S$ and let A be similar to B.

$\therefore\ A = P^{-1}BP$ where $P \in S$ is a non-singular matrix.

Now,

$$\begin{aligned} P^{-1}BP = A &\Rightarrow PP^{-1}BPP^{-1} = PA \\ &\Rightarrow B = PAP^{-1} \\ &\Rightarrow B = (P^{-1})^{-1}A(P^{-1}) \end{aligned}$$

Since P is non-singular $P^{-1} \in S$ is also non-singular.

$\therefore\ B$ is similar to A.

Hence similarity of matrices is symmetric.

Let A be similar to B to B be similar to C. Hence there exist non-singular matrices $P, Q \in S$ such that

$A = P^{-1}BP$ and $B = Q^{-1}CQ$.

Now,
$$\begin{aligned} A &= P^{-1}BP \\ &= P^{-1}(Q^{-1}CQ)P \\ &= (P^{-1}Q^{-1})CQP \\ &= (QP)^{-1}C(QP) \end{aligned}$$

Since $P, Q \in S$ are non-singular, $QP \in S$ is alsc non-singular.

Hence A is similar to C.

$\therefore$ Similarity of matrices is transitive.

Hence similarity of matrices is an equivalence relation.

Problem 2:

If A and B are similar matrices show tha their determinants are same.

Solution. Let A and B be two similar matrices.

$\therefore$ There exists a non-singular matrix P such tha $B = P^{-1}AP$.

Now,
$$\begin{aligned} |B| &= |P^{-1}AP| \\ &= |P^{-1}||A||P| \\ &= |A| \ \left(\text{since } |P^{-1}| = \frac{1}{|P|}\right) \end{aligned}$$

Hence the result.

4.2. Rank of a Matrix

We now proceed to introduce the concept of the rank of a matrix.

Definition:

Let $A = (a_{ij})$ be an $m \times n$ matrix. The rows $R_i = (a_{i1}, a_{i2}, \dots, \dots, a_{in})$ of A can be thought of as elements of F^n. The subspace of F^n generated by the m rows of A is called the row space of A.

Similarly, the subspace of F^m generated by the n columns of A is called the column space of A,

The dimension of the row space (column space) of A is called the row rank (column rank) of A.

Theorem 1:

Any two row equivalent matrices have the same row space and have the same row rank.

Proof:

Let A be an $m \times n$ matrix.

It is enough if we prove that the row space of A is not altered by any elementary row operation.

Obviously, the row space of A is not altered by an elementary row operation of the type $R_i \leftrightarrow R_j$.

Now, consider the elementary row operation

$R_i \rightarrow cR_i$ where $c \in F - \{0\}$.

Since $L(\{R_1, R_2, \dots, R_i, \dots \dots, R_n\}) = L(\{R_1, R_2, \dots \dots \dots cR_i, \dots \dots, R_n\})$ the row space of A is not altered by this type of elementary row operation.

Similarly, we can easily prove that the row space of A is not altered by an elementary row operation of the type $R_i \rightarrow R_i + cR_i$.

Hence row equivalent matrices have the same row space and hence the same row rank.

Similarly, we can prove the following theorem.

Theorem 2:

Any two column equivalent matrices have the same column rank.

Theorem 3:

The row rank and the column rank of any matrix are equal.

Proof:

Let $A = (\,^*, j)$ be an $m \times n$ matrix.

Let $R_1, R_2, \dots \dots \dots, R_m$ denote the rows of A.

Hence $R_i = (a_{i1}, a_{i2}, \dots \dots, a_{in})$.

Suppose the row rank of A is r.

Then the dimension of the row space is r.

Let $v_1 = (b_{11}, \dots, b_{1n}), v_2 = (b_{21}, \dots, b_{2n}), \dots \dots, v_r = (b_{r1}, \dots \dots \dots, b_{rn})$ be a basis for the row space of A.

Then each row is a linear combination of the vectors $v_1, v_2, \dots \dots \dots, v_r$.

Let, $R_1 = k_{11}v_1 + k_{12}v_2 + \cdots \cdots + k_{1r}v_r$
$R_2 = k_{21}v_1 + k_{22}v_2 + \cdots \cdots + k_{2r}v_r$

....

$R_m = k_{m1}v_1 + k_{m2}v_2 + \cdots \cdots + k_{mr}v_r$
where $k_{ij} \in F$.

Equating the i^{th} component of each of the above equations, we get

$$a_{1i} = k_{11}b_{1i} + k_{12}b_{2i} + \cdots \cdots + k_{1r}b_{ri}$$
$$a_{2i} = k_{21}b_{1i} + k_{22}b_{2i} + \cdots \cdots + k_{2r}b_{ri}$$
$$\cdots \cdots \cdots \cdots \cdots$$
$$\dots \cdots \cdots \cdots \cdots \cdots$$
$$a_{mi} = k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots \cdots + k_{mr}b_{ri}$$

Hence

$$\begin{pmatrix} a_{1i} \\ \cdot \\ a_{mi} \end{pmatrix} = b_{1i}\begin{pmatrix} k_{11} \\ \cdot \\ k_{m1} \end{pmatrix} + b_{2i}\begin{pmatrix} k_{12} \\ \cdot \\ k_{m2} \end{pmatrix} + \cdots + b_{ri}\begin{pmatrix} k_{1r} \\ \cdot \\ \cdot \\ k_{mr} \end{pmatrix}$$

Thus, each column of A is a linear combination of r vectors.

Hence the dimension of the column space $\leq r$.

$\therefore$ Column rank of $A \leq r =$ row rank of A.

Similarly, row rank of $A \leq$ column rank of A.

Hence the row rank and the column rank of A are equal.

Definition:

The rank of a matrix A is the common value of its row and column rank.

Note 1:

Since the row rank and the column rank of a matrix are unaltered by elementary row and column operations, equivalent matrices have the same rank. In particular if a matrix A is reduced to its canonical form $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, then rank of $A = r$.

Thus, to find the rank of a matrix A, we reduce A to the canonical form and find the number of non-zero entries in the diagonal.

Note that in the canonical form of the matrix A, there exists an $r \times r$ sub-matrix, namely, I_r, whose determinant is not zero.

Further every $(r + 1) \times (r + 1)$ sub-matrix contains row of zeros and hence its determinant is zero.

Also under any elementary row or column operathe value of a determinant is either unaltered or whiplied by a non-zero constant.

Hence the matrix A is also such that

(i) there exists an $r \times r$ sub-matrix whose determinant is nonzero.

(ii) the determinant of every $(r + 1) \times (r + 1)$ sub-matrix is zero.

Hence one can also define the rank of a matrix A to be if A satisfies (i) and (ii).

Note 2:

Any non-singular matrix of order n is equivlent to the identity matrix and hence its rank is n.

Note 3.

The rank of a matrix is not altered on multipliation by non-singular matrices. since premultiplicaon by a non-singular matrix is equivalent to applying dementary row operaptions and post-multiplication y a non- singular matrix is equivalent to applying dementary column operations.

Solved problems

Problem 1:

Find the rank of the matrix

$$A = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{pmatrix}$$

Solution:

$$A = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 2 & 4 & 3 \\ 4 & 3 & 6 & 7 \\ 0 & 1 & 2 & 7 \end{pmatrix} C_1 \leftrightarrow C_3$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 4 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{pmatrix} \begin{matrix} C_1 \to C_2 - 2C_1 \\ C_3 \to C_3 - 4C_1 \\ C_4 \to C_4 - 3C_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{pmatrix} R_3 \to R_2 - 4R_1$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 1 & 0 & 6 \end{pmatrix} C_3 \to C_3 - 2C_2$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} R_3 \to C_2$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 6 & 0 \end{pmatrix} C_2 \to C_3$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} R_2 \to -\frac{1}{5}R_2, R_3 \to \frac{1}{6}R_3$$

$\therefore$ Rank of $A = 3$.

Problem 2:

Find the rank of the matrix

$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 1 & 0 & 2 \\ 0 & 3 & 4 & 2 \end{pmatrix}$ by examining the determinant minors.

Solution:

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

$\therefore$ Every 3×3 submatrix of A has determinant zero.

Also, $\begin{vmatrix} 1 & 1 \\ 4 & 1 \end{vmatrix} = -3 \neq 0.$

$\therefore$ Rank of $A = 2$.

Exercises

1. Determine the rank of any six matrices of your choice.
2. Find the rank of the following matrices,

(a) $\begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & -3 \\ 6 & -1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 & 2 & 1 \\ 2 & -3 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix}$

(c) $\begin{pmatrix} 1 & -1 & 0 & 2 & 1 \\ 3 & 1 & 1 & -1 & 2 \\ 4 & 0 & 1 & 0 & 3 \\ 9 & -1 & 2 & 3 & 7 \end{pmatrix}$

3. Find the column rank of the matrices

(a) $\begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{pmatrix}$ (b) $\begin{pmatrix} 3 & 1 & -5 & -1 \\ 1 & -2 & 1 & -5 \\ 1 & 5 & -7 & 2 \end{pmatrix}$

4. Find the row rank of the matrix

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

Answers. 2.(a) 2 (b) 3 (c) 3 3. (a) 2 (b) 34.2

4.3. Simultaneous Linear Equations:

In this section we shall apply the theory of matrices developed in the proceeding sections to study the existence of solutions of simultaneous linear equations.

Matrix form of a set of linear equations.

Consider a system of m linear equations in n unknowns $x_1, x_2,.. x_n$ given by

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots\cdots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \cdots\cdots + a_{2n}x_n = b_2 \\ &\ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ &\ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ &a_{m1}x_1 + a_{m2}x_2 + \cdots\cdots + a_{mn}x_n = b_m \end{aligned}$$

Using the concept of matrix multiplication and equality of matrices this system can be written as $AX = B$ where,

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \cdot & \cdot & \ldots\cdots & \cdot \\ \cdot & \cdot & \ldots\cdots & \cdot \\ a_{m1} & a_{m2} & \ldots\cdots & a_{mn} \end{pmatrix}$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}, B = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ b_m \end{pmatrix}$$

The $m \times n$ matrix A is called the coefficient matrix

Definition:

A set of values of $x_1, x_2, \dots \dots \dots, x_1$ which satisfy the above system of equations is calloa solution of the system. The system of equations i said to be consistent if it has at least one solution Otherwise the system is said to be inconsistent.

The $m \times (n+1)$ matrix given by

$$\begin{pmatrix} a_{11} & \dots \dots & a_{1n} & b_1 \\ a_{21} & \dots \dots & a_{2n} & b_2 \\ \cdot & \dots \dots & \cdot & \cdot \\ \cdot & \dots \dots & \cdot & \cdot \\ a_{m1} & \dots \dots & a_{mn} & b_m \end{pmatrix}$$

is called the augmented matrix of the system and is denoted by (A, B).

Thus, the augmented matrix (A, B) is obtained by annexing to A the column matrix B, which becomes the $(n+1)^{\text{th}}$ column in (A, B).

Note. Since every column in A appears in (A, B) the column space of the matrix A is a subspace of the column space of the matrix (A, B).Hence the rank of $A \leq$ rank of (A, B).

Theorem 1:

The system of linear equations

$AX = B$ is consistent iff rank of $A =$ rank of (A, B).

Proof:

Let the system be consistent.

Let $u_1, u_2, \dots \dots, u_n$ be a solution of the system.

Then $B = u_1 C_1 + u_2 C_2 + \cdots \cdots + u_n C_n$ where $C_1, C_2, \cdots, C_n$ denote the columns of A,

Hence the column space of the augmented matrix (A, B), namely

$\langle C_1, C_2, \dots \dots, C_n, B \rangle$ is the same as the column space $\langle C_1, C_2, \dots \dots, C_n \rangle$ of A.

Hence the rank of $A =$ rank of (A, B).

Conversely let rank of $A =$ rank of (A, B).

Then the column rank of $A =$ column rank of (A, B).

$\therefore \dim\langle C_1, C_2, \dots, C_n\rangle = \dim\langle C_1, C_2, \dots, C_n, B\rangle$.

But $\langle C_1, C_2, \dots, C_n\rangle$ is a subspace of $\langle C_1, C_2, \dots \dots, C_n, B\rangle$.

$\therefore B$ is a linear combination of $C_1, C_2, \dots \dots, C_n$.

If $B = u_1C_1 + \cdots + u_nC_n$ then $u_1, u_2, \dots \dots, u_n$ is a solution of the system.

Hence the theorem.

Remark:

The solution of a given system of simultaneous equations is not altered by interchanging any two equations or by multiplying any equation by a nonzero constant or by adding a multiple of one equation to another. Hence, we can reduce the given system of equations to an equivalent system by applying elementary row operations to the augmented matrix. This reduced form will enable us to test for the consistency and to find the solution if it exists. This is illustrated in the following problems.

Solved problems

Problem 1:

Show that the equations
$$\begin{aligned} x + y + z &= 6 \\ x + 2y + 3z &= 14 \\ x + 4y + 7z &= 30 \end{aligned}$$
are consistent and solve them.

Solution:

The given system of equations can be put in the matrix form

$$AX = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 14 \\ 30 \end{pmatrix} = B.$$

∴ The augmented matrix is given by

$$(A, B) = \begin{pmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 14 \\ 1 & 4 & 7 & 30 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 3 & 6 & 24 \end{pmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix} R_3 \to R_3 - 3R_2$$

Hence rank of $A =$ rank of $(A, B) = 2$.

Hence the given system is consistent.

Also the given system of equations reduces to

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 0 \end{pmatrix}$$

$$\therefore\ x + y + z = 6$$

$$y + 2z = 8.$$

Putting $z = c$ we obtain the general solution of the system as

$x = c - 2, y = 8 - 2c, z = c$.

Problem 2:

Verify whether the following system of equations is consistent. If it is consistent, find the solution.

$$x - 4y - 3z = -16$$
$$4x - y + 6z = 16$$
$$2x + 7y + 12z = 48$$
$$5x - 5y + 3z = 0$$

Solution. The matrix form of the system is given by

$$\begin{pmatrix} 1 & -4 & -3 \\ 4 & -1 & 6 \\ 2 & 7 & 12 \\ 5 & -5 & 3 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 16 \\ 48 \\ 0 \end{bmatrix}$$

∴ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & -4 & -3 & -16 \\ 4 & -1 & 6 & 16 \\ 2 & 7 & 12 & 48 \\ 5 & -5 & 3 & 0 \end{bmatrix}$$

$$\begin{pmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \end{pmatrix} R_2 \to R_2 - 4R_1, R_3 \to R_3 - 2R_1, R_4 \to R_4 - 5R_1$$

$$\begin{pmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} R_3 \to R_3 - R_2, R_4 \to R_4 - R_2$$

Rank of A $=$ Rank of (A, B) $= 2$ and hence the system is consistent. Also, the system of equations reduces to

$$\begin{pmatrix} 1 & 4 & 3 \\ 0 & 15 & 18 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 16 \\ 80 \\ 0 \\ 0 \end{pmatrix}$$

∴ $x = 4y = 3z = -16$ and $15y + 18z = 80$.

Putting $z = 0$ we obtain the general solution of the systems as

$x = -(9c/5) + (16/3).$

$y = (6c/5) + (16/3)$

$z = c$

Problem 3:

For What values of is the equations

$$x + y + z = 1$$
$$x + 2y + 4z = n$$
$x + 4y + 10z = n^2$ are connistent?

Solution:

The matrix form of the system in given by

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ y \\ y^2 \end{pmatrix}$$

The augmented matrix is given by

$$(A,B) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \eta \\ 1 & 4 & 10 & \eta^2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta - 1 \\ 0 & 3 & 9 & \eta^2 - 1 \end{pmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta - 1 \\ 0 & 0 & 0 & \eta^2 - 3\eta + 2 \end{pmatrix} R_1 \to R_3 - 3R_2$$

The given system is consistent iff $n^2 - 3n + 2 - 0$

$$\eta = 2 \text{ or } 1\,.$$

Problem 4:

Show that the system of equations

$$x + 2y + z = 11$$
$$4x + 6y + 5z = 11$$
$2x + 2y + 3z = 19$ is inconsistent.

Solution:

The matrix form of the system is given by

$$\begin{pmatrix} 1 & 2 & 1 \\ 4 & 6 & 5 \\ 2 & 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 11 \\ 8 \\ 19 \end{pmatrix}$$

∴ The augmented matrix is given by

$$(A, n) = \begin{pmatrix} 1 & 2 & 1 & 11 \\ 4 & 6 & 5 & 8 \\ 2 & 2 & 3 & 19 \end{pmatrix}$$

$$\therefore \begin{pmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & -2 & 1 & -3 \end{pmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_1 - 2R_1 \end{matrix}$$

$$\therefore \begin{pmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & 0 & 0 & 33 \end{pmatrix} R_3 \to R_3 - R_2$$

∴ Rank of $\Lambda = 2$ and rank of $(\Lambda, B) = 3$.

∴ The given system is inconsistent.

Exercises

1.Solve or prove the inconsistency of the following systems of equations.

(a) $2x - y + 3z = 8$

$x - 2y - z = 4$

$3x + y = 45 = 0$

(b) $x + 2y - 5z = 0$

$3x + 4y + 6z = 0$
$x + y + z = 0$

(c) $x + 2y - 5z = -9$

$3x - y + 2z = 5$
$2x + 3y - z = 3$
$4x - 5y + z = -3$

(d) $x + y + z = 1$

$x + 2y + 3z = 1$

$x + 3y + 5z = 7$

$x + 4y + 7z = 10$

(e) $x - 2y - z - t = -1$

$3x - 2z + 3t = -4$
$5x - 4y + t = -3$

(f)

$$\begin{aligned} x + y + z &= 7 \\ x + 2y + 3z &= 8 \\ y + 2z &= 6 \end{aligned}$$

2.For what values of λ and μ the system of equations

$$\begin{aligned} x + y + z &= 6 \\ x + 2y + 3z &= 10 \\ x + 2y + \lambda z &= \mu \end{aligned}$$

is (a) inconsistent (b) consistent (c) consistent and the solution is unique.

3. Show that the set of all solutions of the system of homogeneous equations $AX = 0$ forms a vector space.

4. Show that a system of n equations in n unknowns given by $AX = Y$ has a unique solution if the $n \times n$ matrix A is nonsingular.

aswers.

1. (a) consistent; $x = y = z = 2$

   (b) consistent; $x = y = z = 0$;

   (c) consistent; $x = \frac{1}{2}, y = \frac{3}{2}, z = \frac{5}{2}$;

   (d) consistent; $x = c - 2, y = 3 - 2c; z = c$

   (e) inconsistent;

   (f) inconsistent;

   (g) inconsistent.

2. If $\lambda = 3$ and $\mu \neq 10$, inconsistent.

If $\lambda = 3$ and $\mu = 10$, consistent.

If $\lambda \neq 3$, consistent and the solution is unique.

4.4. Characteristic Equation and Cayley Hamilton Theorem

Definition:

An expression of the form $A_0 + A_1x + A_2x^2 + \cdots + A_nx^n$ where $A_0, A_1, \dots, A_n$ are square matrices of the same order and $A_n \neq 0$ is called a matrix polynomial of degree n.

For example, $\begin{pmatrix}1 & 2\\0 & 3\end{pmatrix} + \begin{pmatrix}1 & 1\\2 & 1\end{pmatrix}x + \begin{pmatrix}2 & 0\\3 & 1\end{pmatrix}x^2$ is a matrix polynomial of degree 2 and it is simply the matrix $\begin{pmatrix}1 + x + 2x^2 & 2 + x\\2x + 3x^2 & 3 + x + x^2\end{pmatrix}$.

Definition:

Let A be any square matrix of order n and let I be the identity matrix of order n. Then the matrix polynomial given by $A - xI$ is called the characteristic matrix of A. The determinant $|A - xI|$ which is an ordinary polynomial in x of degree n is called the characteristic polynomial of A.

The equation $|A - xI| = 0$ is called the characteristic equation of A.

Example 1:

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

Then the characteristic matrix of A is $A - xI$ given by

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

$\therefore$ The characteristic polynomial of A is

$$\begin{aligned} |A - x|\,| &= \begin{vmatrix}1 - x & 2\\3 & 4 - x\end{vmatrix}\\ &= (1 - x)(4 - x) - 6\\ &= x^2 - 5x - 2\end{aligned}$$

$\therefore$ The characteristic equation of A is $|A - xI| = 0$

$\therefore$ $x^2 - 5x - 2 = 0$ is the characteristic equation of A.

Example 2:

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

The characteristic matrix of A is $A - xI$ given by

$$A - xI = \begin{pmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{pmatrix}$$

The characteristic polynomial of A is

$$\begin{aligned} |A - xI| &= \begin{vmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{vmatrix} \\ &= (1-x)[(1-x)(-x) - 4] - 2(1-x)] \\ &= -x(1-x)^2 - 4(1-x) - 2 + 2x \\ &= -x^3 + 2x^2 - x - 4 + 4x - 2 + 2x \\ &= -x^3 + 2x^2 + 5x - 6 \end{aligned}$$

$\therefore$ The characteristic equation of A is

$$-x^3 + 2x^2 + 5x - 6 = 0$$

(i.e) $x^3 - 2x^2 - 5x + 6 = 0$

Theorem 1: (Cayley Hamilton Theorem).

Any square matrix A satisfies its characteristic equation.

(i.e.) if $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ is the characteristic polynomial of degree n of A then $a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = \mathbf{0}$.

Proof:

Let A be a square matrix of order n.

$$\text{Let } |A - xI| = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \ldots\ldots\ldots\ldots (1)$$

be the characteristic polynomial of A.

Now, $\operatorname{adj}(A-xI)$ is a matrix polynomial of degree $n-1$ since each entry of the matrix $\operatorname{adj}(A-xI)$ is a

cofactor of $A-xI$ and hence is a polynomial of degre $\leq n-1$.

$\therefore$ Let $\operatorname{adj}(A-xI) = B_0 + B_1x + B_2x^2 + \cdots + B_{n-1}x^{n-1}$

Now, $(A-xI)adj(A-xI) = |A-xI|I$.

$(\because (\operatorname{adj}A)A = A(\operatorname{adj}A) = |A|$

$\therefore (A-xI)(B_0 + B_1x + \cdots + B_{n-1}x^{n-1})$
$= (a_0 + a_1x + \cdots + a_nx^n)I$ using (1) and (2)

$\therefore$ Equating the coefficients of the corresponding powers of x we get

$$\begin{aligned} AB_0 &= a_0I \\ AB_1 - B_0 &= a_1I \\ AB_2 - B_1 &= a_2I \\ &\dots \cdots \\ &\dots \dots \\ AB_{n-1} - B_{n-2} &= a_{n-1}I \\ -B_{n-1} &= a_nI \end{aligned}$$

Pre-multiplying the above equations by $I, A, A^2, \dots A^n$ respectively and adding we get

$a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = \mathbf{0}$.

Note:

The inverse of a non-singular matrix can be calculated by using the Cayley Hamilton theorem a follow.

Let $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ be the characteristi polynomial of A.

Then by theorem 1.1 we have

$a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = 0$.

Since $|A-xI| = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n w$ get $a_0 = |A|$

(by putting $x = 0$).

$\therefore \quad a_0 \neq 0 (\because A$ is a non singular matrix.)

$$\therefore \quad I = -\frac{1}{a_0}[a_1 A + a_2 A^2 + \cdots + a_n A^n] \qquad \text{(by3)}$$

$$A^{-1} = -\frac{1}{a_0}[a_1 I + a_2 A + \cdots + a_n A^{n-1}]$$

Solved Problems

Problem 1:

Find the characteristic equation of the matrix

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

Solution:

The characteristic equation of A is given by $|A - \lambda I| = 0$.

(i.e.) $\begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} = 0.$

$(8-\lambda)[(7-\lambda)(3-\lambda) - 16] + 6[-6(3-\lambda) + 8]$

$$+2[24 - 2(7-\lambda)] = 0$$

(i.e.) $(8-\lambda)(\lambda^2 - 10\lambda + 5) + 6(6\lambda - 10)$

$$+2(2\lambda + 10) = 0$$

(i.e.) $(8\lambda^2 - 80\lambda + 40 - \lambda^3 + 10\lambda^2 - 5\lambda)$

$$+(36\lambda - 60) + (4\lambda + 20) = 0$$

(i.e) $\lambda^3 - 18\lambda^2 + 45\lambda = 0$, which represents the characteristic equation of A.

Problem 2. Show that the non-singular matrix

$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ satisfies the equation $A^2 - 2A - 5I = 0$. Hence evaluate A^{-1}.

Solution:

The characteristic polynomial of A is

$$|A - xI| = \begin{vmatrix} 1-x & 2 \\ 3 & 1-x \end{vmatrix} = x^2 - 2x - 5.$$

$\therefore$ By Cayley-Hamilton theorem $A^2 - 2A - 5I = 0$.

$$\therefore\ I = \frac{1}{5}(A^2 - 2A).$$

$$\therefore\ A^{-1} = \frac{1}{5}(A - 2I)$$

$$= \frac{1}{5}\left[\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

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

Problem 3:

Show that the matrix

$A = \begin{pmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{pmatrix}$ satisfies the equation $A(A - I)(A + 2I) = 0$.

Solution:

The characteristic polynomial of A is

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

$= -\lambda^3 - \lambda^2 + 2\lambda$ (verify).

$\therefore$ By Cayley-Hamilton theorem $-A^3 - A^2 + 2A = 0$.

(i.e.) $A^3 + A^2 - 2A = 0$. Hence $A(A^2 + A - 2I) = 0$.

$\therefore\ A(A + 2I)(A - I) = 0$.

Problem 4:

Using Cayley-Hamilton theorem find the inverse of the matrix

$$\begin{pmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{pmatrix}$$

Solution:

Let $A = \begin{pmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{pmatrix}$

The characteristic polynomial of $A = |A - xI|$

$$= \begin{vmatrix} 7-x & 2 & -2 \\ -6 & -1-x & 2 \\ 6 & 2 & -1-x \end{vmatrix}$$

$$= (7-x)[(1+x)^2 - 4] - 2[6(1+x) - 12]$$
$$-2[-12 + 6(1+x)]$$

$$=(3^2 - x)(x^2 + 2x - 3) - 12(x-1) - 12(x-1)$$
$$= -7x^2 + 14x - 21 - x^3 - 2x^2 + 3x - 12x + 12$$
$$-12x - 12$$
$$= -x^3 - 5x^2 - 7x + 3$$

By Cayley-Hamilton theorem

$$-A^3 + 5A^2 - 7A + 3I_3 = 0$$
$$\therefore 4^3 - 54^2 + 74 - 3I_3 = 0$$
$$\therefore 3I_3 = 4^3 - 54^2 + 74$$
$$\therefore 3 = \frac{1}{5}(\mathrm{A}^3 - 5\mathrm{A}^2 + 7\mathrm{A})$$

Pre(or) post multiplying by A^{-1} on boch sides we ge

$$\mathrm{A}^{-1} = \frac{1}{5}[4^2 - 54 \div 7I_3]$$

$$\text{Now. } \mathrm{A}^2 = \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix} \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$$

$$= \begin{pmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{pmatrix} \text{ (verify).}$$

From (1)

Problem 5:

Fid the inverse of the matrix $\begin{pmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{pmatrix}$ mose Crim-Hamitton theorem.

Solution. The characteristic polynomial of A

$$= |A - x|\ | = \begin{vmatrix} 3-x & 3 & 4 \\ 2 & -3-x & 4 \\ 0 & -1 & 1-x \end{vmatrix}$$
$$= -x^3 + x^2 + 11x - 11 \text{ (verify).}$$

∴ By Cayley-Hamilton theorem

$$-A^3 + A^2 + 11A - 11I_3 = 0$$

Hence 11 $I_3 = -(A^3 - A^2 - 11A)$.

∴ $I_3 = -\frac{1}{11}[A^3 - A^2 - 11A]$.

Pre (post) multiplying by A^{-1} on both sides we get

$$A^{-1} = -\frac{1}{11}[A^2 - A - 11I_3]$$
$$= -\frac{1}{11}\left[\begin{pmatrix} 15 & -4 & 28 \\ 0 & 11 & 0 \\ -2 & 2 & -3 \end{pmatrix} - \begin{pmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{pmatrix}\right.$$
$$\left. -11\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] = \begin{pmatrix} -\frac{1}{11} & \frac{7}{11} & -\frac{24}{11} \\ \frac{2}{11} & -\frac{3}{11} & \frac{4}{11} \\ \frac{2}{11} & -\frac{3}{11} & \frac{15}{11} \end{pmatrix}$$

Problem 6:

Verify Cayley Hamilton's theorem for the matrix$A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$.

Solution:

The characteristic equation of A is $|A - \lambda|\ | = 0$.

$$\therefore \begin{vmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{vmatrix} = 0$$
$$\therefore (1-\lambda)(3-\lambda) - 8 = 0$$
$$\therefore \lambda^2 - 4\lambda - 5 = 0.$$

By Cayley Hamilton's theorem A satisfies its characteristic equation.

$\therefore$ We have $A^2 - 4A - 5I = 0$.

Now, $A^2 = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$

$= \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix}$

$4A = \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix}$ and $5I = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$

$\therefore A^2 - 4A - 5I = \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix} - \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix}$

$- \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$

Thus, Cayley Hamilton's theorem is verified.

Problem 7:

Using Cayley Hamilton's theorem for the matrix $A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}$

find (i) A^{-1} (ii) A^4.

Solution. The characteristic equation of A is $|A - \lambda I| = 0$.

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

(i.e) $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$. (Verify)

$\therefore$ By Cayley Hamilton's theorem

$$A^3 - 5A^2 + 8A - 4I = 0 \ldots\ldots (1)$$
$$4I = A^3 - 5A^2 + 8A$$

(i) To find A^{-1} pre multiplying by A^{-1} we get

$$4A^{-1} = A^{-1}A^3 - 5A^{-1}A^2 + 8A^{-1}A$$
$$= A^2 - 5A + 8I (2)$$

Now, $A^2 = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}$

$= \begin{pmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{pmatrix}$

From (2)

$$A^{-1} = \frac{1}{4}\left[\begin{pmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{pmatrix} - \begin{pmatrix} 5 & 0 & -10 \\ 10 & 10 & 20 \\ 0 & 0 & 10 \end{pmatrix} + \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix}\right]$$

$$= \frac{1}{4}\begin{pmatrix} 4 & 0 & 4 \\ 4 & 2 & -8 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\therefore A^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

(ii) To find A^4.

From (1) $A^3 = 5A^2 - 8A + 4I$

$\therefore A^4 = 5A^3 - 8A^2 + 4A$
$= 5[5A^2 - 8A + 4I] - 8A^2 + 4A$

(using (1))

$= 17A^2 - 36A + 20I$

$$= 17\begin{pmatrix}1 & 0 & -6\\ 6 & 4 & 12\\ 0 & 0 & 4\end{pmatrix} - 36\begin{pmatrix}1 & 0 & -2\\ 2 & 2 & 4\\ 0 & 0 & 2\end{pmatrix} + 20\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$= \begin{pmatrix}17 & 0 & -102\\ 102 & 68 & 204\\ 0 & 0 & 68\end{pmatrix} - \begin{pmatrix}36 & 0 & -72\\ 72 & 72 & 144\\ 0 & 0 & 72\end{pmatrix} + \begin{pmatrix}20 & 0 & 0\\ 0 & 20 & 0\\ 0 & 0 & 20\end{pmatrix}$$

$$\therefore A^4 = \begin{pmatrix}1 & 0 & -30\\ 30 & 16 & 60\\ 0 & 0 & 16\end{pmatrix}$$

Exercises:

1.Obtain the characteristic polynomial for the following matrices.

(i) $\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}$ (ii) $\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$

2.Find the characteristic equation of the following matrices.

(i) $\begin{pmatrix}-b & -c\\ 1 & 0\end{pmatrix}$ (ii) $\begin{bmatrix}1 & 0 & 3\\ 2 & 1 & -1\\ 1 & -1 & 1\end{bmatrix}$ (iii) $\begin{pmatrix}8 & -6 & 2\\ -6 & 7 & -8\\ 2 & -4 & 3\end{pmatrix}$

(iv) $\begin{bmatrix}-b & -c & -d\\ 1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}$

3. Verify Cayley-Hamilton theorem for the matrix $A = \begin{pmatrix}1 & 4\\ 2 & 3\end{pmatrix}$ and hence find A^{-1}.

4. If $A = \begin{pmatrix}1 & 0 & 2\\ 0 & 1 & 2\\ 1 & 2 & 0\end{pmatrix}$ prove that $A^3 - 2A^2 - 5A + 6I = 0$.

5. Verify Cayley-Hamilton theorem for A and hence find A^{-1}.

(a) $A=\begin{pmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix}$ (b) $A=\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 3 \end{pmatrix}$

(c) $A=\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$ (d) $A=\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$

6. If $A=\begin{pmatrix} 2 & 4 \\ 1 & 1 \end{pmatrix}$ find A^3 and A^{-3}.

7. Verify that the matrix $A=\begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & -1 \end{pmatrix}$ satisfies its own characteristic equation and hence find A^{-1} and A^4.

UNIT V

Eigen values and Eigen vectors–Properties and problems– Bilinear forms – Quadratic forms – Reduction of quadratic form to diagonal form.

Chapter 5: Sections- 5.1-5.3

5.1. Eigen Values and Eigen Vectors

Definition:

Let A be an $n \times n$ matrix. A number λ. is called an eigen value of A if there exists a nonzero vector $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ such that $AX = \lambda X$ and X is called an eigen vector corresponding to the eigen value λ.

Remark 1:

If X is an eigen vector corresponding to the eigen value λ of A, then αX where α is any non-zero number, is also an eigen vector corresponding to λ..

Remark 2:

Let X be an eigen vector corresponding to the eigen value λ of A. Then $AX = \lambda X$ so that $(A - \lambda I)X = 0$. Thus X is a non-trivial solution of the system of homogeneous linear equations $(A - \lambda I)X = 0$. Hence $|A - \lambda I| = 0$, which is the characteristic polynomial of A. Let $|A - \lambda I| = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_n$.
The roots of this polynomial give the eigen values of A. Hence eigen values are also called characteristic roots.

Properties of Eigen Values:

Property 1: Let X be an eigen vector corresponding, the eigen values λ_1 and λ_2. Then $\lambda_1 = \lambda_2$.

Proof:

By definition $X \neq 0, AX = \lambda_1 X$ and $X = \lambda_2 X$

$\therefore \lambda_1 X = \lambda_2 X$

$\therefore (\lambda_1 - \lambda_2)X = 0$

Since $X \neq 0, \lambda_1 = \lambda_2$.

Property 2:

Let A be a square matrix.

Then (i) the sum of the eigen values of A is equal othe sum of the diagonal elements (trace) of A. (ii) Product of eigen values of A is $|A|$.

Proof:

(i) Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$

The eigen values of A are the roots of the characteristic equation

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Let $|A - \lambda I| = a_0 \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n$ ....

From (1) and (2) we get

$$a_0 = (-1)^n; a_1 = (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn}); \cdots (3)$$

Also, by putting $\lambda = 0$ in (2) we get $a_n = |A|$

Now let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigen values of A.

$\therefore \lambda_1, \lambda_2, \ldots, \lambda_n$ are the roots of (2).

$$\therefore \lambda_1 + \lambda_2 + \cdots + \lambda_n = -\frac{a_1}{a_0}$$

$$= a_{11} + a_{22} + \cdots + a_{nn} (\text{ using (3)})$$

$\therefore$ Sum of the eigen values $=$ trace of A.

(ii) Product of the eigen values = product of the roots

$$= \lambda_1 \lambda_2 \cdots \lambda_n$$
$$= (-1)^n \frac{a_n}{a_0} = \frac{(-1)^n a_n}{(-1)^n}$$
$$= a_n$$
$$= |A|.$$

Property 3:

The eigen values of A and its transpose A^T are the same.

Proof:

It is enough if we prove that A and A^T have the same characteristic polynomial.

Since for any square matrix $M, |M| = |M^T|$ we have,

$$|A - \lambda I| = |(A - \lambda I)^T| = |A^T - (\lambda I)^T|$$
$$= |A^T - \lambda I|$$

Hence the result.

Property 4: If λ is an eigen value of a non-singular matrix A then $\frac{1}{\lambda}$ is an eigen value of A^{-1}.

Proof:

Let X be an eigen vector corresponding to λ.

Then $AX = \lambda X$. Since A is non singular A^{-1} exists.

$$\therefore\ A^{-1}(AX) = A^{-1}(\lambda X)$$
$$IX = \lambda A^{-1} X$$
$$\therefore\ A^{-1}X = \left(\frac{1}{\lambda}\right) X.$$

$\therefore\ \frac{1}{\lambda}$ is an eigen value of A^{-1}.

Corollary: If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigen values of a non-singular matrix A then $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$ are the eigen values of A^{-1}.

Property 5.

If λ is an eigen value of A then $k\lambda$ is an eigen value of kA where k is a scalar.

Proof:

Let X be an eigen vector corresponding to λ. Then $AX = \lambda X$. ……. (1)

Now,

$$\begin{aligned}(kA)X &= k(AX)\\ &= k(\lambda X) \text{ (by (1))}\\ &= (k\lambda)X.\end{aligned}$$

∴ $k\lambda$ is an eigen value of kA.

Property 6:

If λ is an eigen value of A then λ^k is an eigen value of A^k where k is any positive integer.

Proof:

Let X be an eigen vector corresponding to λ.

Then $AX = \lambda X$ ……… (1)

Now,

$$\begin{aligned}A^2X &= (AA)X = A(AX)\\ &= A(\lambda X) \text{ (by (1))}\\ &= \lambda(AX)\\ &= \lambda(\lambda X) \text{ (by (1))}\\ &= \lambda^2 X.\end{aligned}$$

∴ λ^2 is an eigen value of A^2.

Proceeding like this we can prove that λ^k is an eigen value of A^k for any positive integer.

Corollary. If $\lambda_1, \lambda_2, \dots, \lambda_n$ are eigen values of A then $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ are eigen values of A^k for any positive integer k.

Property 7.

Eigen vectors corresponding to distinct eigen values of a matrix are linearly independent.

Proof:

Let $\lambda_1, \lambda_2, \cdots, \lambda_k$ be distinct eigen values of a matrix and let X_i be the eigen vector corresponding to λ_i.

Hence $AX_i = \lambda_i X_i \; (i = 1,2, \dots, k)$

Now, suppose $X_1, X_2, \dots, X_k$ are linearly dependent. Then there exist real numbers $\alpha_1, \alpha_2, \dots, \alpha_k$, not all zero, such that $\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_k X_k = 0$. Among all such relations, we choose one of shortest length, say j.

By rearranging the vectors $X_1, X_2, \dots, X_k$ we may assume that

$\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_j X_j = 0.$

$$\therefore A(\alpha_1 X_1) + A(\alpha_2 X_2) + \cdots + A(\alpha_j X_j) = 0 \dots\dots\dots (2)$$

$$\therefore \alpha_1 (AX_1) + \alpha_2 (AX_2) + \cdots + \alpha_j (AX_j) = 0$$

$$\therefore \alpha_1 \lambda_1 X_1 + \alpha_2 \lambda_2 X_2 + \cdots + \alpha_j \lambda_j X_j = 0 \dots\dots\dots (3)$$

Multiplying (2) by λ_1 and subtracting from (3), we get

$$\alpha_2(\lambda_1 - \lambda_2)X_2 + \alpha_3(\lambda_1 - \lambda_3)X_3 + \cdots + \alpha_j(\lambda_1 - \lambda_j)X_j = 0 \dots\dots\dots (4)$$

and since $\lambda_1, \lambda_2, \dots, \lambda_j$ are distinct and $\alpha_2, \dots, \alpha_j$ are non-zero we have

$\alpha_i(\lambda_1 - \lambda_i) \neq 0;\; i = 2,3, \dots, j.$

Thus (4) gives a relation whose length is $j - 1$, giving a contradiction.

Hence $X_1, X_2, \dots, X_k$ are linearly independent.

Property 8:

The characteristic roots of a Hermitian matrix are all real.

Proof:

Let A be a Hermitian matrix. Hence

$$A = \bar{A}^T \quad \ldots \ldots\ldots(1)$$

Let λ be a characteristic root of A and let X be a characteristic vector corresponding to λ.

$$\therefore \; AX = \lambda X \ldots \ldots .(2)$$

Now,

$$\begin{aligned} AX = \lambda X &\Rightarrow \bar{X}^T AX = \lambda \bar{X}^T X \\ &\Rightarrow (\bar{X}^T AX)^T = \lambda \bar{X}^T X \;\; (\text{since } X^T AX \text{ is} \\ &\Rightarrow X^T A^T (\bar{X}^T)^T = \lambda \bar{X}^T X \\ &\Rightarrow X^T A^T \bar{X} = \lambda \bar{X}^T X \\ &\Rightarrow \bar{X}^T A^T \cdot \bar{X} = \overline{\lambda X^T X} \\ &\Rightarrow \bar{X}^T \bar{A}^T X = \bar{\lambda} X^T \bar{X} \\ &\Rightarrow \bar{X}^T AX = \bar{\lambda} X^T \bar{X} \;\; (\text{using } 1) \\ &\Rightarrow \bar{X}^T \lambda X = \bar{\lambda} X^T \bar{X} \;\; (\text{using } 2) \\ &\Rightarrow \lambda(\bar{X}^T X) = \bar{\lambda}(X^T \bar{X}) \ldots \end{aligned}$$

Now,

$$\begin{aligned} \bar{X}^T X = X^T \bar{X} &= \overline{x_1} x_1 + \overline{x_2} x_2 + \cdots \ldots + \overline{x_n} x_n \\ &= |x_1|^2 + |x_2|^2 + \cdots \ldots + |x_n|^2 \\ &\neq 0 \end{aligned}$$

$\therefore$ From (3) we get $\lambda = \bar{\lambda}$

Hence λ is real.

Corollary.

The characteristic roots of a real symmetric matrix are real.

Proof:

We know that any real symmetric matrix is Hermitian. Hence the result follows from the above property.

Property 9:

The characteristic roots of a skew Hermitian matrix are either purely imaginary or zero.

Proof:

Let A be a skew Hermitian matrix and λ be a characteristic root of A.

$\therefore |A - \lambda I| = 0$

$\therefore |iA - i\lambda I| = 0$

$\therefore$ $i\lambda$ is a characteristic root of iA.

Since A is skew Hermitian iA is Hermitian, $i\lambda$ is real. Hence λ is purely imaginary or zero.

Corollary:

The characteristic roots of a real skew symmetric matrix are either purely imaginary or zero.

Proof:

We know that any real skew symmetric matrix is skew Hermitian.

Hence the result follows from the above property.

Property 10:

Let λ be a characteristic root of a unitary matrix A. Then $|\lambda| = 1$. (i.e) the characteristic roots of a unitary matrix are all the unit modulus.

Proof:

Let λ be a characteristic root of a unitary matrix A and X be a characteristic vector corresponding to λ.

$$\therefore AX = \lambda X \ldots\ldots (1)$$

Taking conjugate and transpose in (1) we get

$(\overline{AX})^T = (\overline{\lambda X})^T$

$$\therefore \bar{X}^T \bar{A}^T = \bar{\lambda} \bar{X}^T \text{ ... ...(2)}$$

Multiplying (1) and (2) we get

$$\left(\overline{X^T A^T}\right)(AX) = \left(\bar{\lambda}\bar{X}^T\right)(\lambda X)$$

$$\therefore \bar{X}^T(\bar{A}^T A)X = \bar{\lambda}\lambda(\bar{X}^T X)$$

Now, since A is an unitary matrix $\bar{A}^T A = I$.

Hence $\bar{X}^T X = (\bar{\lambda}\lambda)\bar{X}^T X$

Since X is non-zero vector $\bar{X}^T$ is also non-zero vector and $\bar{X}^r X = |x_1|^2 + |x_2|^2 + \cdots .. + |x_n|^2 \neq 0$ we get $\lambda\bar{\lambda} = 1$.Hence $|\lambda|^2 = 1$. Hence $|\lambda| = 1$.

Corollary:

Let λ be a characteristic root of an orthogonal matrix A. Then $|\lambda| = 1$.

Since any orthogonal matrix- is unitary the result follows from property 10.

Property 11:

Zero is an eigen value of A if and only if A is a singular matrix.

Proof:

The eigen values of A are the roots of the characteristic equation $|A - \lambda I| = 0$.

Now, 0 is an eigen value of $A \Leftrightarrow |A - 0I| = 0$

$$\Leftrightarrow |A| = 0$$

$$\Leftrightarrow A \text{ is a singular matrix}$$

Property 12.

If A and B are two square matrices of the same order then AB and BA have the same eigen values.

Solution.

Let λ be an eigen value of AB and X be an eigen vector corresponding to λ.

$\therefore \quad (AB)X = \lambda X$

$\therefore \quad B(AB)X = B(\lambda X) = \lambda(BX).$

$\therefore \quad (BA)(BX) = \lambda(BX)$

$\therefore \quad (B\ddot{A})Y = \lambda Y$, where $Y = BX$.

Hence λ is an eigen value of BA.

Also BX is the corresponding eigen vector.

Property 13:

If P and A are $n \times n$ matrices and P is a nonsingular matrix then A and $P^{-1}AP$ have the same eigen values.

Proof:

Let $B = P^{-1}AP$.

To prove A and B have same eigen values, it is enough to prove that the characteristic polynomials of A and B are the same.

Now

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda I| \\ &= |P^{-1}AP - P^{-1}(\lambda I)P| \\ &= |P^{-1}(A - \lambda I)P| \\ &= |P^{-1}||A - \lambda I||P| \\ &= |P^{-1}||P||A - \lambda I| \\ &= |P^{-1}P||A - \lambda I| \\ &= |I||A - \lambda I| \\ &= |A - I\lambda| \end{aligned}$$

$\therefore$ The characteristic equations of A and $P^{-1}AP$ are the same.

Property 14:

If λ is a characteristic root of A then $f(\lambda)$ is a characteristic root of the matrix $f(A)$ where $f(x)$ is any polynomial.

Proof:

Let $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ where $a_0 \neq 0$ and $a_1, a_2, \ldots, a_n$ are all real numbers.

$\therefore\ f(A) = a_0A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI.$

Since λ is a characteristic root of A, λ'' is a characteristic root of A^n for any positive integer n (refer Property 6)

$$\therefore A^nX = \lambda^nX$$
$$A^{n-1}X = \lambda^{n-1}X$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$AX = \lambda X$$
$$\therefore a_0A^nX = a_0\lambda^nX$$
$$a_1A^{n-1}X = a_1\lambda^{n-1}X$$

___

___

$$a_{n-1}AX = a_{n-1}\lambda X$$

Adding the above equations we have

$$a_0A^nX + a_1A^{n-1}X + \cdots + a_{n-1}AX$$
$$= a_0\lambda^nX + a_1\lambda^{n-1}X + \cdots + a_{n-1}\lambda X$$
$$\therefore (a_0A^n + a_1A^{n-1} + \cdots + a_{n-1}A)X$$
$$= (a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda)X$$
$$\therefore (a_0A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI)X$$
$$= (a_0\lambda^n + a_1\lambda^{n-1} + \cdots$$
$$+ a_{n-1}\lambda + a_n)X$$
$$\therefore f(A)X = f(\lambda)X$$

Hence $f(\lambda)$ is a characteristic root of $f(A)$.

Solved Problems

Problem 1:

If X_1, X_2 are eigen vectors corresponding to an eigen value λ then $aX_1 + bX_2 (a, b$ non-zero scalars) is also an eigen vector corresponding to λ.

Solution:

Since X_1 and X_2 are given vectors corresponding to λ, we have

$AX_1 = \lambda X_1$ and $AX_2 = \lambda X_2$.

Hence $A(aX_1) = \lambda(aX_1)$ and $A(bX_2) = \lambda(bX_2)$.

$\therefore\ A(aX_1 + bX_2) = \lambda(aX_1 + bX_2)$.

$\therefore\ aX_1 + bX_2$ is an eigen vector corresponding to λ.

Problem 2:

If the eigen values of

$A = \begin{pmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{pmatrix}$ are 2,2,3 find the eigen values of A^{-1} and A^2.

Solution:

Since 0 is not an eigen value of A, A is a non-singular matrix and hence A^{-1} exists.

Eigen values of A^{-1} are $\frac{1}{2}, \frac{1}{2}, \frac{1}{3}$ and eigen values of A^2 are $2^2, 2^2, 3^2$.

Problem 3:

Find the eigen values of A^5 when

$A = \begin{pmatrix} 5 & 4 & 0 \\ 3 & 6 & 1 \end{pmatrix}$

golution. The characteristic equation of A is obviously $(3-\lambda)(4-\lambda)(1-\lambda) = 0$.

Hence the eigen values of A are 3,4,1.

$\therefore$ The eigen values of A^5 are $3^5, 4^5, 1^5$.

Problem 4:

Find the sum and product of the eigen values of the matrix $\begin{pmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{pmatrix}$ without actually finding the eigen values.

Solution:

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

Sum of the eigen values $=$ trace of $A = 3 + (-2) + 3 = 4.$

Product of the eigen values $= |A|.$

Now,

$$|A| = \begin{vmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{vmatrix}$$
$$= 3(-6+4) + 4(3-4) - 4(-1+2)$$
$$= -6 - 4 - 4 = -14$$

$\therefore$ Product of the eigen values $= -14.$

Problem 5:

Find the characteristic roots of the matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

Solution:

Let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

The characteristic equation of A is given by $|A - \lambda I| = 0.$

$\therefore \begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ -\sin\theta & \cos\theta - \lambda \end{vmatrix} = 0.$

$\therefore \ (\cos\theta - \lambda)^2 - \sin^2\theta = 0.$

$\therefore \ (\cos\theta - \lambda - \sin\theta)(\cos\theta - \lambda + \sin\theta) = 0.$

$\therefore \ [\lambda - (\cos\theta - \sin\theta)][\lambda - (\cos\theta + \sin\theta)] = 0.$

$\therefore$ The two characteristic roots, (the two eigen values) of the matrix are $(\cos\theta - \sin\theta)$ and $(\cos\theta + \sin\theta)$.

Problem 6:

Find the characteristic roots of the matrix $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$

Solution:

The characteristic equation of A is given by $|A - \lambda I| = 0$.

(i.e.) $\begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ -\sin\theta & -\cos\theta - \lambda \end{vmatrix} = 0$

$\therefore\ -(\cos^2\theta - \lambda^2) - \sin^2\theta = 0.$

$\therefore\ \lambda^2 - (\cos^2\theta + \sin^2\theta) = 0.$

$\therefore\ \lambda^2 - 1 = 0.$

$\therefore$ The characteristic roots are 1 and -1.

Problem 7:

Find the sum and product of the eigen values of the matrix

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ without finding the roots of the characteristic equation.

Solution:

Sum of the eigen values of $A =$ trace of $A = a_{11} + a_{22}$.

Product of the eigen values of

$A = |A| = a_{11}a_{22} - a_{12}a_{21}$.

Problem 8:

Verify the statement that the sum of the elements in the diagonal of a matrix is the sum of the eigen values of the matrix $A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}$

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

(i.e.) $\begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} = 0.$

(i.e.) $(-2-\lambda)[(1-\lambda)(-\lambda) - 12] - 2[-2\lambda - 6]$

$$-3[-4 + (1-\lambda)] = 0.$$

(i.e.) $(-2-\lambda)(\lambda^2 - \lambda - 12) + 4(\lambda + 3)$

$$+3(\lambda + 3) = 0.$$

(i.e) $-2\lambda^2 + 2\lambda + 24 - \lambda^3 + \lambda^2$

$$+12\lambda + 4\lambda + 12 + 3\lambda + 9 = 0.$$

(i.e) $-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$

(i.e) $\lambda^3 + \lambda^2 - 21\lambda - 45 = 0.$

This is a cubic equation in λ and hence it has 3 roots and the three roots are the three eigen values of the matrix.

$$\text{The sum of the eigen values } = -\left(\frac{\text{coefficient of } \lambda^2}{\text{coefficient of } \lambda^3}\right)$$
$$= -1.$$

The sum of the elements on the diagonal of the matrix

$A = -2 + 1 + 0 = -1$.Hence the result.

Problem 9:

The product of two eigen values of the matrix $A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$ is 16. Find the third eigen value. What is the sum of the eigen values of A ?

Solution:

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A.

Given, product of 2 eigen values (say) λ_1, λ_2 is 16. $\therefore\ \lambda_1\lambda_2 = 16$

We know that the product of the eigen values is $|A|$.

(i.e) $\lambda_1\lambda_2\lambda_3 = \begin{vmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{vmatrix}$

(i.e)

$$\begin{aligned} 16\lambda_3 &= 6(9-1) + 2(-6+2) + 2(2-6) \\ &= 48 - 8 - 8 \\ &= 32 \end{aligned}$$
$\therefore\ \lambda_3 = 2$

∴ The third eigen value is 2.

Also we know that the sum of the eigen values of

$A =$ trace of $A = 6 + 3 + 3 = 12$

Problem 10:

The product of two eigen values of the matrix $A = \begin{pmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{pmatrix}$ is -12. Find the cigen values of A.

Solution:

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A, Given product of 2 eigen values, say. λ_1 and λ_2 is -12.

$$\therefore \lambda_1 \lambda_2 = -12 \ldots\ldots\ldots.(1)$$

We know that the product of the eigen values is $|A|$.

$$\begin{aligned} \therefore \lambda_1\lambda_2\lambda_3 &= \begin{vmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{vmatrix} \ldots\ldots\ldots(2) \\ \text{i.e } 12\lambda_3 &= -12 \\ \therefore \lambda_3 &= 1 \end{aligned}$$

Also, we know sum of the eigen values $=$ Trace of A.

$$\therefore \qquad \lambda_1 + \lambda_2 + \lambda_3 = 2 + 1 - 3 = 0$$
$$\therefore \lambda_1 + \lambda_2 = -1 \text{ (using (2))(3)}$$

Using (3) in (1) we get $\lambda_1(-1 - \lambda_1) = -12$

$$\lambda_1^2 + \lambda_1 - 12 = 0$$
$$(\lambda_1 + 4)(\lambda_1 - 3) = 0$$
$$\therefore \lambda_1 = 3 \text{ or } -4$$

Putting $\lambda_1 = 3$ in (1) we get $\lambda_2 = -4$. Or putting $\lambda_1 = -4$ in (1) we get $\lambda_2 = 3$.

Thus, the three eigen values are $3, -4, 1$.

Problem 11:

Find the sum of the squares of the eigen values of $A = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$

Solution:

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A.

We know that $\lambda_1^2, \lambda_2^2, \lambda_3^2$ are the eigen values of A^2.

$$A^2 = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$$

$$= \begin{pmatrix} 9 & 5 & 38 \\ 0 & 4 & 42 \\ 0 & 0 & 25 \end{pmatrix}$$

Sum of the eigen values of $A^2 =$ Trace of A^2

$= 9 + 4 + 25$

(i.e) $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 38$

Sum of the squares of the eigen values of $A = 38$.

Problem 12:

Find the eigen values and eigen vectors of the matrix

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

Solution:

The characteristic equation of A is

$|A - \lambda I| = 0.$

$$\therefore \begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix} = 0.$$

$\therefore (1-\lambda)[(5-\lambda)(1-\lambda)-1]-[(1-\lambda)-3].+3[1-3(5-\lambda)] = 0.$

$(1-\lambda)(\lambda^2-6\lambda+4)+(\lambda+2)+3(3\lambda-14) = 0.$

$\lambda^2-6\lambda+4-\lambda^3+6\lambda^2-4\lambda+\lambda+2+9\lambda-42 = 0.$

$\therefore -\lambda^3+7\lambda^2-36 = 0.$ Hence $\lambda^3-7\lambda^2+36 = 0.$

$\therefore (\lambda+2)(\lambda^2-9\lambda+18) = 0.$

Hence $(\lambda+2)(\lambda-6)(\lambda-3) = 0.$

$\therefore \lambda = -2,3,6$ are the three eigen values.

Case (i) Eigen vector corresponding to $\lambda = -2$.

Let $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ be an eigen vector corresponding to $\lambda = -2$.

Hence $AX = -2X$.

(i.e.) $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_1 \\ -2x_2 \\ -2x_3 \end{bmatrix}$

$x_1+x_2+3x_3 = -2x_1$

$x_1+5x_2+x_3 = -2x_2$

$3x_1+x_2+x_3 = -2x_3$

$3x_1+x_2+3x_3 = 0 \ldots\ldots\ldots..(1)$

$x_1+7x_2+x_3 = 0$ .... (2)

$3x_1+x_2+3x_3 = 0 \ldots\ldots\ldots.(3)$

Clearly this system of three equations reduces to two equations only. From (1) and (2) we get

$\because\ x_1 = -2k; x_2 = 0; x_3 = 2k.$

∴ It has only one independent solution and can be obtained by giving any value to k say $k = 1$.

∴ $(-2,0,2)$ is an eigen vector corresponding to $\lambda = -2$.

Case (ii) Eigen vector corresponding to $\lambda = 3$.

Then $AX = 3X$ gives

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

Taking the first 2 equations we get

$$\frac{x_1}{-5} = \frac{x_2}{5} = \frac{x_3}{-5} = k(\text{ say }).$$

$\therefore x_1 = -k; x_2 = k; x_3 = -k.$

Taking $k = 1$ (say) $(-1,1,-1)$ is an eigen vector corresponding to $\lambda = 3$.

Case (iii) Eigen vector corresponding to $\lambda = 6$.

We have $AX = 6X$.

Hence

$$\begin{aligned} -5x_1 + x_2 + 3x_3 &= 0 \\ x_1 - x_2 + x_3 &= 0 \\ 3x_1 + x_2 - 5x_3 &= 0 \end{aligned}$$

Taking the first two equations we get

$$\frac{x_1}{4} = \frac{x_2}{8} = \frac{x_3}{4} = k$$

∴ $x_1 = k; x_2 = 2k; x_3 = k$. It satisfies the third equation also.

Taking $k = 1$ (say) $(1,2,1)$ is an eigen vector corresponding to $\lambda = 6$.

Problem 13:

Find the eigen values and eigen vectors of the matrix

$$A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$$

Solution:

The characteristic equation of A is $|A-\lambda I|=0$.

$$\therefore \begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0.$$

$$\therefore\ (6-\lambda)[(3-\lambda)^2-1]+2[(2\lambda-6)+2]$$
$$+2(2-6+2\lambda)=0.$$

$$\therefore\ (6-\lambda)(8+\lambda^2-6\lambda)+4\lambda-8+4\lambda-8=0.$$

$$\therefore 48+6\lambda^2-36\lambda-8\lambda-\lambda^3+6\lambda^2+8\lambda-16=0.$$

$$\therefore\ -\lambda^3+12\lambda^2-36\lambda+32=0.$$

Hence $\lambda^3-12\lambda^2+36\lambda-32=0$.

$$\therefore\ (\lambda-2)(\lambda-2)(\lambda-8)=0.$$

$\therefore$ The eigen values are 2,2,8.

We now find the eigen vectors.

Case (i) $\lambda=2$.

The eigen vector $X=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is got from $AX=2X$.

$$\therefore\ \begin{aligned} 6x_1-2x_2+2x_3&=2x_1 \\ -2x_1+3x_2-x_3&=2x_2 \\ 2x_1-x_2+3x_3&=2x_3 \end{aligned}$$
$$\therefore\ \begin{aligned} 4x_1-2x_2+2x_3&=0 \\ -2x_1+x_2-x_3&=0 \\ 2x_1-x_2+x_3&=0 \end{aligned}$$

The above three equations are equivalent to the single equation

$2x_1-x_2+x_3=0$.

The independent eigen vectors can be obtained by giving arbitrary values to any two of the unknowns x_1,x_2,x_3.

Giving $x_1 = 1; x_2 = 2$ we get $x_3 = 0$.

Giving $x_1 = 3; x_2 = 4$ we get $x_3 = -2$.

$\therefore$ Two independent vectors corresponding to $\lambda = 2$ are $(1,2,0)$ and $(3,4,-2)$.

Case (ii) $\lambda = 8$.

The eigen vector $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is got from $AX = 8X$.

$\therefore -2x_1 - 2x_2 + 2x_3 = 0 \quad \ldots\ldots\ldots..(1)$

$-2x_1 - 5x_2 - x_3 = 0 \ldots\ldots\ldots\ldots(2)$

$2x_1 - x_2 - 5x_3 = 0 \ldots\ldots\ldots\ldots.(3)$

From (1) and (2) we get

$\frac{x_1}{12} = \frac{x_2}{-6} = \frac{x_3}{6} = k$ (say).

$\therefore \; x_1 = 2k; x_2 = -k; x_3 = k.$

Giving $k = 1$ we get an eigen vector corresponding to 8 as $(2,-1,1)$.

Problem 14:

Find the eigen values and eigen vectors of the matrix $A = \begin{pmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

(i.e.) $\begin{vmatrix} 2-\lambda & -2 & 2 \\ 1 & 1-\lambda & 1 \\ 1 & 3 & -1-\lambda \end{vmatrix} = 0.$

$\therefore \; (2-\lambda)[-(1-\lambda)(1+\lambda) - 3] + 2[-(1+\lambda) - 1] + 2[3 - (1-\lambda)] = 0$

$\therefore (2-\lambda)(\lambda^2 - 4) - 2(2+\lambda) + 2(2+\lambda) = 0$

$\therefore \; 2\lambda^2 - 8 - \lambda^3 + 4\lambda - 4 - 2\lambda + 4 + 2\lambda = 0$

$\therefore\ -\lambda^3 + 2\lambda^2 + 4\lambda - 8 = 0$

Hence $\lambda^3 - 2\lambda^2 - 4\lambda + 8 = 0$

$\therefore (\lambda - 2)(\lambda^2 - 4) = 0$

Hence $(\lambda - 2)(\lambda - 2)(\lambda + 2) = 0$

$\therefore\ \lambda = 2,2,-2$ are the three eigen values.

Case (i) $\lambda = 2$.

Let $X = (x_1, x_2, x_3)$ be an eigen vector corresponding to $\lambda = 2, X$ is got from $AX = 2X$.

(i.e.) $\begin{bmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix}$

The eigen vector corresponding to $\lambda = 2$ is given oy the equations

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

$$\text{(i.e.)} \ -x_2 + x_3 = 0 \ldots\ldots..(1)$$

$$x_1 - x_2 + x_3 = 0 \ldots\ldots\ldots\ldots(2)$$

$$x_1 + 3x_2 - 3x_3 = 0 \ldots\ldots\ldots.(3)$$

taking (1) and (2) we get $\frac{x_1}{0} = \frac{x_2}{1} = \frac{x_3}{1} = k$ (say).

$$\therefore\ x_1 = 0; x_2 = k; x_3 = k.$$

taking $k = 1$, we get $(0,1,1)$ as an eigen vector. rresponding to $\lambda = 2$.

Case (ii) $\lambda = -2$.

Corresponding to $\lambda = -2$ we have $AX = -2X$.

$$\therefore\ 2x_1 - 2x_2 + 2x_3 = -2x_1$$
$$x_1 + x_2 + x_3 = -2x_2$$
$$x_1 + 3x_2 - x_3 = -2x_3$$
$$\therefore\ 2x_1 - x_2 + x_3 = 0$$
$$x_1 + 3x_2 + x_3 = 0$$
$$x_1 + 3x_2 + x_3 = 0$$

$\therefore$ Taking the first two equations we get,

$$\frac{x_1}{-4} = \frac{x_2}{-1} = \frac{x_3}{7} = k \text{ (say)}.$$

$\therefore\ x_1 = -4k; x_2 = -k; x_3 = 7k.$

Taking $k = 1$ we get $(-4, -1, 7)$ as an eigen vector corresponding to the eigen value $\lambda = -2$.

Exercises

1. For each of the following matrices find the characteristic vectors corresponding to each characteristic root.

   (a) $\begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$

   (b) $\begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$

   (c) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

   (d) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & i & 0 & 0 \\ 2 & 1/2 & -i & 0 \\ 1/3 & -i & \pi & -1 \end{pmatrix}$

2. For what value of k is 3 a characteristic root of $\begin{pmatrix} 3 & 1 & -1 \\ 3 & 5 & -k \\ 3 & k & -1 \end{pmatrix}$

3. Find the characteristic roots and the corresponding characteristic vectors of A^3+A^2+A+I if $A=\begin{pmatrix}1 & -1 & -1\\ 1 & -1 & 0\\ 1 & 0 & -1\end{pmatrix}$
4. Prove that if λ is a characteristic root of the matrix A then $\lambda-k$ is a characteristic root of $A-kl$.
5. Show that the characteristic root of a triangular matrix are just the diagonal elements of the matrix.
6. Show that every orthogonal matrix of odd order should have 1 or -1 as a characteristic root.
7. Show that if λ is a characteristic root of a unitary matrix then so is $\frac{1}{\lambda}$.

Answers.

1. (a) 0,3,15. The corresponding characteristic vectors are $(1,2,2)$; $(2,-1,-2)$; $(2,-2,1)$.

   (b) 1,1,5; The characteristic vectors corresponding to 1 is a linear combination of $(2,-1,0)$ and $(1,0,-1)$. A characteristic vector corresponding to 5 is $(1,1,1)$.

   (c) 1,1,1. A characteristic vector is $(1,0,0)$.
2. $k=2$.

5.2. Bilinear forms:

Introduction

Consider a finite dimensional, inner product space V er the field **R** of real numbers. The inner product is function from $V\times V$ to **R** satisfying the following mditions

(i) $\langle \alpha u_1+\beta u_2, v\rangle=\alpha\langle u_1,v\rangle+\beta\langle u_2,v\rangle$

(ii) $\langle u,\alpha v_1+\beta v_2\rangle=\alpha\langle u,v_1\rangle+\beta\langle u,v_2\rangle$

In other words, the inner product is a scalar valued function of the two variables u and v and is linear function in each of the two variables. This pe of scalar valued functions are called bilinear rms. In this chapter we study bilinear forms on finite mensional vector spaces.

1. Bilinear forms

Definition.

Let V be a vector space over a field F. bilinear form on V is a function $f: V \times V \to F$ ach that

(i) $f(\alpha u_1 + \beta u_2, v) = \alpha f(u_1, v) + \beta f(u_2, v)$

(ii) $f(u, \alpha v_1 + \beta v_2) = \alpha f(u, v_1) + \beta f(u, v_2)$ where $\alpha, \beta \in F$ and $u_1, u_2, v_1, v_2 \in V$.

In other words, f is linear as a function of any one f the two variables when the other is fixed.

Examples

1. Let V be a vector space over **R**. Then an inner product on V is a bilinear form on V.
2. Let V be any vector space over a field F. Then the zero function $\hat{0}: V \times V \to F$ given by $\hat{0}(u, v) = 0$ is a bilinear form.

   For,

$$\begin{aligned} \hat{\mathbf{0}}(\alpha u_1 + \beta u_2, v) &= \mathbf{0} \\ &= \alpha \mathbf{0} + \beta \mathbf{0} \\ &= \alpha \hat{\mathbf{0}}(u_1, v) + \beta \hat{\mathbf{0}}(u_2, v) \end{aligned}$$

Similarly

$$\hat{0}(u, \alpha v_1 + \beta v_2) = \alpha 0(u, v_1) + \beta \hat{0}(u, v_2)$$

3. Suppose V is a vector space over a field F. Let f_1 and f_2 be two linear functionals on V, (ie) f_1 and f_2 are linear transformations from V to F. Then $f: V \times V \to F$ defined by $f(u, v) = f_1(u) f_2(v)$ is a bilinear form.

For, $f(\alpha u_1 + \beta u_2, v)$

$$= f_1(\alpha u_1 + \beta u_2) f_2(v)$$
$$= [\alpha f_1(u_1) + \beta f_1(u_2)] f_2(v)$$

(since f_1 is linear)

$$= \alpha f_1(u_1) f_2(v) + \beta f_1(u_2) f_2(v)$$
$$= \alpha f(u_1, v) + \beta f(u_2, v)$$

Similarly,

$$f(u, \alpha v_1 + \beta v_2) = \alpha f(u, v_1) + \beta f(u, v_2).$$

Exercises

1. Show that the function f defined by

   $f(x, y) = x_1 y_1 + x_2 y_2 + \cdots \ldots \ldots + x_n y_n$ where $x = (x_1, x_2, \ldots \ldots, x_n)$ and $y = (y_1, y_2, \ldots \ldots \ldots, y_n)$ is a bilinear form on $V_n(F)$.

2. Which of the following are bilinear forms on $V_2(\mathbf{R})$?

   Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

   (a) $f(x, y) = 1$.

   (b) $f(x, y) = (x_1 - y_1)^2 + x_2 y_2$.

   (c) $f(x, y) = (x_1 + y_1)^2 - (x_1 - y_1)^2$.

   (d) $f(x, y) = x_1 y_2 - x_2 y_1$.

Answers. (c) and (d) are bilinear forms.

Notation. Let V be a vector space over a field F. Then the set of all bilinear forms on V is denoted by $L(V, V, F)$.

Theorem 1:

Let V be a vector space over a field F. Then $L(V,V,F)$ is a vector space over F under addition and scalar multiplication defined by

$$(f+g)(u,v) = f(u,v) + g(u,v) \text{ and}$$
$$(\alpha f)(u,v) = \alpha f(u,v)$$

Proof:

Let $f, g \in L(V,V,F)$ and $\alpha_1 \in F$.

We claim that $f+g$ and $\alpha_1 f \in L(V,V,F)$.

$$\begin{aligned}(f+g)&(\alpha u_1 + \beta u_2, v)\\ &= f(\alpha u_1 + \beta u_2, v) + g(\alpha u_1 + \beta u_2, v)\\ &= \alpha f(u_1,v) + \beta f(u_2,v) + \alpha g(u_1,v) + \beta g(u_2,v)\\ &= \alpha[f(u_1,v) + g(u_1,v)] + \beta[f(u_2,v) + g(u_2,v)]\\ &= \alpha[(f+g)(u_1,v)] + \beta[(f+g)(u_2,v)]\end{aligned}$$

Similarly, we can prove that

$$(f+g)(u, \alpha v_1 + \beta v_2) = \alpha[(f+g)(u,v_1)] + \beta[(f+g)(u,v_2)]$$

Hence $(f+g) \in L(V,V,F)$.

*Also $(\alpha_1 f)(\alpha u_1 + \beta u_2, v)$

$$\begin{aligned}&= \alpha_1 f(\alpha u_1 + \beta u_2, v)\\ &= \alpha_1[\alpha f(u_1,v) + \beta f(u_2,v)]\\ &= \alpha_1 \alpha f(u_1,v) + \alpha_1 \beta f(u_2,v)\\ &= \alpha[(\alpha_1 f)(u_1,v)] + \beta[(\alpha_1 f)(u_2,v)]\end{aligned}$$

Similarly

$$(\alpha_1 f)(u, \alpha v_1 + \beta v_2) = \alpha[(\alpha_1 f(u,v_1)] + \beta[(\alpha_1 f)(u,v_2)]$$

$\therefore\ \alpha_1 f \in L(V,V,F)$.

The remaining axioms of a vector space can be easily verified.

Matrix of a bilinear form. Let f be a bilinear form on V. Fix a basis $\{v_1, v_2, \dots, v_n\}$ for V.

Let $u = \alpha_1 v_1 + \cdots \ldots + \alpha_n v_n$ and $v = \beta_1 v_1 + \cdots \ldots + \beta_n v_n$.

Then $f(u, v)$

$$= f(\alpha_1 v_1 + \cdots .. + \alpha_n v_n, \beta_1 v_1 + \cdots .. + \beta_n v_n)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \beta_j f(v_i; v_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \alpha_i \beta_j \text{ where } f(v_i, v_j) = a_{ij}$$

$$= (\alpha_1, \ldots \ldots, \alpha_n) \begin{pmatrix} a_{11} \ldots a_{1n} \\ \cdots \cdots . \\ a_{n1} \ldots a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \ldots \\ \beta_n \end{pmatrix}$$

$\therefore \quad f(u, v) = XAY^T$ where

$X = (\alpha_1, \ldots, \alpha_n), A = (a_{ij})$ and $Y = (\beta_1, \ldots, \beta_n)$

The $n \times n$ matrix A is called the matrix of the bilinear form with respect to the chosen basis.

Conversely, given any $n \times n$ matrix $A = (a_{ij})$ the $f: V \times V \to \dot{F}$ defined by $f(u, v) = XAY^T$ is a bilinear form on V and $f(v_i, v_j) = a_{ij}$. Also, if g is any other bilinear form on V such that $g(v_i, v_j) = a_{ij}$, then $f = g$ (verify).

Solved Problems

Problem 1:

Let f be the bilinear form defined on $V_2(\mathbf{R})$ by $f(x, y) = x_1 y_1 + x_2 y_2$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Find the matrix of f.

(i) W.r.t. the standard basis $\{e_1, e_2\}$.

(ii) w.r.t. the basis $\{(1,1), (1,2)\}$.

Solution.

(i) $f(e_1, e_1) = f((1,0), (1,0))$

$= 1 \times 1 + 0 \times 0 = 1$

Similarly

$$f(e_1, e_2) = 0$$
$$f(e_2, e_1) = 0$$
$$f(e_2, e_2) = 1.$$

$\therefore$ The matrix of f is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

(ii) Let $v_1 = (1,1)$ and $v_2 = (1,2)$.

Then

$$f(v_1, v_1) = 1 + 1 = 2$$
$$f(v_1, v_2) = 1 + 2 = 3$$
$$f(v_2, v_1) = 1 + 2 = 3$$
$$f(v_2, v_2) = 1 + 4 = 5.$$

$\therefore$ The matrix of f is $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$

Exercises:

1. Find the matrix of the bilinear form $f(x, y) = x_1y_2 - x_2y_1$ with respect to the standard basis in $V_2(\mathbf{R})$.
2. Find the matrix of the bilinear from f defined on $V_3(\mathbf{R})$ by $f(x, y) = x_1y_1 + x_3y_3$ where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ w.r.t.

   (a) standard basis

   (b) $\{(1,1,0), (0,1,1), (1,0,1)\}$.

   Answers.
3. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
4. (a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

Theorem 2:

Let V be a vector space of dimension n ver a field F. Fix a basis $\{v_1, v_2, \dots\dots, v_n\}$ for V. Then the function $\varphi: L(V,V,F) \to M_n(F)$ which associates with each bilinear form $f \in L(V,V,F)$ the $n \times n$ matrix (a_{ij}) where $f(v_i, v_j) = a_{ij}$ is an isomorphism.

Proof:

Clearly φ is $1 - 1$ and onto.

Now, let $f, g \in L(V,V,F)$ and $\alpha \in F$.

Let $\varphi(f) = (a_{ij})$ and $\varphi(g) = (b_{ij})$.

Then $$\begin{aligned}(f+g)(v_i, v_j) &= f(v_i, v_j) + g(v_i, v_j) \\ &= a_{ij} + b_{ij}\end{aligned}$$

$$\therefore \varphi(f+g) = (a_{ij} + b_{ij}) = (a_{ij}) + (b_{ij}) = \varphi(f) + \varphi(g).$$

Also $(\alpha f)(v_i, v_j) = \alpha f(v_i, v_j) = \alpha a_{ij}$

$$\therefore \ \varphi(\alpha f) = (\alpha a_{ij}) = \alpha(a_{ij}) = \alpha\varphi(f).$$

Thus φ is an isomorphism.

Corollary. $L(V,V,F)$ is a vector space of dimension n^2.

5.3 Quadratic forms:

Definition:

A bilinear form f defined on a vector space V is called a symmetric bilinear form if $f(u,v) = f(v,u)$ for all $u, v \in V$.

Examples

(i) Let V be a vector space over **R**. Then any inner product defined on V is a symmetric bilinear form.

(ii) The bilinear form $\hat{0}$ defined in example 2 of 5.2 is a symmetric bilinear form.

(iii) Let f be a bilinear form on V. Then the bilinear form f_1 defined by $f_1(u,v) = f(u,v) + f(v,u)$ is a symmetric bilinear form.

Theorem 3:

A bilinear form f defined on V is symmetric iff its matrix (a_{ij}) w.r.t any one basis $\{v_1, v_2, \dots \dots, v_n\}$ is symmetric.

Proof:

Let f be a symmetric bilinear form.

Now,

$$a_{ij} = f(v_i, v_j)$$
$$= f(v_j, v_i) \text{ (since } f \text{ is symmetric)}$$
$$= a_{ji}$$

$\therefore$ (a_{ij}) is a symmetric matrix.

Conversely, let (a_{ij}) be a symmetric matrix.

Hence $A = A^T$Then

$$f(u,v) = XAY^T$$
$$= (XAY^T)^T \text{ (since } XAY^T \text{ is a } 1 \times 1 \text{ matrix)}$$
$$= YA^TX^T$$
$$= YAX^T$$
$$= f(v,u)$$

f is a symmetric bilinear form.

Definition:

Let f be a symmetric bilinear form defined by V. Then the quadratic form associated with f is the mapping $q: V \to F$ defined by $q(v) = f(v,v)$. The matrix of the bilinear form f is called the matrix of the associated quadratic form q.

Examples

1. Consider the bilinear form f defined on $V_n(F)$ by $f(u, v) = x_1y_1 + x_2y_2 + \cdots \ldots + x_ny_n$; $u = (x_1, \ldots \ldots, x_n), v = (y_1, \ldots, y_n)$. Then the quadratic form q associated with f is given by

$$q(u) = f(u, u) = x_1^2 + \cdots \ldots + x_n^2.$$

2. Let A be a symmetric matrix of order n associated with the symmetric bilinear form f. Then the corresponding quadratic form is given by

$$q(X) = XAX^T = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

For example, consider the symmetric matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 6 \end{pmatrix}$$

The quadratic form q determined by A w.r.t. the standard basis for V_3(R) is given by

$$q(v) = \left(x_1, x_2, x_3 \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 6 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right)$$
$$= x_1^2 + 4x_2^2 + 6x_3^2 + 4x_1x_2 + 14x_2x_3 + 6x_1x_3.$$

3. Consider the diagonal matrix

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

The quadratic form q determined by A w.r. the standard basis for $V_3(\mathbf{R})$ is given by

$$q(v) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
$$= x_1^2 + 2x_2^2 + 3x_3^2.$$

We say that this quadratic form q is in the diagonal form.

4. Consider the quadratic form defined on $V_2(R)$ by $q(x_1, x_2) = 2x_1^2 + x_1x_2 + x_2^2$.

Then the symmetric matrix associated with q can be found as follows.

Let

$$2x_1^2 + x_1x_2 + x_2^2 = (x_1, x_2)\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= ax_1^2 + 2bx_1x_2 + cx_2^2$$

$$\therefore\ a = 2; b = \frac{1}{2}; c = 1.$$

$$\therefore\ A = \begin{bmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$$

Exercises

1. Find the quadratic forms associated with the following matrices w.r.t. the standard basis.

(a) $\begin{pmatrix} 2 & -3 & 1 \\ -3 & 2 & 4 \\ 1 & 4 & -5 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & -2 & -4 \\ 3 & -4 & -3 \end{pmatrix}$

2. Find the matrices for the following quadratic forms.

(a) $x_1^2 + 4x_1x_2 + 3x_2^2$ in $V_2(\mathrm{R})$ (b) $2x_1^2 + x_2^2 + 3x_1x_2$ in $V_2(\mathrm{R})$

(c) $2x_1^2 + x_3^2 - 6x_1x_2$ in $V_3(\mathrm{R})$ (d) x_1x_2 in $V_4(\mathrm{R})$

(e) $x_1x_2 + x_4^2$ in $V_4(\mathrm{R})$.

Answers.

3. (a) $2x_1^2 + 2x_2^2 - 5x_3^2 - 6x_1x_2 + 8x_2x_3 + 2x_1x_3$.

(b) $x_1^2 - 2x_2^2 + 3x_3^2 + 4x_1x_2 - 8x_2x_3 + 6x_1x_3$.

4. (a) $\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$ (b) $\begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}$

(c) $\begin{pmatrix} 2 & -3 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Theorem 4:

Let f be a symmetric bilinear form defined on V. Let q be the associated quadratic form.

(i) $f(u,v) = \frac{1}{4}\{q(u+v) - q(u-v)\}$

(ii) $f(u,v) = \frac{1}{4}\{q(u+v) - q(u) - q(v)\}$

Proof:

(i) $\frac{1}{4}\{q(u+v) - q(u-v)\}$

$$=\frac{1}{4}\{f(u+v, u+v) - f(u-v, u-v)\}$$

$$=\frac{1}{4}\{f(u,u) + f(u,v) + f(v,u) + f(v,v)$$
$$-f(u,u) + f(u,v) + f(v,u) - f(v,v)\}$$

$$=\frac{1}{4}\{4f(u,v)\}$$

$$=f(u,v).$$

Proof of (ii) is similar.

Note. the above theorem shows that if f is a symmetric bilinear form and q the assouated quadratic form, then $f(u,v)$ can be determined from q.

Exercises

1. If q is a quadratic form prove that
$q(u+v+w)-q(u+v)-q(v+w)-q(u+w)+q(u)+q(v)+q(w)=0.$
2. Show that if q_1 is the quadratic form associated with the bilinear form f_1 and q_2 is the quadratic form associated with the bilinear form f_2 then q_1+q_2 is the quadratic form associated with the bilinear form f_1+f_2.

5.4. Reduction of a quadratic form ω the diagonal form:

we have seen that a quadratic form associated with a diagonal matrix of order n is of the form

$$a_1x_1^2+a_2x_2^2+\cdots\ldots+a_nx_n^2$$

which is known as the diagonal form. Now, we prove that any quadratic form can be reduced to the diagonal form by means of a non- singular linear transformation. The method of reduction which we describe below is due to Lagrange.

Consider the quadratic form

$$\varphi=\varphi(x_1,x_2,\ldots\ldots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j$$

$$=a_{11}x_1^2+\cdots\ldots+a_{nn}x_n^2+2a_{12}x_1x_2+\cdots\ldots+2a_{n(n-1)}x_nx_{n-1}.$$

Case (i) Suppose at least one of $a_{11},\ldots,a_{mn}$ is not zero. We assume, without loss of generality, that $a_{11}\neq 0$.Then

$$\varphi = (a_{11}x_1^2 + 2a_{12}x_1x_2 + \cdots \ldots + 2a_{1n}x_1x_n) + \sum_{i,j=2}^{n} a_{ij}x_ix_j$$

$$= a_{11}\left(x_1^2 + 2\frac{a_{12}}{a_{11}}x_1x_2 + \cdots \ldots + 2\frac{a_{1n}}{a_{11}}x_1x_n\right) + \varphi_1(x_2, \ldots \ldots x_n) \text{ (say)}$$

$$= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n\right)^2 + \varphi_2(x_2, \ldots \ldots \ldots, x_n)(\text{ say })$$

Now, putting $y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \cdots + \frac{a_{1n}}{a_{11}}x_n$, $y_2 = x_2, \ldots, y_n = x_n, \varphi$ reduces to

$$\varphi = \alpha_1 y_1^2 + \varphi_2(y_2, \ldots, y_n) \ldots \ldots \ldots (1)$$

where $\alpha_1 = a_{11}$

Case (ii) Suppose $a_{11} = a_{22} = \cdots \ldots = a_{nn} = 0$. We still have $a_{ij} \neq 0$ for some i, j such that $i \neq j$.

Without loss of generally we assume that $a_{12} \neq 0$.

Then the non-singular linear transformation

$$x_1 = y_1, x_2 = y_1 + y_2, x_3 = y_3, \ldots \ldots, x_n = y_n$$

changes the quadratic form φ to another quadratic form in which the term y_1^2 is present.

Now applying the method of case (i) φ can be reduced to the form (1). Treating φ_2 in the same way we get

$\varphi_2 = \alpha_2 z_2^2 + \varphi_3(z_2, \ldots, z_n)$ so that

$$\varphi = \alpha_1 z_1^2 + \alpha_2 z_2^2 + \varphi_3(z_2, \ldots, z_n)$$

Continuing this process of reduction, we obtain φ in the form $\varphi = \alpha_1 w_1^2 + \cdots \ldots + \alpha_r w_r^2$.

Solved problems

Problem 1:

Reduce the quadratic form

$x_1^2 + 4x_1x_2 + 4x_1x_3 + 4x_2^2 + 16x_2x_3 + 4x_3^2$ to the diagonal form.

Solution:

Let

$$\varphi = x_1^2 + 4x_1x_2 + 4x_1x_3 + 4x_2^2 + 16x_2x_3 + 4x_3^2$$
$$= (x_1 + 2x_2 + 2x_3)^2 + 8x_2x_3$$

Putting $x_1 + 2x_2 + 2x_3 = y_1, x_2 = y_2, x_3 = y_2 + y_3$ we get

$$\varphi = y_1^2 + 8y_2^2 + 8y_2y_3$$
$$= y_1^2 + 8\left(y_2 + \frac{1}{2}y_3\right)^2 - 2y_3^2.$$

Putting $z_1 = y_1, z_2 = y_2 + \frac{1}{2}y_3, z_3 = y_3$ we get

$\varphi = z_1^2 + 8z_2^2 - 2z_3^2$ where $z_1 = x_1 + 2x_2 + 2x_3$

$$z_2 = \frac{1}{2}(x_2 + x_3)$$
$$z_3 = x_3 - x_2$$

Problem 2:

Reduce the quadratic form $2x_1x_2 - x_1x_3 + x_1x_4 - x_2x_3 + x_2x_4 - 2x_3x_4$ to the diagonal form using Lagrange's method.

Solution:

Let $\varphi = 2x_1x_2 - x_1x_3 + x_1x_4 - x_2x_3$

$$+x_2x_4 - 2x_3x_4.$$

Putting $x_1 = y_1; x_2 = y_1 + y_2; x_3 = y_3$ and $x_4 = y_4$, we get

$$\varphi = 2y_1^2 + 2y_1y_2 - 2y_1y_3 + 2y_1y_4 - y_2y_3 + y_2y_4 - 2y_3y_4$$
$$= 2(y_1^2 + y_1y_2 - y_1y_3 + y_1y_4) - y_2y_3 + y_2y_4 - 2y_3y_4$$

$$= 2\left(y_1 + \frac{1}{2}y_2 - \frac{1}{2}y_3 + \frac{1}{2}y_4\right)^2 - \frac{1}{2}y_2^2 - \frac{1}{2}y_3^2 - \frac{1}{2}y_4^2 - y_3y_4$$

Putting $z_1 = y_1 + \frac{1}{2}y_2 - \frac{1}{2}y_3 + \frac{1}{2}y_4; z_2 = y_2; z_3 = y_3$; and $z_4 = y_4$ we get

$$\varphi = 2z_1^2 - \frac{1}{2}z_2^2 - \frac{1}{2}z_3^2 - \frac{1}{2}z_4^2 - z_3z_4$$

$$= 2z_1^2 - \frac{1}{2}z_2^2 - \frac{1}{2}(z_3^2 + 2z_3z_4 + z_4^2)$$

$$= 2z_1^2 - \frac{1}{2}z_2^2 - \frac{1}{2}(z_3 + z_4)^2$$

Putting $w_1 = z_1, w_2 = z_2, w_3 = z_3 + z_4, w_4 = w_4$

we get $\varphi = 2w_1^2 - \frac{1}{2}w_2^2 - \frac{1}{2}w_3^2$

here

$$w_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{1}{2}x_4$$

$$w_2 = -x_1 + x_2; w_3 = x_3 + x_4; w_4 = x_4.$$

Exercises:

Reduce the following quadratic forms to agonal form.

1. $x_1^2 + 2x_2^2 - 7x_3^2 - 4x_1x_2 + 8x_1x_3$
2. $2x_1^2 + 5x_2^2 + 19x_3^2 - 24x_4^2 + 8x_1x_2 + 12x_1x_3 + 8x_1x_4 + 18x_2x_3 - 8x_2x_4 - 16x_3x_4$
3. $2x_1x_2 - x_1x_3 + x_2x_3$
4. $-2x_1x_2 + 2x_2x_3 - 2x_3x_4 + 2x_1x_4$
5. $(x_1x_2x_3)\begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & -2 \\ 4 & -2 & 18 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$
6. $(x_1x_2x_3)\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & -1 \\ 2 & -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$

Answers:

1. $z_1^2 - 2z_2^2 + 9z_3^2$ 2. $z_1^2 - 3z_2^2 + 4z_3^2$

3. $2z_1^2 - \frac{1}{2}z_2^2 + \frac{1}{2}z_3^2$ 4. $z_1^2 - z_2^2 + z_3^2 - z_4^2$

5. $z_1^2 + 2z_2^2 - 48z_3^2$ 6. $z_1^2 - z_2^2 + 8z_4^2$

Study Learning Material Prepared by

Dr. S. KALAISELVI M.SC., M.Phil., B.Ed., Ph.D.,

ASSISTANT PROFESSOR,

DEPARTMENT OF MATHEMATICS,

SARAH TUCKER COLLEGE (AUTONOMOUS),

TIRUNELVELI-627007.

TAMIL NADU, INDIA.