SEMESTER – V Linear Algebra

L	T	P	C
3	2	0	4

CORE PAPER – VII
ABSTRACT ALGEBRA II (75 Hours) (SMMA51)

Objectives:

- To facilitate a better understanding of vector space
- To solve problems in matrices

Unit I **Vector Spaces :** Definition and examples – elementary properties – subspaces – linear transformation – fundamental theorem of homomorphism **16L.**

Unit II Span of a set – linear dependence and independence – basis and dimension – theorems **14L**

Unit III Rank and nullity Theorem – matrix of a linear transformation
Inner product space : Definition and examples – orthogonality – orthogonal complement – Gram Schmidt orthognalisation process. **15L**

Unit IV **Matrices :** Elementary transformation – inverse – rank -Cayley Hamilton Theorem-Applications of Cayley Hamilton Theorem **15L**

Unit V Eigen values and Eigen vectors – Properties and problems-Bilinear Forms-Quadratic Forms-Reduction of quadratic form to diagonal form **15L**

Text Book:

Arumugam & Issac – Modern Algebra

Books for Reference :

- Shama .J.N and Vashistha .A.R, "Linear Algebra", Krishna Prakash Nandir, 1981.
- John B. Fraleigh, "A First Course in Abstract Algebra", 7th edition, Pearson, 2002.
- Strang G., "Introduction to Linear Algebra", 4th edition, Wellesly Cambridge Press, Wellesly, 2009.
- Artin M., "Abstract Algebra", 2nd edition, Pearson, 2011

ABSTRACT ALGEBRA - II

UNIT - I
VECTOR SPACE

Definition and Examples

Definition: A non-empty set V is said to be a vector space over a field Fif

(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) $1u = u$ for all $u \in V$.

Remark

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

   $F \times V \rightarrow V$ defined by $(\alpha, v) \rightarrow \alpha v$.

Examples

1. $R \times R$ is a vector space over a field R under the 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 $(R\times R,+)$ is an abelian group.

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

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

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

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

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

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

$=\alpha u+\alpha v.$

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

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

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

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

Obviously $1u = u$

∴ $R \times R$ is a vector space over R.

2. $R^n= \{(x_1, x_2, . . . ,x_n) : x_i\in R, 1 \le i \le n\}$. Then $R^n$is a vector space over R under addition and scalar multiplication defined by $(x_1,x_2,...,x_n)+(y_1,y_2,...,y_n)= (x_1+y_1,x_2+y_2,...,x_n+y_n)$ and $\alpha(x_1,x_2,...,x_n)= (\alpha x_1,\alpha x_2,...,\alpha x_n)$.

Proof:

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

Theinverseof$(x_1,x_2,...,x_n)$I s$(-x_1,-x_2,...,-x_n)$.

Hence$(R^n,+)$ is an abelian group.

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

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

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

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

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

$=\alpha(x_1,x_2,...,x_n)+\alpha(y_1,y_2,...,y_n)=\alpha u+\alpha v.$

Similarly$(\alpha+\beta)u=\alpha u+\beta uand\alpha(\beta u) = (\alpha\beta)u$.

∴ $1u = u$.

∴$R^n$is vector spaceoverR.

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

Theorem: Let V be a vector space over a field F ,Then

(i) $\alpha 0 = 0$ 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) $\alpha v= 0 \Rightarrow \alpha = 0$ or $v =0$.

Proof:

(i) $\alpha 0 = \alpha(0 + 0) = \alpha 0 + \alpha 0$. Hence $\alpha 0 = 0$.

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

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

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 is nothing to prove.

$\therefore$ Let α ƒ= 0. Then $\alpha^{-1} \in F$.

Now, $v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1}0=0$.

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: Let V be a vector space over a field F. A non-empty subset W of V is a subspace of V if and only if W is closed with respect to vector addition and scalar multiplication 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: Let V be a vector space over a field F. A non-empty subset W ofVis a subspace of V if and only if $u, v \in W$ 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 $u \in W \Rightarrow \alpha \in F$ and $u \in W \Rightarrow \alpha u \in W$.

Hence 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. $W = \{(a, 0, 0): a \in R\}$ is a subspace of R^3,

For, let $u = (a, 0, 0)$, $v = (b, 0, 0) \in W$ and $\alpha,\beta \in R$.

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 R^3.

<u>Solved problems</u>

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

Solution.

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

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

Solution. Let $A= \{(a,0,0): a \in R\}$, $B = \{(0,b,0): b \in R\}$.

Clearly A and B are subspaces of R^3.

However A∪B is not a subspace of R^3.

For,(1,0,0)and(0,1,0)∈A∪B.But(1,0,0)+(0,1,0)=(1,1,0)∉A∪B.

Problem:If A and B are subspaces of Vprove that A + B = {v ∈ V: v =a + b, a ∈ A, b ∈B} is a subspace of V. Further show that A + B is the smallest subspace containing A and B. (ie.,) If W isany subspace of V containing A and B then W contains A +B.

Solution. Let v_1, v_2∈ A + B and α ∈ F.

Then v_1= a_1+ b_1, v_2= a_2+ $b_2$where a_1, a_2∈A,andb_1,b_2∈B.

Now,$v_1+v_2=(a_1+b_1)+(a_2+b_2)=(a_1+a_2)+(b_1+b_2)$∈A+B.

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

HenceA+Bis a subspace of V.

Clearly A⊆A+B and B ⊆ A + B.

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

Weshall prove that A+B⊆W.

Let v∈A+B.

Then v=a+b where a∈A and b∈B. Since A⊆W,a∈W.

Similarly b∈W and a+ b=v∈W.

Therefore A+B⊆Wsothat A+B is the smallest subspace of Vcontaining A and B.

Problem: Let A and B be subspace of a vector space V. Then A∩B = {0} if and only if every vector v∈A+B can be uniquely expressed in the form v=a+bwhere a ∈A and b ∈B.

Solution. Let A∩B={0}. Letv∈A+B.

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

Then$a_1-a_2=b_2-b_1$.

But a_1-a_2∈A and b_2-b_1∈B.

Hence a_1-a_2,b_2-b_1∈A∩B.

SinceA∩B={0}, $a_1-a_2=0$ 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 ∈ A and b ∈ B is unique.

Conversely suppose that anyelement in A + B can be uniquely expressed in the forma+ b where a∈A and b∈B.

We claim that A∩ B= {0}.

If $A\cap B = \{0\}$, let $v \in A\cap B$ and $v = 0$. Then $0 = v - v = 0 + 0$.

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

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: $V = A\oplus B$ If and only if everyelement of V can be uniquely expressed in the form a+b where $a\in A$ and $b\in B$.

Examples

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

Theorem: 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 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 subspace $\Rightarrow \alpha v_1 - \alpha v_2 \in W \Rightarrow \alpha v_1 \in W + \alpha v_2 \Rightarrow W + \alpha v_1 = W + \alpha v_2$.

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

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

Hence V/W is an abelian group.

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

$\alpha[(W+v_1)+(W+v_2)]=\alpha(W+v_1+v_2)$

$=W+\alpha(v_1+v_2)$

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

$= (W+\alpha v_1)+(W+\alpha v_2)$

$=\alpha(W+v_1)+\alpha(W+v_2)$

$(\alpha+\beta)(W+v_1)=W+(\alpha+\beta)v_1$

$=W +\alpha v_1+\beta v_1$

$=(W+\alpha v_1)+(W +\beta v_1)$

$=\alpha(W+v_1)+\beta(W+v_1)$

$\alpha[\beta(W+v_1)]=\alpha(W+\beta v_1)$

$=W+\alpha\beta v_1$

$1(W+v_1)=W+1v_1$

$=W+v_1$

Hence V/W is a vector space.

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

Linear transformation

Definition Let V and W be a vector space over a field F. A mapping $T: V \rightarrow W$ 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 space 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 then 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 a identity linear transformation.

Theorem: *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 a 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 = {0}.

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

Theorem[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 $V_1 = \ker T = \{v : v \in V \text{ and } T(v) = 0\}$

Clearly T (0) = 0.

Hence $0 \in \ker T = V_1$

$\therefore V_1$ is non-empty subset of V.

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

$\therefore T(u) = 0$ and $T(v) = 0$.

Now $T(\alpha u + \beta v) = T(\alpha u) + T(\beta v)$

$= \alpha T(u) + \beta T(v)$

$= \alpha 0 + \beta 0 = 0$ and so $\alpha u + \beta v \in \ker T$.

Hence kerT 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 $v_1 \in V$.

$\therefore T(v)=T(v_1+w)$

$=T(v_1) + T(w) = 0 + T(w)$

$=T(w)$

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

$\therefore \varphi$ is 1-1.

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

$\Rightarrow T(v) = T(w)$

$\Rightarrow T(v) - T(w) = 0$

$\Rightarrow T(v) + T(-w) = 0$

$\Rightarrow T(v-w)=0$

$\Rightarrow v-w \in \ker T=V_1$

$\Rightarrow v \in V_1+w$

$\Rightarrow V_1+v=V_1+w$.

Φ is onto.

Let $w \in W$.

Since T is onto, there exists $v \in V$ such that $T(v)=w$ and so $\varphi(V_1+v)=w$.

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

Hence φis an isomorphism.

Theorem: 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 vector space over F.

An element of A+B is of the form (a+b) where a∈Aand b∈B. But A + a = A.

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

Now, consider f: B $\rightarrow \frac{A+B}{A}$

Defined $\frac{A+B}{A}$ is of the form A + b.

Now, consider f: B $\rightarrow \frac{A+B}{A}$ by f(b) = A+b.

Clearly f is onto.

Also $f(b_1+b_2)=A+(b_1+b_2)$

$= (A+b_1)+(A+b_2)$

$=f(b_1)+f(b_2)$ and

$f(\alpha b_1)= A+\alpha b_1=\alpha(A+b_1) = \alpha f(b_1)$.

Hence f is a linear transformation.

Let K be the kernel off.

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

Now,A+b=Aifandonlyifb∈A.HenceK=A∩Band so $\frac{A+B}{A} \cong \frac{B}{A\cap B}$

Theorem: LetVandW bevectorspacesoverafieldF.LetL(V,W)represent thesetofalllineartransformationsfromVtoW.ThenL(V,W)itselfisavectorspace over F under addition and scalar multiplication defined by (f + g)(v) = f(v) + g(v) and (αf)(v) =αf(v).

Proof. Let f, g ∈L(V, W) and v_1, v_2∈ V .

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

= 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)

Also (f + g)(αv) = f(αv) + g(αv) = αf(v) + αg(v) = α[f(v) + g(v)] = α(f + g)(v).

Hence (f + g) ∈L(V, W).

Now,(αf)(v_1+ v_2) = (αf)(v_1) + (αf)(v_2) = αf(v_1) + αf(v_2)

= α[f(v_1) + f(v_2)] = αf(v_1+ v_2).

Also(αf)(βv)=α[f(βv)]=α[βf(v)]=β[αf(v)]=β[(αf)(v)].

Henceαf∈L(V,W). Addition defined on L(V, W) is obviously commutative andassociative.

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

Further (−f): V → Wdefined 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)isavectorspaceoverF.

UNIT - II
SPAN OF A SET

Definition:

Let V be a vector space over a field F.Let $v_1, v_2, \ldots, v_n \in V$. Then an element of the form $\alpha_1 v_1+\alpha_2 v_2+\cdots+\alpha_n v_n$ where $\alpha_i \in F$ is called a linear combination of the vectors $v_1, v_2, \ldots, v_n$.

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 byL(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 $\alpha_1, \alpha_2, \ldots, \alpha_n \in F$.

Theorem: 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.,) 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, $w = \beta_1 w_1+ \beta_2 w_2+ \cdots + \beta_m w_m$ where $w_j \in S$ $\beta_j \in F$.

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

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

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

Hence $\alpha v + \beta w \in L(S)$ and so L(S) is asubspace of S.

(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 $S \subseteq W$.

Let $u \in L(S)$.

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

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

Hence $L(S) \subseteq W$.

Note: $L(S)$ is called the subspace spanned(generated) 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 $L(S) = \{\alpha e_1 + \beta e_2 : \alpha, \beta \in R\} = \{(\alpha, \beta, 0) : \alpha, \beta \in R\}$

(b) Let $S = \{e_1, e_2, e_3\}$. Then $L(S) = \{\alpha e_1 + \beta e_2 + \gamma e_3 : \alpha, \beta, \gamma \in R\} = \{(\alpha, \beta, \gamma) : \alpha, \beta, \gamma \in R\} = V_3(R)$ Thus $V_3(R)$ is spanned by $\{e_1, e_2, e_3\}$.

2. In $V_n(R)$ let $e_1 = (1, 0, \cdots, 0)$; $e_2 = (0, 1, 0, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$.

Let $S = \{e_1, e_2, \ldots, e_n\}$. Then $L(S) = \{\alpha_1 e_1 + \alpha_2 e_2 + \alpha_n e_n : \alpha_i \in R\} = \{(\alpha_1, \alpha_2, \ldots, \alpha_n) : \alpha_i \in R\} = V_n(R)$

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

Theorem: 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) $L(S) = S$ if and only if 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 $\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_n s_n \in L(T)$.

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 $s_1, s_2, \ldots, s_m \in S$ and $s_{m+1}, \ldots, s_n \in T$.

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_m s_m \in L(S)$ and $\alpha_{m+1} s_{m+1} + \cdots + \alpha_n s_n \in L(T)$.

Therefore $S = (\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_m s_m) + (\alpha_{m+1} s_{m+1} + \cdots + \alpha_n s_n) \in L(S) + L(T)$.

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

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

Conversely, let Sbe asubspace of V.

Then the smallest subspace containing S is S itself.

Hence$L(S)=S$.

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

Linear Independence

In $V_3(R)$, let $S = \{e_1, e_2, e_3\}$. We have seen that $L(S) = V_3(R)$. Thus S is a subset of$V_3(R)$ which spans the whole space $V_3(R)$.

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. $V_3(R)$ is a finite dimensional vector space.
2. $V_n(R)$ is a finite dimensional vector space, since $S=\{e_1,e_2,...,e_n\}$ is a finite sub- set of $V_n(R)$ such that $L(S) = V_n(R)$. In general if F is any field $V_n(F)$ is a finite dimensional vector space over F.

Definition: Let V be a vector space over a field F. A finite set of vectors $v_1,v_2,...,v_n$inVissaidtobelinearlyindependentif$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n=0\Rightarrow\alpha_1=\alpha_2=\cdots=\alpha_n=0$.If$v_1,v_2,...,v_n$are not linearly independent ,then they are said to be linearly dependent.

Note: If $v_1,v_2,...,v_n$are called linearly dependent then there exists scalars$\alpha_1,\alpha_2,...,\alpha_n$not all zero such that $\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n= 0$.

Examples:

1. In$V_n(F)$,$\{e_1,e_2,...,e_n\}$ is a linearly independent set of vectors, for $\alpha_1e_1+\alpha_2e_2+\cdots+\alpha_ne_n=0$.

   $\Rightarrow\alpha_1(1,0,...,0)+\alpha_2(01,...,0)+\cdots+\alpha_n(0,0,...,1)=(0,0,...,0)$

   $\Rightarrow(\alpha_1, \alpha_2, . . . , \alpha_n) = (0, 0, . . . , 0) \Rightarrow \alpha_1= \alpha_2= \cdots = \alpha_n= 0$.

2.In$V_3(R)$thevectors(1,2,1),(2,1,0)and(1,−1,2)arelinearlyindependent.

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

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

$\therefore \quad \alpha_1+2\alpha_2+\alpha_3=0 \quad \cdots(1)$

$2\alpha_1+\alpha_2-\alpha_3=0 \quad \cdots(2)$

$\alpha_1+2\alpha_3=0 \quad \cdots(3)$

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

$\therefore$The given vectors are linearlyindependent.

3.In$V_3(R)$thevectors(1,4,−2),(−2,1,3)and(−4,11,5)arelinearlydependent.For, let$\alpha_1(1,4,-2)+\alpha_2(-2,1,3)+\alpha_3(-4,11,5)=(0,0,0)$

$\therefore \quad \alpha_1-2\alpha_2-4\alpha_3=0 \quad \cdots(1)$

$4\alpha_1+\alpha_2+11\alpha_3=0\cdots(2)$

$-2\alpha_1+3\alpha_2+5\alpha_3=0\cdot\cdot(3)$

From (1) and (2),

$\alpha_1=-18k, \alpha_2=-27k, \alpha_3=9k$.These values of α_1, $\alpha_2$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.

Theorem: Anysubsetofalinearlyindependentsetislinearlyindependent.

Proof: LetVbeavectorspaceoverafieldF.

Let$S=\{v_1,v_2,...,v_n\}$bealinearly independent set.

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

Suppose $S'$is a linearly dependent set.

Then there exist$\alpha_1,\alpha_2,...,\alpha_k$inFnotallzero,suchthat$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_kv_k=0$.

Hence $\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_kv_k+0v_{k+1}+\cdots+0v_n=0$isanon-triviallinearcombination giving the zero vector. Here S is a linearly dependent set which is a contradiction.
Hence S'islinearlyindependent.

Theorem: Anysetcontainingalinearlydependentsetisalsolinearlydependent.

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

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

Theorem: Let$S=\{v_1,v_2,...,v_n\}$bealinearlyindependentsetofvectorsina vectorspaceVoverafieldF. TheneveryelementofL(S)canbeuniquelywrittenin theform$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$,where$\alpha_i\in F$.

Proof.BydefinitioneveryelementsofL(S)isoftheform$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$.

Now,$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n=\beta_1v_1+\beta_2v_2+\cdots+\beta_nv_n$.

Hence$(\alpha_1-\beta_1)v_1+(\alpha_2-\beta_2)v_2+\cdots+(\alpha_n-\beta_n)v_n=0$.

SinceSisalinearlyindependentset,$\alpha_i-\beta_i=0$foralli.

$\therefore \alpha_i=\beta_i$for all i. Hencethetheorem.

Theorem:$S=\{v_1,v_2,...,v_n\}$bealinearlyindependentsetofvectorsina vector space V if and only if there exists a vector $v_k\in S$ such that $v_k$is a linear combinationoftheprecedingvectors$v_1,v_2,...,v_{k-1}$.

Proof: Suppose$v_1,v_2,...,v_n$arelinearlydependent.
Thenthereexist$\alpha_1,\alpha_2,...,\alpha_n\in F$, not all zero, such that $\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n=0$.
Let k be the largest integer for which α_kƒ= 0.
Then $\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_kv_k=0$. $\therefore \alpha_kv_k=-\alpha_1v_1-\alpha_2v_2-\cdots-\alpha_{k-1}v_{k-1}$.
$\therefore v_k=(-\alpha^{-1}\alpha_1)v_1+\cdots+(-\alpha^{-1}\alpha_{k-1})v_{k-1}$.

$\therefore v_k$is a linear combination of thepreceding vectors.

Conversely,

suppose there exists a vector $v_k$such that$v+k=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_{k-1}v_{k-1}$.

Hence$-\alpha_1v_1-\alpha_2v_2-\cdots-\alpha_{k-1}v_{k-1}+v_k+0v_{k+1}+\cdots+0v_n=0$.

Sincethecoefficientof$v_k=1$, wehave$S=\{v_1,v_2,...,v_n\}$islinearlydependent.

Example: In$V_3(R)$,letS=[(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 dependentset.

Theorem:LetVbeavectorspaceoverF.Let$S=\{v_1,v_2,...,v_n\}$and L(S)=W.Thenthereexists alinearlyindependent subsets$S'$of SsuchthatL(S')= W.

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

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 precedingvectors.Let$S_1=\{v_1,v_2,...,v_{k-1},v_{k+1},...,v_n\}$.(ie.,)$S_1$isobtainedby deleting the vector $v_k$from S.

Weclaim that $L(S_1)=L(S)=W$.

Since $S_1 \subseteq S$, $L(S_1) \subseteq L(S)$.

Now, let $v \in L(S)$.

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

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

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

∴ 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 $L(S') = W$.

Basis and dimension:

Definition: A linearly independent subset S of a vector space V which spans thewholespaceViscalledabasis ofthevectorspace.

Theorem:

Any finite dimensional vector space V contains a finite number of linearlyindependentvectorswhichspanV.(ie.,)Afinitedimensionalvectorspacehas abasisconsistingofafinitenumberofvectors.

Proof: Since V is finite dimensional there exists a finite subset S of V such that L(S)=V.ClearlythissetScontainsalinearlyindependentsubset$S'=\{v_1,v_2,...,v_n\}$ suchthat$L(S')=L(S)=V$HenceS'isabasisforV.

Theorem:LetVbeavectorspaceoverafieldF.Then$S=\{v_1,v_2,...,v_n\}$is abasisforVifandonlyifeveryelementofVcanbeuniquelyexpressedasalinear combination of element ofS.

Proof: Let S be a basis for V.

Then by definition S is linearly independent and L(S)=V.

HencebytheoremeveryelementofV canbeuniquelyexpressedasa linear combination of elements ofS.

Conversely, suppose every element of V can be uniquely expressed as a linearcombinationofelementsofS.

ClearlyL(S)=V.

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

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

Thuswehaveexpressed 0asalinearcombinationof vectors of S in two ways.

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

Hence S is linearly independent. Hence S isabasis.

Examples

1.S={(1,0,0),(0,1,0),(0,0,1)}isabasisforV_3(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(R) has been expressed uniquely as a linear combination of the elements of S and hence S is a basis for V_3(R).

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

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

4. {1, i} a basis for the vector space C overR.

Theorem: Let V be a vector space over a field F. LetS={v_1,v_2,...,v_n} spanV.LetS = {w_1,w_2,...,w_n}bealinearlyindependentsetofvectorsin V.Then m ≤n.

Proof. Since L(S) = V, every vector in V and in particular w_1, is a linear combination ofv_1,v_2,...,v_n.

HenceS_1={w_1,v_1,v_2,...,v_n}is a linear independent set of vectors. Hence there exists a vector v_kf= $w_1$in S_1which is a linear combination of the preceding vectors.

LetS_2={w_1,v_1,...,v_{k-1},v_{k+1},...,v_n}.

Clearly,L(S_2)=V.

Hence$w_2$is a linearcombinationofthevectorsinS_2.

HenceS_3={w_2,w_1,v_1,...,v_{k-1},v_{k+1},...,v_n} is linearly dependent. Hence there exists a vector in S_3which is a linear combination of the preceding vectors. Since the w_i's are linearly independent, this vectorcannotbe $w_2$or $w_1$and hence must be some $v_j$wherejk(say, with j >k).

Deletionof$v_j$fromthesetS$_3$givesthesetS$_4$={w_2,w_1,v_1,...,v_{k-1},v_{k+1},...,v_{j-1},v_{j+1},...,v_n}of n vectors spanningV.

Inthisprocess,ateachstepweinsertonevectorfrom{w_1, w_2, . . . ,w_m} and delete one vector from {v_1, v_2, . . . , v_n}.

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

Hence w_{n+1} is a linear combination of $w_1, w_2, \ldots, w_n$.

Hence $\{w_1, w_2, \ldots, w_n, w_{n+1}, \ldots, w_n\}$ is linearly dependent which is a contradiction.

Hence $m \leq n$.

Theorem: 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 $S = \{v_1, v_2, \ldots, v_n\}$.

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

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

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.

Theorem: 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 $m < n$ cannot span V.

Proof.

(i) Let $S = \{v_1, v_2, \cdots, v_n\}$ 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, $m \leq n$ 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, $S = \{v_1, v_2, \cdots, v_n\}$ is a basis for V and hence linearly independent.

Hence by theorem $n \leq m$ which is a contradiction. Hence S' cannot span V.

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

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

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

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

Now, consider $S_1 = \{v_{1,2}, \ldots, v_r, v_{r+1}\}$.

We shall prove that S_1 is linearly independent by showingthatnovectorinS_1isalinearcombinationoftheprecedingvectors.

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

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

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 basiscontainingS.

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

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

By theorem , we can find $w_1, w_2, \ldots, w_s \in V$ suchthat$S' = \{v_1, v_2, \cdots, v_r, w_1, w_2, \ldots, w_s\}$isabasisofV.Now, let $B = L(\{w_1, w_2, \ldots, w_s\})$.

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

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

Hence $v = \alpha_1 v_1 + \cdots + \alpha_r v_r = \beta_1 w_1 + \cdots + \beta_s w_s$

$\therefore \alpha_1 v_1 + \cdots + \alpha_r v_r - \beta_1 w_1 - \cdots - \beta_s w_s = 0$.

Now,sinceS'islinearlyindependent,$\alpha_i = 0 = \beta_j$ for all i and j.

Hence $v = 0$. Thus $A \cap B = \{0\}$.

Now, let $v \in V$.

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

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

Definition:Let V be a vector space and $S = \{v_1, v_2, \ldots, v_n\}$ be a set of independent vectors in V . Then S is called a maximal linear independent set ifforeveryv∈V−S,theset{v,v_1,v_2,...,v_n} islinearlydependent.

Definition.LetS={v_1,v_2,...,v_n}beasetofvectorsinVandletL(S)=V. ThenSiscalledaminimal generatingsetIf foranyv$_i$∈S,L(S−{v_i})=V.

Theorem: LetV beavectorspaceoverafieldF.LetS={v_1,v_2,...,v_n}⊆V. Then the following areequivalent.

(i) S is a basis for V.

(ii) S is a maximal linearly independentset.

(iii) S is a minimal generatingset.

Proof.(i)⇒(ii)LetS={v_1,v_2,...,v_n}beabasisforV.Thenbytheoremany n+1vectorsinVarelinearlydependentandhenceSisamaximallinearlyindependent set.

(ii)⇒(iii)LetS={v_1,v_2,...,v_n}beamaximallinearlyindependentset.thatS is a basis for V we shall prove that L(S) = V .

Obviously L(S) ⊆V .

Now,letv∈V.

Ifv∈S,thenv∈L(S).(sinceS⊆L(S))

Ifv∉S,S′={v_1,v_2,...,v_n,v}is a linearly dependent set (since S is a maximal independent set)

∴There exists a vectorinS′whichisalinearcombinationofthepreceedingvectors.Sincev_1,v_2,...,$v_n$are linearly independent, this vector must be v. Thus v is a linear combination of v_1, v_2, . . . ,v_n. Therefore v ∈L(S).

Hence V ⊆L(S). Thus V = L(S).

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

If S is not minimal, there exists v_i∈ S such that L(S − {v_i}) = V .

Since S is a linearly independent, S − {v_i} is also linearly independent. Thus S − {v_i} is a basis consisting of n − 1 elementswhichisacontradiction.

HenceSisaminimalgeneratingset.

(iii)⇒((i)

LetS = {v_1, v_2, . . . , v_n} be a minimal generating set. To prove that S is a basis, we haveto 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:*Anyvectorspaceof dimension*n*overafield*F*isisomorphicto*$V_n(F)$.

Proof. LetVbeavectorspaceofdimensionn.Let$\{v_1,v_2,...,v_n\}$beabasisforV.

Thenweknowthatif$v \in V$,vcanbewrittenuniquelyas$v=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n$, where$\alpha_i \in F$.

Now,considerthemap$f:V \to V_n(F)$givenby$f(\alpha_1v_1+\cdots+\alpha_nv_n)= (\alpha_1,\alpha_2,...,\alpha_n)$.

Clearlyfis1-1andonto.

Let$v,w \in V$.

Then$v=\alpha_1v_1+\cdots+\alpha_nv_nandw=\beta_1v_1+\cdots+\beta_nv_n$.

$f(v + w) = f[(\alpha_1+ \beta_1)v_1+ (\alpha_2+ \beta_2)v_2+ \cdots + (\alpha_n+ \beta_n)v_n]$

$=((\alpha_1+\beta_1),(\alpha_2+\beta_2),\cdots,(\alpha_n+\beta_n))$

$=(\alpha_1,\alpha_2,\cdots,\alpha_n)+(\beta_1,\beta_2,\cdots,\beta_n)$

$= f(v)+f(w)$

Also$f(\alpha v)=f(\alpha\alpha_1v_1+\cdots+\alpha\alpha_nv_n)$

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

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

$= \alpha f(v)$.

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 areisomorphic.

Theorem:. 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$\{v_1,v_2,...,v_n\}$beabasisforV.

WeshallprovethatT(v_1),T(v_2),...,T(v_n) are linearly independent and that they span W.

Now,$\alpha_1T(v_1)+\alpha_2T(v_2)+\cdots+ \alpha_nT(v_n) =0$

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

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

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

$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ (since $v_1, v_2, \ldots, v_n$ arelinearlyindependent).

$\therefore T(v_1), T(v_2), \ldots, T(v_n)$ arelinearlyindependent.

Now, let $w \in W$. Then since T is onto, there exists a vector $v \in V$.

suchthat $T(v) = w$.

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

Thenw= $T(v) = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n)$

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

Thuswisa linear combinationofthevectors$T(v_1), T(v_2) \ldots, T(v_n)$.

$\therefore T(v_1), T(v_2) \ldots, T(v_n)$ spanW and hence is a basis forW.

Corollary: Two finite dimensional vector space V and W over a field F are isomorphicifandonlyiftheyhavethesamedimension.

Theorem: Let V and W be finite dimensional vector spaces over a field F. Let$\{v_1, v_2, \cdots, v_n\}$beabasisforVandlet$w_1, w_2, \ldots, w_n$beanynvectorsinW(not necessarilydistinct)Thenthereexistsauniquelineartransformation$T: V \to W$such that$T(v_i) = w_i, i = 1, 2, \ldots, n$.

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

WedefineT$(v) = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n$.

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

Let$x = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_nandy = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_n v_n$

$\therefore (x+y) = (\alpha_1 + \beta_1) v_1 + (\alpha_2 + \beta_2) v_2 + \cdots + (\alpha_n + \beta_n) v_n$

$\therefore T(x+y) = (\alpha_1 + \beta_1) w_1 + (\alpha_2 + \beta_2) w_2 + \cdots + (\alpha_n + \beta_n) w_n$.

$= (\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n) + (\beta_1 w_1 + \beta_2 w_2 + \cdots + \beta_n w_n)$

$= T(x) + T(y)$

Similarly $T(\alpha x) = \alpha T(x)$.

Hence T is a linear transformation.

Also $v_1 = 1v_1+0v_2+\cdots+0v_n$.

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

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

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

Let $v=\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_nv_n \in V$.

$T'(v) = \alpha_1T'(v_1)+\alpha_2T'(v_2)+\cdots+\alpha_nT'(v_n)$

$=\alpha_1w_1+\alpha_2w_2+\cdots+\alpha_nw_n$

$= T(v)$.

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

(i) dim W ≤ dim V.

(ii) $\dim\left(\frac{V}{W}\right) = dimV - dimW$

Proof.

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

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

Let $S=\{w_1,w_2,...,w_m\}$ 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 . Let $S=\{w_1,w_2,...,w_m,v_1,v_2,\cdots,v_r\}$ be a basis for V. Then $m+r=n$. Now, we claim $S' = \{W + v_1, W + v_2, . . . , W + v_r\}$ is a basis for $\frac{V}{W}$.

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

$\Rightarrow (W + \alpha_1v_1) + (W + \alpha_2v_2) + \cdots + (W + \alpha_rv_r) = W$

$\Rightarrow W +\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r=W$

$\Rightarrow \alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r \in W$.

Now,since$\{w_1,w_2,\cdots,w_m\}$isabasisforW,$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r=\beta_1w_1+\beta_2w_2+\cdots+\beta_mw_m$.

Therefore$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r-\beta_1w_1-\beta_2w_2-\cdots-\beta_mw_m=0$.

Hence$\alpha_1=\alpha_2=\cdots=\alpha_r=\beta_1=\beta_2=\cdots=\beta_m=0$andsoS′isalinearlyindependentset.

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

Letv=$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r+\beta_1w_1+\beta_2w_2+\cdots+\beta_mw_m$. Then

$W + v = W + (\alpha_1v_1+ \alpha_2v_2+ \cdots + \alpha_rv_r+ \beta_1w_1+ \beta_2w_2+ \cdots + \beta_mw_m)$

$=W+(\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r)$(since$\beta_1w_1+\beta_2w_2+\cdots+\beta_mw_m\in W$

$=(W+\alpha_1v_1)+ (W+\alpha_2v_2)+\cdots+(W+\alpha_rv_r)$

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

Hence S′spans$\frac{V}{W}$of thatS′isabasisfor$\frac{V}{W}$ and dim$\frac{V}{W}$=r=n−m=dimV−dimW.

Theorem:Let V be a finite dimensional vector space over a field F . LetA

and B be subspaces of V . Then dim (A + B) = dim A + dim B − dim (A ∩ B)

Proof. A and B are subspaces ofV. Hence A ∩ B is subspace of V.

Let dim(A∩B) = r

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

Since A ∩ B is a subspace of A and B, S is a part of a basis for A and B.

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

Weshallprovethat$\{v_1,v_2,\ldots,v_r,u_1,u_2,\ldots,u_s,w_1,w_2,\ldots,w_t\}$beabasisforA+B.

Let$\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r+\beta_1u_1+\beta_2u_2+\cdots+\beta_su_s+\gamma_1w_1+\gamma_2w_2,\cdots+\gamma_tw_t=0$.

Then $\beta_1u_1+\beta_2u_2+\cdots+\beta_su_s=-(\alpha_1v_1+\alpha_2v_2+\cdots+\alpha_rv_r)-(\gamma_1w_1+\gamma_2w_2,\cdots+\gamma_tw_t)\in B$.

Hence $\beta_1u_1+ \beta_2u_2+ \cdots + \beta_su_s\in B$.

Also $\beta_1u_1+ \beta_2u_2+ \cdots + \beta_su_s\in A$.

Hence $\beta_1u_1+\beta_2u_2+\cdots+\beta_s ussss_s\in A\cap B$andso$\beta_1u_1+\beta_2u_2+\cdots+\beta_su_s=\delta_1v_1+\delta_2v_2+\cdots+\delta_rv_r$.

$\beta_1u_1+\beta_2u_2+\cdots+\beta_su_s-\delta_1v_1-\delta_2v_2-\cdots-\delta_rv_r=0$.

Thus $\beta_1=\beta_2=\cdots=\beta_s=\delta_1=\delta_2=\cdots=\delta_r=0$ (Since $\{u_1,u_2,...,u_s,v_1,v_2,...,v_r\}$ is linearly independent)

$\therefore$ Similarly we can prove $\gamma_1=\gamma_2=\cdots=\gamma_t=0$.

Thus $\alpha_i=\beta_j=\gamma_k=0$ for all $1 \le i \le r$; $1 \le j \le s$; $1 \le k \le t$. Thus S' is a linearly independent set.

Clearly S' spans A + B and so S' is a basis for A + B. Hence dim (A + B) = r + s + t.

Also $\dim A=r+s$; $\dim B=r+t$ and $\dim(A\cap B)=r$.

Hence $\dim A+\dim B-\dim A\cap B=(r+s)+(r+t)-r=r+s+t=\dim(A+B)$.

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 $\dim V=\dim A+\dim B$.

UNIT - III
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. Let $T:V\to W$ be a linear transformation. Then $\dim V = rankT + nullityT$.

Proof.

We know that $V/\ker T = T(V)$

$$\therefore \dim V - \dim(\ker T) = \dim(T(V))$$

$$\therefore \dim V - nullityT = rankT$$

$$\therefore \dim V = nullity + rankT$$

Note. $\ker T$ is also called **null space** of *T*.

Example. Let *V* denote the set of all polynomials of $\deg ree \le n$ 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 $\dim V = n+1$, $rankT=n$

Definition. A linear transformation $T:V\to W$ is called ***non-singular*** if *T* is 1-1; otherwise *T*is called singular.

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 $\{w_1, w_2, \ldots, w_m\}$ for W.

Let $T:V\to W$ be a linear transformation. We have seen that T is completely specified by the elements $T(v_1, v_2, \ldots, v_m)$ Now, let

$$T(v_1) = a_{11}w_1 + a_{12}w_2 + \ldots + a_{1n}w_n$$

$$T(v_2) = a_{21}w_1 + a_{22}w_2 + \ldots + a_{2n}w_n$$

...(1)

....

$$T(v_m) = a_{m1}w_1 + a_{m2}w_2 + \ldots + a_{mn}w_n$$

Hence $T(v_1, v_2, \ldots, v_m)$ 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.

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

Such an array of *mn* elements of *F* arranged in *m* rows and *n* columns is know 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 and $m \times n$ 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 $T : V_2(R) \to V_2(R)$ 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 $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

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

$$T(e_1) = (1,1) = w_1$$

$$T(e_2) = (0,1) = (1/2)w_1 - (1/2)w_2$$

Hence the matrix representing T is $\begin{bmatrix} 1 & 0 \\ 1/2 & -1/2 \end{bmatrix}$

Solved Problems

Problem 1.

Obtain the matrix representing the linear transformation $T:V_3(R)\to V_3(R)$ given by $T(a,b,c)=(3a.a-b,2a+b+c)$ w.r.t. the standard basis $\{e_1,e_2,e_3\}$.

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 $\begin{bmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Problem 2.

Find the linear transformation $T:V_3(R)\to V_3(R)$ denoted by the matrix $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$ 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, $(a,b,c)=a(1,0,0)+b(0,1,0)+c(0,0,1)$

$$=ae_1+be_2+ce_3$$

$$\therefore T(a,b,c)=T(ae_1+be_2+ce_3)$$

$$=aT(e_1)+bT(e_2)+cT(e_1)$$

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

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

This is the required linear transformation.

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 a field F. Let $\alpha \in F$. We define $\alpha A=(\alpha a_{ij})$

Theorem.

The set $M_{m\times n}(F)$ of all $m \times n$ matrices over the field F is a vector space of dimension *mn*over F under matrix addition and scalar multiplication defined above.

Proof

Let $A=(a_{ij})$ and $B=(b_{ij})$ be two $m \times n$ matrices over a 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 are 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) $IA=A$

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 $(i,j)^{th}$ place and 0 in the other places. We have *mn*matrices of this form. Also any matrix $A=(a_{ij})$ can be written as $A=\sum a_{ij}E_{ij}$. Hence A is a linear combination of the matrices E_{ij} are linearly independent. Hence these *mn* matrices form a bases for the space of all $m \times n$ matrices. Therefore the dimension of the vector space is *mn.*

Theorem

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.

$L(V,W)$ is a vector space of dimension *mn* over F. Now, we shall prove that the vector space $L(V,W)$ is isomorphic to the vector space $M_{m\times n}(F)$ is of dimension *mn*, it follows that $L(V,W)$ is also of dimension *mn*

Fix a basis $\{v_1, v_2, \ldots, v_m\}$ for V and an ordered basis $\{w_1, w_2, \ldots, w_m\}$ for W.

We know that any linear transformation

$T \in L(V,W)$ can be represented by an $m \times n$ matrix over F.

Let 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 $M(T_2) = (b_{ij})$

$$M(T_1) = (a_{ij}) \Rightarrow T_1(v_i) = \sum_{j=1}^{n} a_{ij} w_j$$

$$M(T_2) = (b_{ij}) \Rightarrow T_2(v_i) = \sum_{j=1}^{n} b_{ij} w_j \text{ s}$$

$$\therefore (T_1 + T_2)(v_i) = \sum_{j=1}^{n} (a_{ij} b_{ij}) w_j$$

$$\therefore M(T_1 + T_2) = (a_{ij} + b_{ij})$$

$$= (a_{ij}) + (b_{ij})$$

$$= M(T_1) + M(T_2)$$

Similarly $M(\alpha T_1) = \alpha M(T_1)$

Hence M is the required isomorphism from $L(V,W)$ to $M_{m\times n}(F)$

Definition and examples

Definition. Let V be a vector space over F. An inner product of V is a function which assigns 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) $\langle u,v\rangle=\overline{\langle u,v\rangle}$, where $\overline{\langle u,v\rangle}$ is the complex conjugate of $\langle u,v\rangle$.

(iv) $\langle u,v\rangle\geq 0$ 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 an 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, $\langle u,\alpha v\rangle=\overline{\langle \alpha u,v\rangle}$

$=\overline{\alpha\langle u,v\rangle}$

$=\bar{\alpha}\overline{\langle u,v\rangle}$

$=\bar{\alpha}\langle u,v\rangle$

Note 3. $\langle u,v+w\rangle=\langle u,v\rangle+\langle v,w\rangle$

For, $\langle u,v+w\rangle=\overline{\langle v+w,u\rangle}$

$=\overline{\langle v,u\rangle+\langle w,u\rangle}$

$=\overline{\langle v,u\rangle}+\overline{\langle w,u\rangle}$

$=\langle u,v\rangle+\langle u,w\rangle$

Note 4. $\langle u,0\rangle=\langle 0,v\rangle=0$

For, $\langle u,0\rangle = \langle u,00\rangle = 0\langle u,0\rangle = 0$

Similarly $\langle 0,v\rangle = 0$.

Examples.

1. $V_n(R)$ is a real inner product space with inner product defined by

$\langle x,y\rangle = x_1y_1 + x_2y_2 + \ldots + x_ny_n$

$x = (x_1, x_2, \ldots, x_n)$ and

$y = (y_1, y_2, \ldots y_n)$

This is called the standard inner product on $V_n(R)$.

Proof.

Let $x, y, z \in V_n(R)$ and $\alpha \in R$.

(i) $\langle x+y, z\rangle = (x_1 + y_1)z_1 + (x_2 + y_2)z_2 + \ldots + (x_n + y_n)z_n$

$= (x_1z_1 + x_2z_2 + \ldots + x_nz_n) + (y_1z_1 + y_2z_2 + \ldots + y_1z_n)$

$= \langle x,z\rangle + \langle y,z\rangle$

(ii) $\langle \alpha x, y\rangle = \alpha x_1y_1 + \alpha x_2y_2 + \ldots + \alpha x_ny_n$

$= \alpha(x_1y_1 + x_2y_2 + \ldots + x_ny_n)$

$= \alpha\langle x,y\rangle$

(iii) $\langle x,y\rangle = x_1y_1 + x_2y_2 + \ldots + x_ny_n$

$= y_1x_1 + y_2x_2 + \ldots + y_nx_n$

$= \langle y,x\rangle$

(iv) $\langle x,x\rangle = x_1^2 + x_2^2 + \ldots + x_n^2 \geq 0$ and

$\langle x,x\rangle = 0$ iff $x_1^2 = x_2^2 = \ldots = x_n^2 = 0$

$\langle x,x\rangle = 0$ iff $x = 0$

2. $V_n(C)$ is a complex inner product space with inner product defined by $\langle x, y\rangle = x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n}$ where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$

Proof.

Let $x, y, z \in V_n(C)$ and $\alpha \in C$

(i) $\langle x+y, z\rangle = (x_1 + y_1)\overline{z_1} + (x_2 + y_2)\overline{z_2} + \ldots + (x_n + y_n)\overline{z_n}$

$= \left(x_1\overline{z_1} + x_2\overline{z_2} + \ldots + x_n\overline{z_n}\right) + \left(y_1\overline{z_1} + y_2\overline{z_2} + \ldots + y_1\overline{z_n}\right)$

$= \langle x, z\rangle + \langle y, z\rangle$

(ii) $\langle \alpha x, y\rangle = \alpha x_1\overline{y_1} + \alpha x_2\overline{y_2} + \ldots + \alpha x_n\overline{y_n}$

$= \alpha\left(x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n}\right)$

$= \alpha\langle x, y\rangle$

(iii) $\overline{\langle y, x\rangle} = \overline{y_1\overline{x_1} + y_2\overline{x_2} + \ldots + y_n\overline{x_n}}$

$= \overline{y_1}x_1 + \overline{y_2}x_2 + \ldots + \overline{y_n}x_n$

$= x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n}$

$= \langle x, y\rangle$

(iv) $\langle x, x\rangle = x_1\overline{x_1} + x_2\overline{x_2} + \ldots + x_n\overline{x_n}$

$= |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 \geq 0$ and

$\langle x, x\rangle = 0$ iff $x_1^2 = x_2^2 = \ldots = x_n^2 = 0$

$\langle x, x\rangle = 0$ iff $x = 0$

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 $\alpha \in R$

(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) $\langle \alpha f,g\rangle = \int_0^1 \alpha f(t)g(t)dt$

$= \alpha \int_0^1 f(t)g(t)dt$

$= \alpha\langle g,h\rangle$

(iii) $\langle f,g\rangle = \int_0^1 f(t)g(t)dt$

$= \int_0^1 g(t)f(t)dt$

$= \langle g,h\rangle$

(iv) $\langle f,f\rangle = \int_0^1 [f(t)]^2 dt \geq 0$ and

$\langle f,f\rangle = 0$ iff $f = 0$

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) $\|f\|^2 = \langle f, f\rangle$

$= \int_0^1 [f(t)]^2 dt$

$= \int_0^1 (t+2)^2 dt$

$= \int_0^1 (t^3-7t-6)dt$

$= \int_0^1 (t^3+4t+4)dt$

$= \left[\frac{t^3}{3} + 2t^2 + 4t\right]_0^1$

$= \frac{1}{3} + 2 + 4$

$= \frac{19}{3}$

$\|f\| = \frac{\sqrt{19}}{\sqrt{3}}$

Theorem. 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) $\|x + y\| = \|x\| + \|y\|$ (Triangle inequality).

Proof.

(i) $\|x\| = \sqrt{\langle x, x\rangle} \geq 0$ and $\|x\| = 0$ iff $x = 0$.

(ii) $\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle$

$= \alpha \langle x, \alpha x\rangle$

$= \alpha \overline{\alpha} \langle x, x\rangle$

$= |\alpha|^2 \|x\|^2$

$\|\alpha x\| = |\alpha| \|x\|$

(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 $0 \leq \langle z, z\rangle$

$$=\left\langle y-\frac{\langle y,x\rangle}{\|x\|^2}x,\, y-\frac{\langle y,x\rangle}{\|x\|^2}x\right\rangle$$

$$=\langle y,y\rangle-\frac{\overline{\langle y,x\rangle}}{\|x\|^2}\langle y,x\rangle-\frac{\langle y,x\rangle}{\|x\|^2}\langle x,y\rangle+\frac{\langle y,x\rangle\overline{\langle y,x\rangle}}{\|x\|^2\|x\|^2}\langle x,x\rangle$$

$$=\|y^2\|-\frac{\overline{\langle y,x\rangle}\langle y,x\rangle}{\|x\|^2}-\frac{\langle y,x\rangle\langle x,y\rangle}{\|x\|^2}+\frac{\langle y,x\rangle\overline{\langle y,x\rangle}}{\|x\|^2}$$

$$=\|y^2\|-\frac{\langle x,y\rangle\overline{\langle x,y\rangle}}{\|x\|^2}$$

$$\therefore 0\le \|x\|^2\|y\|^2-|\langle x,y\rangle|^2$$

$$\therefore |\langle x,y\rangle|\le \|x\|\ \|y\|$$

(iv) $\|x+y\|^2=\langle x+y,x+y\rangle$

$$=\langle x,x\rangle+\langle x,y\rangle+\langle y,x\rangle+\langle y,y\rangle$$

$$=\|x\|^2+\langle x,y\rangle+\overline{\langle x,y\rangle}+\|y\|^2$$

$$=\|x\|^2+2\,\mathrm{Re}\langle x,y\rangle+\|y\|^2$$

$$\le\|x\|^2+2|\langle x,y\rangle|+\|y\|^2$$

$$\le^2 +2\|x\|\|y\|+\|y\|^2$$

$$\le(\|x\|+\|y\|)^2$$

$$\therefore\|x+y\|=\|x\|+\|y\|$$

Orthogonality

Definition. Let V be an inner product space and $x, y\in V$ let x is said to be ***orthogonal*** to *y*if $\langle x,y\rangle=0$

Note 1. *x*is orthogonal to $y\Rightarrow\langle x,y\rangle=0$

$\Rightarrow \overline{\langle x, y \rangle} = \bar{0}$

$\Rightarrow \langle y, x \rangle = 0$

$\Rightarrow$ *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 orthogonal to *y*

Note 4. 0 is orthogonal to every vector in V 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 n 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 $\{e_1, e_2, \ldots e_n\}$ in R^n or C^n is an orthogonal set with respect to the standard inner product.

Theorem. Let $S = \{v_1, v_2, \ldots, v_n\}$ be an orthogonal set of non zero vectors in an inner product space V. then S is linearly independent.

Proof.

Let $\alpha_1 v_1, \alpha_2 v_2, \ldots \alpha_n v_n = 0$

Then $\langle \alpha_1 v_1, \alpha_2 v_2, \ldots, \alpha_n v_n, v_1 \rangle = \langle 0, v_1 \rangle = 0$

$\therefore \alpha_1 \langle v_1, v_1 \rangle + \alpha_2 \langle v_2, v_1 \rangle + \ldots + \alpha_n \langle v_n, v_1 \rangle = 0$

$\therefore \alpha_1 \langle v_1, v_1 \rangle = 0$ (since S is orthogonal)

$\therefore \alpha_1 = 0$ (since $v_1 \neq 0$)

Similarly $\alpha_2 = \alpha_3 = \ldots = \alpha_n = 0$

Hence S is linearly independent.

Theorem. Let $S=\{v_1,v_2,\ldots,v_n\}$ be an orthogonal set of non zero vectors in an inner product space V. let $v\in V$ and $v=\alpha_1 v_1,\alpha_2 v_2,\ldots,\alpha_n v_n$. Then $\alpha_k = \frac{\langle v, v_k\rangle}{\|v_k\|^2}$

Proof. $\langle v,v_k\rangle=\langle \alpha_1 v_1,\alpha_2 v_2,\ldots,\alpha_n v_n,v_k\rangle$

$$=\alpha_1\langle v_1,v_k\rangle+\alpha_2\langle v_2,v_k\rangle+\ldots+\alpha_k\langle v_k,v_k\rangle+\ldots+\alpha_n\langle v_n,v_k\rangle$$

$=\alpha_k\langle v_k,v_k\rangle$ (since S is orthogonal)

$$=\alpha_k\|v_k\|^2$$

$$\therefore \alpha_k = \frac{\langle v, v_k\rangle}{\|v_k\|^2}$$

Theorem. Every finite dimensional inner product space has an orthonormal basis

Proof.

Let V be a finite dimensional inner product space. Let $\{v_1,v_2,\ldots v_n\}$ be a basis for V. From this basis we shall construct an orthonormal basis $\{w_1,w_2,\ldots w_n\}$ by means of a construction know 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 $\{w_1,w_2,\ldots,w_k\}$. Then put

$$w_{k+1}=v_{k+1}-\sum_{j=1}^{k}\frac{\langle v_{k+1},w_j\rangle}{\|w_k\|^2}w_k$$

We claim that $w_{k+1}\neq 0$. For, if $w_{k+1}=0$ then v_{k+1} is a linear combination of $\{w_1,w_2,\ldots,w_k\}$ and hence is a linear combination of $\{v_1,v_2,\ldots,v_k\}$ which is a contradiction since $\{v_1,v_2,\ldots,v_{k+1}\}$ are linearly independent

Also

$$\langle w_{k+1},w_1\rangle=\langle v_{k+1},w_1\rangle-\sum_{j=1}^{k}\frac{\langle v_{k+1},w_j\rangle}{\|w_j\|^2}\langle w_j,w_i\rangle$$

$$=\langle v_{k+1},w_i\rangle-\frac{\langle v_{k+1},w_i\rangle}{\|w_i\|^2}\langle w_i,w_i\rangle$$

$$=\langle v_{k+1},w_i\rangle-\langle v_{k+1},w_i\rangle$$

$$=0$$

Thus, continuing in this way we ultimately obtain a non zero orthogonal set $\{w_1,w_2,\ldots,w_n\}$

By theorem this set is linearly independent and hence a basis

To obtain an orthonormal basis we replace each w_i by $\frac{w_i}{\|w_i\|}$

Solved Problems

Problem 1. Apply Gram-Schmidt orthogonalisation process to construct an orthonormal basis for $V_3(R)$ with the standard inner 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 $w_2 = v_2 - \dfrac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1$

$= (1,3,1) - (1,0,1)$

$= (0,3,0)$

$\|w_2\|^2 = 9$

Also, $\langle w_2, w_3 \rangle = 0 + 6 + 0 = 6$ and $\langle w_1, v_3 \rangle = 3 + 0 + 1 = 4$

Now, $w_3 = v_3 - \dfrac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \dfrac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2$

$= (3,2,1) - \dfrac{4}{2}(1,0,1) - \dfrac{6}{9}(0,3,0)$

$= (3,2,1) - 2(1,0,1) - \dfrac{2}{3}(0,3,0)$

$= (1,0,-1)$

$\therefore \|w_3\|^2 = 2$

$\therefore$ The orthogonal basis is $\{(1,0,1), (0,3,0), (1,0,-1)\}$

Hence the orthonormal basis is

$\left\{ \left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{1}{\sqrt{2}} \right), (0,1,0), \left(\dfrac{1}{\sqrt{2}}, 0, \dfrac{-1}{\sqrt{2}} \right) \right\}$

Problem 2. Let V be the set of all polynominals of degree ≤ 2 together with the zero polynomial. V is a real inner product space with inner product defined by $\langle f, g \rangle = \int_{-1}^{1} f(x) g(x) dx$.

Starting with the basis $\{1, x, x^2\}$, obtain an orthonormal basis for V.

Solution.

Let $v_1 = 1; v_2 = x$ and $v_3 = x$

Let $w_1 = v_1$

Then $\|w_1\|^2 = \langle w_1, w_1 \rangle = \int_{-1}^{1} 1dx = 2$

$\|w_1\| = \sqrt{2}$

$$w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2$$

$$= x^2 - \frac{1}{2}\int_{-1}^{1} x^2 dx - \left[\frac{3x}{2}\right]\int_{-1}^{1} x^3 dx$$

$$= x^2 - \frac{1}{3}$$

$$\|w_3\|^2 = \langle w_3, w_3 \rangle = \int_{-1}^{1} \left(x^2 - \frac{1}{3}\right)^2 dx = \frac{8}{45}$$

Hence the orthogonal basis is $\left\{1, x, x^2 - \frac{1}{3}\right\}$

The required orthonormal basis is $\left\{\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{2}x, \frac{\sqrt{10}}{4}(3x^2 - 1)\right\}$

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) $S^{\perp} = \{x / x \in V \; and \; \langle x, u \rangle = 0 \;\; for\, all\, u \in S\}$

Examples

1. $V^{\perp} = \{0\}$ and $\{0\}^{\perp} = V$ since 0 is the only vector which is orthogonal to every vector

2. Let $S = \{(x.0.0) / x \in R\} \subseteq V_3(R)$ with standard inner product. Then $S^{\perp} = \{(0, y, z) / y, z \in R\}$

(i.e) The orthogonal complement of the *x*-axis is the *yz* plane

Theorem

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

Let *V* be a finite dimensional inner product space. Let *W* be a subspace of *S*. Then *V* is the direct sum of *W* and $W^{\perp}$ (i.e) $V = W \oplus W^{\perp}$

Proof.

(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 every 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 $\{v_1, v_2, ..., v_r\}$ be an orthonormal bases for W. Let $v \in V$

Consider $v_0 \in v - \langle v, v_1 \rangle v_1 - \langle v, v_2 \rangle v_2 - ... - \langle v, v_r \rangle v_r$

$\therefore \langle v_0, v_i \rangle = \langle v, v_i \rangle - \langle v, v_i \rangle \langle v_i, v_i \rangle$ (since $\langle v_i, v_j \rangle = 0$ if $i \neq j$

$= \langle v, v_i \rangle - \langle v, v_i \rangle$ (since $\langle v_i, v_j \rangle = 1$)

$= 0$

$\therefore v_0$ is orthogonal to each of $\{v_1, v_2, ..., v_r\}$ and hence is orthogonal to every vector in W. Hence $v_0 \in W^{\perp}$ and $v = [\langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + ... + \langle v, v_r \rangle v_r] + v_0 \in W + W^{\perp}$ $V = W \oplus W^{\perp}$

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. Let V be a finite dimensional inner product space. Let W be a subspace of V. Then $(W^{\perp})^{\perp} = W$

Proof.

Let $w \in W$. Then for any $u \in W^{\perp}$, $\langle w,u \rangle = 0$

Hence $w \in (W^{\perp})^{\perp}$. Thus $W \subseteq (W^{\perp})^{\perp}$ **...(1)**

Now by theorem $V = W \oplus W^{\perp}$

Also $V = W^{\perp} \oplus (W^{\perp})^{\perp}$

Hence $\dim W = \dim(W^{\perp})^{\perp}$ **...(2)**

From (1) and (2) we get $(W^{\perp})^{\perp} = W$

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 $u \in S_2$

But $S_1 \subseteq S_2$. Hence $\langle u,v \rangle = 0$ for all $u \in S_1$

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 $W_1 \in W_1 + W_2$

$\therefore \ (W_1+W_2)^\perp \subseteq W_1^\perp$ (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 $v_2 \in W_2$

$\therefore \ \langle w,u \rangle = \langle w, v_1 + v_2 \rangle$

$= \langle w, v_1 \rangle + \langle w, v_2 \rangle$

=0 + 0 (since $v_1 \in W_1$ and $v_2 \in W_2$) = 0

Hence $w \in (W_1+W_2)^\perp$

$W_1^\perp \cap W_2^\perp \in (W_1+W_2)^\perp$ ....(2)

From (10 and (2) we get

$(W_1+W_2)^\perp = W_1^\perp \cap W_2^\perp$

(ii) Proof is similar to that of (i)

UNIT - IV
THEORY OF MATRICES

Introduction

In this chapter we shall develop the general theory of matrices. Throughout this chapter we deal with matrices whose entries are from the field F of real or complex numbers.

Algebra of Matrices

We have already seen that an$m \times n$ matrix A is an arrat of mn numbers a_{ij} where $1 \leq i \leq m, 1 \leq j \leq n$ arranged in *m* rows and *n* columns as follows

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

We shall denote this matrix by the symbol (a_{ij}). If *m*=*n*, A is called a square matrix of order *n*

Definition. Two matrices A=(a_{ij}) and B=(b_{ij}) are said to be equal if A and B have the same number of rows and columns and the corresponding entries in the two matrices are same.

Additional of matrices. We have already defined the addition of two $m \times n$ matrix A=a_{ij} and B=(b_{ij}) by A+B=$(a_{ij} + b_{ij})$

We note that we can add two matrices iff they have the same number of rows and columns.

Example. If A= $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 9 & 5 \end{bmatrix}$ and B= $\begin{bmatrix} 0 & 4 \\ 2 & 1 \\ -1 & 0 \end{bmatrix}$ then A+B=$\begin{bmatrix} 1 & 6 \\ 5 & 5 \\ 8 & 5 \end{bmatrix}$

Remark. The set of all $m \times n$ matrices is an abelian group under matrix addition. The $m \times n$ matrix with each entry 0 is the zero matrix and is denoted by 0 and the additive inverse of matrix A=(a_{ij}) is $(-a_{ij})$ and is denoted by $-A$

If A=$a_{ij}$is any matrix and α is any number (real or complex) we have defined the matrix αA by $\alpha A = (\alpha a_{ij})$

The set of all $m \times n$ matrices over the field R under matrix addition and scalar multiplication defined above is a vector space. This result is true if R is replaced by C or by any field F

We now proceed to define multiplication of matrices. We have already defined the multiplication of 2×2 matrices, which we generalise in the following definition

Definition. Let A=a_{ij} be an$m \times n$ matrix and B=(b_{ij}) be an $n \times p$ matrix. We define the product AB as the $m \times p$ matrix (c_{ij}) where the ij^{th} entry (c_{ij}) is given by

$$c_{ij} = a_{i1}b_{1i} + a_{i2}b_{2i} + \cdots + a_{in}b_{ni} = \sum_{k=1}^{n} a_{ik}b_{ki}$$

Note 1. The product AB of two matrices is defined only when the number of columns of A is equal to the number of rows of B.

Note 2. The entry c_{ij} of the product AB is found by multiplying i^{th} row of A and the j^{th} column of B. To multiply a row and a column, we multiply the corresponding entries and add.

Solved Problems

Problem 1. Show that the matrix $A = \begin{bmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{bmatrix}$ satisfies the equation $A(A-I)(A+2I) = 0$

Solution

$$A - I = \begin{bmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -3 & 1 \\ 3 & 0 & 3 \\ -5 & 2 & -5 \end{bmatrix}$$

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

Now

$$A(A-I)(A+2I) = \begin{bmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{bmatrix} \begin{bmatrix} 1 & -3 & 1 \\ 3 & 0 & 3 \\ -5 & 2 & -5 \end{bmatrix} \begin{bmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} -12 & -4 & -12 \\ -9 & -3 & -9 \\ 21 & 7 & 21 \end{bmatrix} \begin{bmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = 0$$

$A(A-I)(A+2I) = 0$

Problem 2.

Prove that $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^{n} = \begin{bmatrix} \lambda^{n} & n\lambda^{n-1} \\ 0 & \lambda^{n} \end{bmatrix}$

Solution. We prove this result by induction of n. when $n = 1$ result is obviously true. Let us assume that the result is true for $n = k$

$$\therefore \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^{k} = \begin{bmatrix} \lambda^{k} & k\lambda^{k-1} \\ 0 & \lambda^{k} \end{bmatrix}$$

$$\therefore \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^{k} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^{k} & k\lambda^{k-1} \\ 0 & \lambda^{k} \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

$$= \begin{bmatrix} \lambda^{k+1} & \lambda^k + k\lambda^k \\ 0 & \lambda^{k+1} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda^{k+1} & (k+1)\lambda^k \\ 0 & \lambda^{k+1} \end{bmatrix}$$

∴The result is true for $n = k + 1$

Hence the result is true for all positive integers n.

Definition. Let A=(a_{ij}) be an$m \times n$ matrix. Then the $n \times m$ matrix B=(b_{ij}) where $b_{ij} = a_{ij}$ is called the transpose of the matrix A and it is denoted by A^T. Thus A^T is obtained from the matrix A by interchanging its rows and columns and the (ij^{th}) entry of $A^T = (ji^{th})$ entry of A.

For example, if $A = \begin{bmatrix} 1234 \\ 2101 \\ 0315 \end{bmatrix}$ then $A^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 3 & 0 & 1 \\ 4 & 1 & 5 \end{bmatrix}$ clearly if A is an$m \times n$ matrix. Then the $n \times m$ matrix

Theorem. Let A and B be two $m \times n$ matrices. Then

(i) $(A^T)^T = A$

(ii) $(A + B)^T = A^T + B^T$

Proof.

(i) The (ij^{th}) entry of $(A^T)^T$

=(ij^{th}) entry of A^T

$=$ (ij^{th})entry of A

∴ $(A^T)^T = A$

(ii) The (ij^{th}) entry of $(A + B)^T$

=(ji^{th}) entry of $A + B$

=(ji^{th}) entry of A + (ji^{th}) entry of B

=(ij^{th}) entry of A^T + (ji^{th}) entry of B^T

=(ij^{th}) entry of $A^T + B^T$

∴$(A + B)^T = A^T + B^T$

Theorem.

Let A be an$m \times n$ matrix and B be an $n \times p$ matrix. Then $(AB)^T = B^T A^T$

Proof.

By hypothesis AB is defined and it is an$m \times p$ matrix. Hence $(AB)^T$is a $p \times m$ matrix

Further B^T is a $p \times n$ matrix and A^T is an$n \times m$ matrix

Hence, the product $B^T A^T$ is defined and it is a $p \times m$ matrix.

Now, let A=(a_{ij}), B=(b_{ij}) and $(AB) = (c_{ij})$

Then $(i,j)^{th}$ entry of

$$(AB) = (c_{ij}) = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$(AB)^T = (c_{ji}) = \sum_{k=1}^{n} a_{jk} b_{ki}$$

Now the i^{th} row of B^T is the i^{th} column of B and it consists of the elements $b_{1i}, b_{2i}, \dots b_{ni}$. Also the j^{th} column of B^T is the j^{th} row of A and it consists of the elements $a_{1i}, a_{2i}, \dots a_{ni}$

$$= \sum_{k=1}^{n} b_{ki} a_{jk}$$

$= (i,j)^{th}$ entry of $(AB)^T$

Hence $(AB)^T = B^T A^T$

Definition. Let A=(a_{ij}) be a matrix with entries from the field of complex numbers. The conjugate of A, denoted by $\bar{A}$, is defined by $\bar{A} = \overline{(a_{ij})}$.

$\bar{A}^T$is called the conjugate transpose of the matrix A.

For example if $A = \begin{bmatrix} 2 & 2+i & -i \\ 1+i & -3 & 4+3i \end{bmatrix}$ then $\bar{A} = \begin{bmatrix} 2 & 2-i & i \\ 1-i & -3 & 4-3i \end{bmatrix}$

Theorem. Let A and B matrices with entries from C. Then

(i) $\overline{(\bar{A})} = A$.

(ii) $\overline{A+B} = \bar{A} + \bar{B}$

(iii) $\overline{kA} = \bar{k}\bar{A}$, where $k \in C$.

(iv) $A = \bar{A} \Leftrightarrow$all entries of A are real

(v) $\overline{AB} = \bar{A}\bar{B}$

(vi) $(\bar{A})^T = \overline{A^T}$

The proof of the above results are immediate consequences of the corresponding properties of complex numbers.

Types of Matrices

Definition. An$1 \times n$ matrix is called a row matrix. Thus a row matrix is consists of 1 row and columns.

It is of the form $(a_{11}, a_{12}, \dots a_{1n})$

Definition. An$m \times 1$ matrix is called a column matrix. Thus a column matrix is consists of 1 column and rows.

It is of the form $\begin{bmatrix} a_{11} \\ a_{12} \\ \cdot \\ \cdot \\ \cdot \\ a_{1n} \end{bmatrix}$

Definition. Let A=(a_{ij}) be a square matrix. Then the elements $(a_{11}, a_{22}, \dots a_{nn})$ are called the diagonal elements of A and the diagonal elements constitute what is known as the principal diagonal of the matrix A. A square matrix is called a diagonal matrix if all the entries which do not belong to the principal are zero. Hence in a diagonal matrix $a_{ij} = 0$ if $i \neq j$

For example $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is a diagonal matrix

Definition. A diagonal matrix in which all the entries of the principal diagonal are equal is called a scalar matrix

For example $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ is a scalar matrix

Definition. A square matrix (a_{ij}) is called an upper triangular matrix if all the entries above the principal diagonal are zero

Hence $a_{ij} = 0$ whenever $i < j$ is an upper triangle matrix.

Definition. A square matrix (a_{ij}) is called a lower triangle matrix if all the entries below the principal diagonal are zero

Hence $a_{ij} = 0$ whenever $i > j$ in an lower triangular matrix

For example $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ is an lower triangular matrix $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 2 & 3 & 0 \\ 2 & 3 & 2 & 4 \end{bmatrix}$ is upper triangular

Clearly a square matrix is a diagonal matrix iff it is both lower triangular and upper triangular.

Definition. A square matrix $A = (a_{ij})$ is said to be symmetric if $a_{ij} = a_{ji}$ for all i, j

Example.

$\begin{bmatrix} a & b \\ b & a \end{bmatrix}$, $\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$, $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 5 \\ 3 & 0 & 6 & 7 \\ 4 & 5 & 7 & 8 \end{bmatrix}$ are symmetric matrices.

Theorem. A square matrix A is symmetric iff$A = A^T$

Proof. Let A be a symmetric matrix

Then the $(i,j)^{th}$ entry of A

$= (j,i)^{th}$ entry of A

$= (i,j)^{th}$ entry of A^T

Hence $A = A^T$

Conversely let $A = A^T$

Then $(i,j)^{th}$ entry of A

$= (i,j)^{th}$ entry of A^T

$= (j,i)^{th}$ entry of A

Hence A is symmetric

Theorem. Let A be any square matrix. Then $A + A^T$ is symmetric

Proof. $(A + A^T)^T = A^T + (A^T)^T$

$= A^T + A$

$= A + A^T$

Hence $A + A^T$ is symmetric

Theorem. Let A and B be symmetric matrices of order n. Then

(i) $A + B \text{ is symmetric}$

(ii) AB is symmetric iff $AB = B$

(iii) $AB + BA$ is symmetric

(iv) If A is symmetric, then kA is symmetric where $k \in F$.

Proof.

(i) $(A + B)^T = A^T + B^T$

$= A + B$ (since A and B are symmentric)

$\therefore$ $A + B$is symmetric

(ii) AB is symmetric

$$\Leftrightarrow (AB)^T = AB$$

$$\Leftrightarrow B^T A^T = AB$$

$$\Leftrightarrow BA = A$$

(iii) $(AB + B\ \ )^T = (AB)^T + (BA)^T$

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

$= BA + A$ (since A and B are symmetric)

$= AB + B$

$\therefore$ $AB + BA$is symmetric

(iv) $(kA)^T = kA^T = k$ since A is symmetric

$\therefore$ $kA$is symmetric

Definition. A square matrix $A = (a_{ij})$ is said to be skew symmetric if $a_{ij} = -a_{ji}$, for all i, j

Note. Let A be a skew symmetric matrix. Then $a_{ij} = -a_{ji}$. Hence $2a_{ij} = 0$ (ie) $a_{ij} = 0$, for all i. Thus in a skew symmetric matrix all the diagonal entries are zero

$\begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix}, \begin{bmatrix} 0 & -2 & 1 \\ 2 & 0 & -3 \\ -1 & 3 & 0 \end{bmatrix}$ Are examples of skew symmetric matrices

Theorem. A square matrix A is skew symmetric matrix iff$A = -A^T$

Proof is similar to that of by theorem

Theorem. Let A be any square metrix. Then $A - A^T$ is skew symmetric

Proof.

$(A = -A^T)^T = A^T - (A^T)^T$

$= A^T - A$

$= -(A^T - A)$

Hence $A - A^T$ is skew symmetric

Theorem. Any square matrix A can be expressed uniquely as the sum of a symmetric matrix and a skew symmetric matrix.

Proof. Let A be any square matrix

Then $A + A^T$ is skew symmetric matrix (by Theorem)

$\therefore \frac{1}{2}(A + A^T)$is also a symmetric matrix

Also $\frac{1}{2}(A - A^T)$ is also a symmetric matrix (by above theorem)

Now, $A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T)$

$\therefore$A is the sum of a symmetric matrix and a skew symmetric matrix

Now, to prove the uniqueness, let $A = R + S$ where S is a symmetric matrix and R is a skew symmetric matrix. We claim that $S = \frac{1}{2}(A + A^T)$ and $R = \frac{1}{2}(A - A^T)$

$A = S + R \ldots.(i)$

$\therefore A^T = (S + R)^T$

$= S^T + R^T$

$= S - R$ (since S is symmetric and R is skew symmetric)

$\therefore A^T = S - R \ldots (ii)$

From (i) and (ii) we get $S = \frac{1}{2}(A + A^T)$ and $R = \frac{1}{2}(A - A^T)$

Theorem. Let A and B be skew symmetric matrices of order n. Then

(i) A+B is skew s$ymmetric$

(ii) kA is skew symmetric where $k \in F$

(iii) A^{2n} is a symmetric matrix and A^{2n+1} is a skew symmetric matrix where n is any positive integer.

Proof.

Let A, B be skew symmetric

(i) $(A + B)$

$= -A - B$

$= -(A + B)$

$\therefore A + B \text{ is skew symmetric}$

(ii) Proof is similar to that of (i)

(iii) Let m be any positive integer

Then $(A^m)^T = (AA \ldots m \text{ times})^T$

$= A^T A^T \ldots A^T (m \text{ times})$

$=(-A)(-A) \ldots (-A)(m \text{ times})$ (since $A^T = -A$)

$= (-1)^m A^m$

$$(A^m)^T = \begin{cases} A^m \text{ if } m \text{ is even} \\ -A^m \text{ if } m \text{ is odd} \end{cases}$$

$A^m$is symmetric when m is even and skew symmetric when m is odd

Definition. A square matrix $A = (a_{ij})$ is said to be Hermitian matrix if$a_{ij} = -a_{ij}$ for all i, j. A is said to be a skew Hermitian matrix iff $a_{ij} = -\overline{a_{ij}}$ for all i, j.

Note

1. Any hermitian matrix over R is a symmetric matrix and any skew Hermitian matrix over R is a skew symmetric matrix.
2. Let $A = (a_{ij})$ be a hermitian matrix. Then $a_{ii} = -\overline{a_{ii}}$ and hence a_{ii} is real for all i.
3. Let $A = (a_{ij})$ be a skew hermitian matrix. Then $a_{ii} = -\overline{a_{ii}}$ and hence $a_{ii} = 0$ or purely imaginary for all i.

Theorem. Let A be a square matrix

(i) A is Hermitianiff$A = \bar{A}^T$

(ii) A is skew Hermitianiff$A = -\bar{A}^T$

Proof. The result is an immediate consequence of the definition

Theorem. Let A and B be square matrices of the same order. Then

(i) A, B are Hermitian $\Rightarrow A + B$ is Hermitian

(ii) A, B are skew Hermitian $\Rightarrow A + B$ is skew Hermitian

(iii) A is Hermitian $\Rightarrow iA$ is Hermitian

(iv) A is skew Hermitian $\Rightarrow iA$ is skew Hermitian

(v) A is Hermitian and k is real $\Rightarrow kA$ is Hermitian

(vi) A is skew Hermitian and k is real $\Rightarrow kA$ is skew Hermitian

(vii) A, B are Hermitian $\Rightarrow AB + BA$ is Hermitian

(viii) A, B are Hermitian $\Rightarrow AB - BA$ is Hermitian

Proof. We shall prove (i), (iii) and (vii)

(i) $\overline{(A + B)}^T = (\bar{A} + \bar{B})^T$

$= \bar{A} + \bar{B}$

$= A + B$ (since A and B are Hermitian)

$\therefore A + B$is Hermitian

(ii) $\overline{-(\iota A)}^T =-(\overline{\iota\ \ })^T$

$= \bar{A}^T$

$= iA$ (since A is Hermitian)

$\therefore iA$is skew Hermitian

(vii) $\overline{(AB+BA)}^T = (\overline{AB}+\overline{BA})^T$

$$= (\bar{A}\bar{B}+\bar{B}\bar{A})^T$$

$$= (\bar{A}\bar{B})^T + (\bar{B}\bar{A})^T$$

$$= \bar{B}^T\bar{A}^T + \bar{A}^T\bar{B}^T$$

$$= BA + AB$$

$$= AB + B$$

$\therefore AB + B$ is Hermitian

Theorem. Let A be any square matrix. Then

(i) $A+\bar{A}^T$ is Hermitian

(ii) $A-\bar{A}^T$ is skew Hermitian

Proof.

(i) Let $A+\bar{A}^T = B$

$$\overline{B} = \bar{A} + A^T$$

$$\therefore \overline{B}^T = \bar{A} + A^{T^T}$$
$$= \bar{A}^T + A$$

(ii) Proof is similar to that of (i)

Theorem. Any square matrix A can be uniquely expressed as the sum of a Hermitian matrix and a skew Hermitian matrix.

Proof.

The proof is similar to that of the Theorem

Definition. A real square matrix A is said to be orthogonal if $AA^T = A^TA = I$

Example

$A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$is an orthogonal matrix (verify).

Theorem. Let A and B be orthogonal matrices of the same order. Then

(i) A^T is orthogonal

(ii) AB is orthogonal

Proof

(i) $A^T(A^T)^T = A^T A = I$ (since A is orthogonal)

similarly we can prove $(A^T)^T A^T = I$

$\therefore A^T$is orthogonal

(ii) $(AB)(AB)^T = (AB)(B^T A^T)$

$$= A(BB^T)A^T$$

$$= AIA^T$$

$$= AA^T$$

$$= I$$

Similarly $(AB)(AB)^T = I$

Hence AB is orthogonal

Definition. A square matrix A is said to be an unitary matrix if $A\bar{A}^T = \bar{A}^T A = I$

For example $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ is unitary.

Note. Any matrix over R is an orthogonal matrix

Theorem. If A and B are unitary matrices of the same order, then AB is also an unitary matrix

Proof. Similar to the proof of (ii) of the above theorem

The Inverse of a Matrix.

A 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse iff $|A| = ad - b$ 0 and the inverse of A is given by $\frac{1}{|A|}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Such matrices are called non-singular. In this section we shall describe the method of finding the inverse of any non-singular matrix of order n.

Determinants. We can associate with any $n \times n$ matrix $A = (a_{ij})$ over a field F an element of F given by the determinant $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$

If value can be determined in the usual way and it is denoted by $|A|$

For example

(i) If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then $|A| = ad - bc$

(ii) If $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{bmatrix}$ then $A = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = 1$

Definition. A square matrix A is said to be singular if $|A| = 0$

A is called a non singular matrix if $|A| \neq 0$

Theorem. The rule of multiplying two matrices is same as the rule for multiplying two determinants.

Hence if A and B are two $n \times n$ matrices. $|AB| = |A||B|$.

Theorem. The product of any two non-singular matrices is non-singular.

Proof. Let A and B be two non-singular matrices of the same order. Then $|A| \neq 0$ and $|B| \neq 0$

$\therefore |AB| = |A||B|$

Hence AB is non singular matrix.

Note. Sum of two non-singular matrices need not be non-singular. For, if A is non-singular matrix then $-A$ is also a non-singular matrix and $A + (-A)$ is the zero matrix which is obviously a singular matrix

Definition. Let $A = (a_{ij})$ be an$n \times n$ metrix. If we delete the row and the column containing the element (a_{ij}) we obtain a square matrix of order $n - 1$ and the determinant of this square matrix is called the minor of the element (a_{ij}) and is denoted by (M_{ij})

The minor M_{ij} multiplied by $(-1)^{i+j}$ is called the cofactor of the element a_{ij} and is denoted by A_{ij}

$\therefore A_{ij} = (-1)^{i+j} M_{ij}$

Example. Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Corresponding to the 9 elements a_{ij}, we get 9 minors of A. For example, the minor of $a_{11}$is

$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$and the minor of a_{23} is $M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}$

The cofactor of a_{11} is $A_{11} = (-1)^2 M_{11} = M_{11}$

The cofactor of a_{23} is $A_{23} = (-1)^{2+3} M_{23} = -M_{23}$

Definition. Let $A = (a_{ij})$ be a square matrix. Let A_{ij} denote the co-factor of a_{ij}. The transpose of the matrix A_{ij} is called the adjoint to adjugate of the matrix A and is denoted be $adjA$

Thus the $(i,j)^{th}$ entry of $adjA$ is A_{ji}

Note. If A is a square matrix of order n then $adjA$ is also a square matrix of order n.

Example. Let $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix}$

Then $A_{11} = \begin{vmatrix} 1 & -1 \\ 1 & 3 \end{vmatrix} = 4$

$A_{12} = \begin{vmatrix} 3 & -1 \\ -2 & 3 \end{vmatrix} = -7$

Similarly other co-factors can be calculated and we get

$$adjA = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = \begin{bmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{bmatrix}$$

We notice that

$AadjA = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix} = (adjA)A$ (verify)

Theorem. Let A be any square matrix of order n. Then $(adjA)A = A(adjA) = |A|I$ where I is the identity matrix of order n.

Proof. The $(i,j)^{th}$ element of $\left(A(adjA)\right)$

$= \sum_{k=1}^{n} a_{ik} A_{jk}$

$= \begin{cases} 0 \; if \; i \neq j \\ |A| \; if \; i = j \end{cases}$

$\therefore A(adjA) = \begin{bmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & a_{n2} & \dots & |A| \end{bmatrix} = |A|I$

Similarly $(adjA)A = |A|I$

Hence $(adjA)A = A(adjA) = |A|I$

Note. Suppose $|A| \neq 0$. Now, consider the matrix $B = \frac{1}{|A|} \, adj \; A$

Then $AB = A \; \left[\frac{1}{|A|} \, adj \; A\right]$

$= \frac{1}{|A|}(A \; adj \; A)$

$= \frac{1}{|A|}|A|I$

$= I$

Similarly $BA = I$. Thus $AB = BA = I$

Definition. Let A be a square matrix of order n. A is said to be invertible in there exists a square matrix B of order n such that $AB = BA = I$ and B is called the inverse of A and is denoted by A^{-1}

Note. The invertible matrices are precisely the units of the ring $M_n(F)$

Theorem. A square matrix A of order n is non singulariff A is invertible

Proof. Suppose A is invertible.

Then there exists a matrix B such that $AB = BA = I$

Hence $|AB| = |I| = 1$

$\therefore |A||B| = 1$

Hence $|A| \neq 0$ so that A is non-singular.

Conversely, let A be non-singular. Hence $|A| \neq 0$

Now, consider the matrix $B = \frac{1}{|A|}\, adj\ A$

Then $AB = BA$ (refer the above Note)

∴ A is invertible and A is the inverse of A.

Solved problem

Problem1. Compute the inverse of the matrix $A = \begin{bmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix}$

Solution.

$$|A| = \begin{vmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{vmatrix} = -1$$

Since $|A| \neq 0$, A is non-singular

Hence A^{-1} exist and is given by $A^{-1} = \frac{adj\ A}{|A|}$

Now, we find $adjA = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$ where $A_{ij}, (i,j = 1,2,3)$ are cofactors ofa_{ij}

$$A_{11} = \begin{vmatrix} 6 & -5 \\ -2 & 2 \end{vmatrix} = 2;$$

$$A_{12} = -\begin{vmatrix} 15 & -5 \\ 5 & 2 \end{vmatrix} = 5$$

$$A_{13} = \begin{vmatrix} -15 & 6 \\ 5 & -2 \end{vmatrix} = 0$$

$$A_{21} = -\begin{vmatrix} -1 & 1 \\ -2 & 2 \end{vmatrix} = 0$$

$$A_{22} = \begin{vmatrix} 2 & 1 \\ 5 & 2 \end{vmatrix} = -1$$

$$A_{23} = -\begin{vmatrix} 2 & -1 \\ 5 & -2 \end{vmatrix} = -1$$

$A_{31} = \begin{vmatrix} -1 & 1 \\ 6 & -5 \end{vmatrix} = -1$

$A_{32} = -\begin{vmatrix} 2 & 1 \\ -15 & -5 \end{vmatrix} = -5$

$A_{33} = \begin{vmatrix} 2 & -1 \\ -15 & 6 \end{vmatrix} = -3$

Hence $adj\ A = \begin{bmatrix} 2 & 0 & -1 \\ 5 & -1 & -5 \\ 0 & -1 & -3 \end{bmatrix}$

$A^{-1} = \frac{1}{-1} = \begin{bmatrix} 2 & 0 & -1 \\ 5 & -1 & -5 \\ 0 & -1 & -3 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 1 \\ -5 & 1 & 5 \\ 0 & 1 & 3 \end{bmatrix}$

Problem 2.

if$\omega = e^{2\pi i/3}$ find the inverse of the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}$

solution.

We note that $\omega^3 = 1$

$\therefore |A| \neq 0$, A is non-singular. Hence A^{-1}exists and is given by $A^{-1} = \frac{adj\ A}{|A|}$

Now, $adj\ A = \begin{bmatrix} \omega^2 - \omega & \omega^2 - \omega & \omega^2 - \omega \\ \omega^2 - \omega & \omega - 1 & 1 - \omega^2 \\ \omega^2 - \omega & 1 - \omega^2 & \omega - 1 \end{bmatrix}$

$\therefore A^{-1} = \frac{1}{3(\omega^2 - \omega)} \begin{bmatrix} \omega^2 - \omega & \omega^2 - \omega & \omega^2 - \omega \\ \omega^2 - \omega & \omega - 1 & 1 - \omega^2 \\ \omega^2 - \omega & 1 - \omega^2 & \omega - 1 \end{bmatrix}$

$$= \frac{1}{3\omega} \begin{bmatrix} \omega & \omega & \omega \\ \omega & 1 & -1 - \omega \\ \omega & -1 - \omega & 1 \end{bmatrix}$$

Problem 3.

Show that a square matrix A is orthogonal iff$A^{-1} = A^T$

Solution.

Suppose A is orthogonal. Then $AA^T = I$

$\therefore |A\ A^T| = |I| = 1$

$\therefore |A||\ A^T| = 1$

$\therefore |A||A| = 1$

$\therefore |A| \neq 0$and hence A is non-singular

$\therefore A^{-1}$exists.

Now, $A^{-1}(A\ A^T) = A^{-1}I$

$\therefore (A^{-1}A)A^T = A^{-1}$

$\therefore IA^T = A^{-1}$

$\therefore A^T = A^{-1}$

Conversely, let $A^T = A^{-1}$

Then $AA^T = AA^{-1} = I$ similarly $AA^T = I$

Hence A is orthogonal

Problem 4. Show that a square matrix A is involutoryiff$A = A^{-1}$

Solution. Suppose A is involutory. Then $A^2 = I$.

Hence $|A^2| = 1$

$\therefore |A^2| = |A||A| = 1$

$\therefore |A| \neq 0$and hence A is non-singular

$\therefore A^{-1}$exists

Now, $A^{-1}(AA) = A^{-1}I$

$\therefore (A^{-1}A)A = A^{-1}$

$\therefore IA = \ ^{-1}$

$\therefore A = A^{-1}$

Conversely, let $A = A^{-1}$

Then $A^2 = AA - A\ \ ^{-1} = 1$

$\therefore A$is involutory.

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 F
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 $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & -1 \end{bmatrix} A_1 = \begin{bmatrix} 3 & -1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$

$A_2 = \begin{bmatrix} 2 & 2 \\ 4 & 1 \\ 6 & -1 \end{bmatrix} A_3 = \begin{bmatrix} 1 & 2 \\ 5 & 7 \\ 3 & -1 \end{bmatrix}$ $A_1$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 $R_i \to cR_j$

(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 $m \times n$ 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 form the identity matrix by applying a single elementary row or column operation is called an elementary matrix

For example, $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$ are elementary matrices obtained from the identity matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ by applying the elementary operations$R_1 \leftrightarrow R_2$

$R_1 \to 4R_1, R_3 \to R_3 + 2R_2$ respectively

Exercise. Give examples of elementary matrices of order 4.

Theorem. Any elementary matrix is non-singular.

Proof.

The determinant of the identity matrix of any order is 1. Hence the determinant of an elementary matrix obtained by interchanging any two rows is -1. The determinant of an interchanging any two obtained by multiplying any row by$k \neq 0$is k. The determinant of an elementary matrix obtained by adding a multiple of one row with another row is 1. Hence any elementary matrix is non-singular.

Solved problems.

Problem 1.

Reduce the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{bmatrix}$ to the canonical form.

Solution. $A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{bmatrix}$

$$\sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{bmatrix} R_2 \to R_2 - R_1 \& R_3 \to R_3 - R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{bmatrix} C_2 \to C_2 - 2C_1 \& C_3 \to C_3 - C_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} C_2 \to C_3 + 3C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} R_2 \to -R_2$$

Problem 2. Find the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix}$

Solution.

$$\begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 1 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} AR_2 \to R_2 - 3R_1 \& R_3 \to R_3 + 2R_1$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 0 & 14 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 5 & -1 & 1 \end{bmatrix} AR_3 \to R_3 - R_2$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{bmatrix} A R_1 \to R_1 - \frac{1}{7}R_3 \quad , \quad R_2 \to R_2 + \frac{1}{2}R_3 \& R_3 \to \frac{1}{14}R_3$$

$$\Rightarrow A^{-1} = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{bmatrix}$$

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 $B = P^{-1}AP$.

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 $R_i = (a_{i1}, a_{i2}, \dots a_{in})$ 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.

Definition. The rank of a matrix A is the common value of its row and column rank

Solved Problems

Problem 1.

Find the rank of the matrix $A = \begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{bmatrix}$

Solution.

$$A = \begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 4 & 3 \\ 4 & 3 & 6 & 7 \\ 0 & 1 & 2 & 7 \end{bmatrix} C_1 \leftrightarrow C_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{bmatrix} C_1 \to C_2 + 2C_1, \quad C_3 \to C_3 + 4C_1, \quad C_4 \to C_4 + 3C_1$$

$$\sim\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -5 & -10 & -5\\ 0 & 1 & 2 & 7\end{bmatrix} R_2 \to R_2 + 4R_1$$

$$\sim\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -5 & 0 & 0\\ 0 & 1 & 0 & 6\end{bmatrix} C_3 \to C_3 - 2C_2, C_4 \to C_4 - C_2$$

$$\sim\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -5 & 0 & 0\\ 0 & 0 & 0 & 6\end{bmatrix} R_3 \to R_3 + \frac{1}{5}R_2$$

$$\sim\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & -5 & 0 & 0\\ 0 & 0 & 6 & 0\end{bmatrix} C_2 \leftrightarrow C_3$$

$$\sim\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix} R_2 \to -\tfrac{1}{5}R_2,\;\; R_3 \to \tfrac{1}{6}R_3$$

$\therefore$ Rank of $A = 3$

Problem 2. Find the rank of the matrix $A = \begin{bmatrix}1 & 1 & 1 & 1\\ 4 & 1 & 0 & 2\\ 0 & 3 & 4 & 2\end{bmatrix}$ by examining the determinant minors.

Solution.

$$\begin{bmatrix}1 & 1 & 1\\ 4 & 1 & 0\\ 0 & 3 & 4\end{bmatrix} = 0 = \begin{bmatrix}1 & 1 & 1\\ 1 & 0 & 2\\ 3 & 4 & 2\end{bmatrix}$$

$$\begin{bmatrix}1 & 1 & 1\\ 4 & 1 & 2\\ 0 & 3 & 2\end{bmatrix} = 0 = \begin{bmatrix}1 & 1 & 1\\ 4 & 0 & 2\\ 0 & 4 & 2\end{bmatrix}$$

$\therefore$Every 3×3 submatrix of A has determinant zero.

Also $\begin{vmatrix}1 & 1\\ 4 & 1\end{vmatrix} = -3 \neq 0$

$\therefore$ Rank of $A = 2$

Characteristic Equation and Caylay 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 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

$$A - xI = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$= \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix}.$$

$\therefore$ The characteristic polynomial of A is $|A - xI| = \begin{vmatrix} 1-x & 2 \\ 3 & 4-x \end{vmatrix}$

$$= (1-x)(4-x) - 6$$

$$= x^2 - 5x - 2$$

$\therefore$ 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

$$|A - xI| = \begin{vmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{vmatrix}$$

$$= (1-x)[(1-x)(-x) - 4] - 2(1-x)$$

$$= -x(1-x)^2 - 4(1-x) - 2 + 2x$$

$$= -x^3 + 2x^2 - x - 4 + 4x - 2 + 2x$$

$$= -x^3 + 2x^2 + 5x - 6$$

$\therefore$The characteristic equation of A is

$$-x^3 + 2x^2 + 5x - 6 = 0$$

(i.e) $x^3 - 2x^2 - 5x + 6 = 0$

Theorem (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 $a_0 + a_1A + a_2A^2 + \cdots + a_nA^n = 0$

Proof

Let A be a square matrix of order n.

Let $|A - xI| = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$.....(i)

Be the characteristic polynomial of A

Now, $adj\ (A - xI)$ is a matrix polynomial of degree $n - 1$ since each entry of the matrix $adj\ (A - xI)$is a cofactor of $A - xI$ and hence is a polynomial of degree $\leq n - 1$

∴ Let $adj\ (A - xI) = B_0 + B_1x + B_2x^2 + \cdots + B_{n-1}x^{n-1}$ ....(2)

Now, $(A - xI)adj\ (A - xI) = |A - xI|I$ (since $(adj\ A)A = A(adj\ A) = |A|I)$

$(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n)I$using (1) and (2)

∴ Equating the coefficients of the corresponding powers of x we ger

$AB_0 = a_0I$

$AB_1 - B_0 = a_1I$

$AB_2 - B_1 = a_2I$

……..

……..

$AB_{n-1} - B_{n-2} = a_{n-1}I$

$-B_{n-1} = a_nI$

Pre-multiplying the above equations by $I, A, A^2, \ldots, A^n$ respectively and adding we get

$a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = 0$

Note. The inverse of a non-singular matrix can be calculated by using the cayley Hamilton theorem as follows.

Let $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ be the characteristic polynomials of A

Then by theorem we have $a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = 0$ ....(3)

Since $|A - xI| = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ we get $a_0 = |A|$ (by putting $x = 0$)

∴$a_0 \neq 0$ (since A is a non singular matrix)

$\therefore I = -\frac{1}{a_0}[a_1 A + a_2 A^2 + \cdots + a_n A^n]$ (by (3))

$\therefore A^{-1} = -\frac{1}{a_0}[a_1 I + a_2 A + \cdots + a_n A^{n-1}]$

Solved problems.

Problem 1.

Find the characteristic equation of the matrix $A = \begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$

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.) $(8\lambda^2 - 80\lambda + 40 - \lambda^3 + 10\lambda^2 - 5\lambda) + (36\lambda - 60) + (4\lambda + 10) = 0$

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

∴Bycaylay-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 $\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$

Solution.

Let $A = \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$

The characteristic polynomial of $A = |A - xI| = \begin{vmatrix} 7-x & 2 & -2 \\ -6 & -1-x & 2 \\ 6 & 2 & -1-x \end{vmatrix}$

$= (7-x)[(1+x)^2 - 4] - 2[6(1+x) - 12] - 2[-12 + 6(1+x)]$

$= (7-x)(x^2 + 2x - 3) - 12(x-1) - 12(x-1)$

$= 7x^2 + 12x - 21 - x^3 - 2x^2 + 3x - 12x + 12 - 12x + 12$

$= -x^3 + 5x^2 - 7x + 3$

∴byCayley-Hamilton Theorem

$-A^3 + 5A^2 - 7A + 3I_3 = 0$

$\therefore A^3 - 5A^2 + 7A - 3I_3 = 0$

$\therefore 3I_3 = A^3 - 5A^2 + 7A$

$\therefore I_3 = \frac{1}{3}(A^3 - 5A^2 + 7A)$

Pre (or post) multiplying by A^{-1} on both sides we get

$\therefore A^{-1} = \frac{1}{3}(A^2 - 5A + 7I_3)$...(1)

Now, $A^2 = \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$

$= \begin{bmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{bmatrix}$

∴from (1)

$$A^{-1} = \frac{1}{3}\left(\begin{bmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{bmatrix} - \begin{bmatrix} 35 & 10 & -10 \\ -30 & -5 & 10 \\ 30 & 10 & -5 \end{bmatrix} + 7\begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}\right)$$

$$= \frac{1}{3}\begin{bmatrix} -3 & -2 & 2 \\ 6 & 5 & -2 \\ -6 & -2 & 5 \end{bmatrix}$$

Problem 5.

Find the inverse of the matrix $\begin{bmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix}$ using Caylay-Hamilton theorem.

Solution.

The characteristic polynomial of $A = |A - xI| = \begin{bmatrix} 3-x & 3 & 4 \\ 2 & -3-x & 4 \\ 0 & -1 & 1-x \end{bmatrix}$

$= -x^3 + x^2 + 11x - 11$

$\therefore$byCayley-Hamilton Theorem

$-A^3 + A^2 + 11A - 11 = 0$

$\therefore A^3 - A^2 - 11A + 11I_3 = 0$

Hence $11I_3 = -(A^3 - A^2 - 11A)$

$I_3 = -\frac{1}{11}(A^3 - A^2 - 11A)$

Pre (post) multiplying by A^{-1} on both sides we get

$A^{-1} = -\frac{1}{11}(A^2 - A - 11I_3)$

$$= -\frac{1}{11}\left[\begin{bmatrix} 15 & -4 & 28 \\ 0 & 11 & 0 \\ -2 & 2 & -3 \end{bmatrix} - \begin{bmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix} - 11\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right] = \begin{bmatrix} -\frac{1}{11} & \frac{7}{11} & -\frac{24}{11} \\ \frac{2}{11} & -\frac{3}{11} & \frac{4}{11} \\ \frac{2}{11} & -\frac{3}{11} & \frac{15}{11} \end{bmatrix}$$

Problem 6. Verify Cayley Hamilton's theorem foe the matrix $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$

Solution.

The characteristic equation of A is $|A - \lambda I| = 0$

$\therefore \begin{vmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{vmatrix} = 0$

$\therefore (1-\lambda)(3-\lambda) - 8 = 0$

$\therefore \lambda^2 - 4\lambda - 5 = 0$

By byCayley-Hamilton Theorem A satisfies its characteristic equation

$\therefore$We have $A^2 - 4A - 5I = 0$

Now, $A^2 = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix}$

$4A = \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix}and5I = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$

$A^2 - 4A - 5I = \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix} - \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix} - \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$

The Cayley-Hamilton Theorem is verified

Theorem 7.

Using Cayley-Hamilton Theorem for matrix $A = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix}$ 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.) $\lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0$

By Cayley-Hamilton Theorem

$A^3 - 5A^2 + 8A - 4I_3 = 0$ …. (1)

$4I_3 = A^3 - 5A^2 + 8A$

(i) To find A^{-1} pre multiplying by A^{-1} we get

$4A^{-1} = A^{-1}A^3 - 5A^{-1}A^2 + 8A^{-1}A$

$4A^{-1} = A^2 - 5A + 8I$

$A^{-1} = \frac{1}{4}(A^2 - 5A + 8I)$ … (2)

Now, $A^2 = \begin{bmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{bmatrix}$

From (2)

$$A^{-1} = \frac{1}{4}\left(\begin{bmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 0 & -10 \\ 10 & 10 & 20 \\ 0 & 0 & 10 \end{bmatrix} + \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix}\right)$$

$$A^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

(ii) To find A^4

From (1) $A^3 = 5A^2 - 8A + 4I_3$

$= 5(5A^2 - 8A + 4I) - 8A + 4I_3$ (using (1))

$= 17A^2 - 36A + 20I$

$$= 17\begin{bmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{bmatrix} - 36\begin{bmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix} + 20\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 17 & 0 & -102 \\ 102 & 684 & 204 \\ 0 & 0 & 68 \end{bmatrix} - \begin{bmatrix} 36 & 0 & -72 \\ 72 & 72 & 144 \\ 0 & 0 & 72 \end{bmatrix} + \begin{bmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{bmatrix}$$

$$\therefore A^4 = \begin{bmatrix} 1 & 0 & -30 \\ 30 & 16 & 260 \\ 0 & 0 & 16 \end{bmatrix}$$

UNIT - V
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 non-zero vector $X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$ such that $AX = \lambda X$ and X is called an eigen vector correcponding 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 $AX = \lambda$ 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 to the eigen values $\lambda_1$and λ_2. Then $\lambda_1 = \lambda_2$

Proof . By definition $X \neq 0, AX = \lambda_1 X$ and $AX = \lambda_2 X$

$\therefore \lambda_1 X = \lambda_2 X$

$\therefore (\lambda_1 - \lambda_2)X = 0$

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 to the sum of the diagonal elements (trace) of A

(ii) product of eigen values of A is $|A|$

Proof.

(i) Let $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$

The eigen values of A are the roots of the characteristic equation

$$|A-\lambda I| = \begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} = 0 \dots (1)$$

Let $|A-\lambda I| = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_n$ ... (2)

From (1) and (2) weget

$a_0 = (-1)^n;\ a_1 = (-1)^{n-1}(a_{11} + a_{22} + \cdots + a_{nn})$;...(3)

Also by putting $\lambda = 0$ is (2) we get $a_n = |A|$

Now let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigen values of A.

$\therefore \lambda_1, \lambda_2, \dots, \lambda_n$ are the roots of (2)

$\therefore \lambda_1 + \lambda_2 + \cdots + \lambda_n = -\frac{a_1}{a_0} = a_{11} + a_{22} + \cdots + a_{nn}$ (using (3))

∴sum of the eigen values= trace of A.

(ii) Product of the eigen values =product of the roots

$= \lambda_1\lambda_2 \dots \lambda_n$

$= (-1)^n \dfrac{a_n}{a_0}$

$= \dfrac{(-1)^n a_n}{(-1)^n}$

$= a_n$

$= |A|$

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

$$|A-\lambda I| = |(A-\lambda I)^T| = |(A)^T - (\lambda I)^T| = |A^T - \lambda\ |$$

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

$\therefore A^{-1}(AX) = A^{-1}(\lambda X)$

$IX = \lambda A^{-1}X$

$\therefore A^{-1}X = \left(\frac{1}{\lambda}\right)X$

$\therefore \left(\frac{1}{\lambda}\right)$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 toλ

Then $AX = \lambda$ ...(1)

Now, $(kA)X = k(AX)$

$= k(\lambda X)$ (by (1))

$= (k\lambda)X$

$\therefore k\lambda$ 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, $A^2X = (AA)X = A(AX)$

$= A(\lambda X$ (by(1))

$= \lambda(AX)$

$= \lambda(\lambda X)$ (by (1))

$= \lambda^2 X$

$\lambda^2$is an eigen value of A^2

Proceeding like this we can prove that λ^k is an eigen value of A^k where k is any positive integer

Corollary. If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the 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 $\lambda_1, \lambda_2, \dots, \lambda_n$ be distinct eigen values of a matrix and let X_i be the eigen vector corresponding to λ_i

Hence $AX_i = \lambda_i X_i (i = 1,2,3, \dots k)$ ... (1)

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 i.

By rearranging the vectors $X_1, X_2, \dots, X_k$ we may assume that

$\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_j X_j = 0$...(2)

$\therefore A(\alpha_1 X_1) + A(\alpha_2 X_2) + \cdots + A(\alpha_j X_j) = 0$

$\therefore \alpha_1 (AX_1) + \alpha_2 (AX_2) + \cdots + \alpha_j (AX_j) = 0$

$\therefore \alpha_1 (\lambda_1 X_1) + \alpha_2 (\lambda_2 X_2) + \cdots + \alpha_j (\lambda_j X_j) = 0$...(3)

Multiplying (2) by λ_1 and subtracting from (3), we get

$\therefore \alpha_2 (\lambda_1 - \lambda_2) X_2 + \alpha_3 (\lambda_1 - \lambda_3) X_3 + \cdots + \alpha_j (\lambda_1 - \lambda_j) X_j = 0$...(4)

And since $\lambda_1, \lambda_2, \dots, \lambda_n$ 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 dependent.

Property 8. The characteristic roots of a Hermitian matrix are all real

Proof.

Let A be a Hermitian matrix

Hence $A = A^{-1}$...(1)

Let λ be a characteristic root of A and let S be a characteristic vector corresponding to λ

$\therefore AX = \lambda X$....(2)

Now

$AX = \lambda \quad \Rightarrow \bar{X}^T AX = \lambda \bar{X}^T X$

$\Rightarrow (\bar{X}^T AX)^T = \lambda \bar{X}^T X$ (since$X^T AX$ is a 1×1 matrix)

$\Rightarrow X^T A^T (\bar{X}^T)^T = \lambda \bar{X}^T X$

$\Rightarrow X^T A^T \bar{X} = \lambda \bar{X}^T X$

$\Rightarrow \overline{X^T A^T \bar{X}} = \overline{\lambda X^T X}$

$\Rightarrow \bar{X}^T \bar{A}^T X = \bar{\lambda} X^T \bar{X}$

$\Rightarrow \bar{X}^T AX = \bar{\lambda} X^T \bar{X}$ (using 1)

$\Rightarrow \bar{X}^T \lambda X = \bar{\lambda} X^T \bar{X}$ (using 2)

$\Rightarrow \lambda (\bar{X}^T X) = \bar{\lambda} (X^T \bar{X})$...(3)

Now,

$\bar{X}^T X = X^T \bar{X} = \overline{x_1} x_1 + \overline{x_2} x_2 + \cdots + \overline{x_n} x_n$

$= |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$

$\neq 0$

∴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

$|A - \lambda I| = 0$

$|iA - i\lambda I| = 0$

$i\lambda$is a characteristic root of iA

Since A is skew HermitianiA 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 characteristic root of an 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 an unitary matrix A and X be a characteristic vector corresponding to λ

∴$AX = \lambda X$….(1)

Taking conjugate and transpose in (1) we get

$(\overline{AX})^T = \left(\overline{\lambda X}\right)^T$

∴$\bar{X}^T \bar{A}^T = \bar{\lambda}\bar{X}^T$…..(2)

Multiplying (1) and (2) we get

$\overline{X^T A^T}(AX\text{-}) = (\bar{\lambda}\bar{X}^T)(\lambda X)$

$\therefore \bar{X}^T(\bar{A}^T A)X = (\bar{\lambda}\lambda)(\bar{X}^T X)$

Now, since A is an unitary matrix $\bar{A}^T A = 1$

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

$(\bar{X}^T X) = |x_1|^2 + |x_2|^2 + \cdots + |x_n|^2 \neq 0$we get $(\bar{\lambda}\lambda) = 1$

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 $A \Leftrightarrow |A - 0I| = 0$

$\Leftrightarrow |A| = 0$

$\Leftrightarrow A$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.

$\therefore (AB)X = \lambda X$

$\therefore B(AB)X = B(\lambda X) = \lambda(BX$

$\therefore (BA)(BX) = \lambda(BX$

$\therefore (BA)Y = \lambda Y$ 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 non-singular 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 $|B - \lambda I| = |P^{-1}AP - \lambda I|$

$= |P^{-1}AP - P^{-1}(\lambda I)P|$

$= |P^{-1}(A - \lambda I)P|$

$= |P^{-1}||A - \lambda I||P|$

$= |P^{-1}||P||A - \lambda I|$

$= |P^{-1}P||A - \lambda I|$

$= |I||A - \lambda I|$

$= |A - \lambda I|$

∴The characteristic equation of A and and $P^{-1}AP$ have the same eigen values

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 $f(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$ where $a_0 \neq 0$ and $a_0, a_1, \ldots a_n$ are all real numbers

$\therefore f(A) = a_0 A^n + a_1 A^{n-1} + \cdots + a_{n-1}A + a_n I$

Since λ is a characteristic root of A, λ^n is a characteristic root of A^n for any positive integer n (refer property 6)

$\therefore A^n X = \lambda^n X$

$A^{n-1}X = \lambda^{n-1}X$

………..

………..

$AX = \lambda X$

$a_0 A^n X = a_0 \lambda^n X$

$a_1 A^{n-1} X = a_1 \lambda^{n-1} X$

…………

…………

$a_{n-1} AX = a_{n-1} \lambda X$

Adding the above equations we have

$a_0 A^n X + a_1 A^{n-1} X + \cdots + a_{n-1} AX = a_0 \lambda^n X + a_1 \lambda^{n-1} X + \cdots + a_{n-1} \lambda X$

$\therefore (a_0 A^n + a_1 A^{n-1} + \cdots + a_{n-1} A)X = (a_0 \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda)X$

$\therefore (a_0 A^n + a_1 A^{n-1} + \cdots + a_{n-1} A + a_n I)X = (a_0 \lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n)X$

$\therefore f(A)X = f(\lambda)X$

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 scalar) is also an eigen vector corresponding to λ

Solution.

Since X_1, X_2 are eigen vectors corresponding to an eigen value λ, we have

$AX_1 = \lambda X_1$and$AX_2 = \lambda X_2$

And hence $A(aX_1) = \lambda(aX_1)$ and $A(bX_2) = \lambda(bX_2)$

$\therefore A(aX_1 + bX_2) = \lambda(a\ \ _1 + bX_2)$

$\therefore (aX_1 + bX_2$ is an eigen vector corresponding to λ

Problem 2.

If the eigen values of $A = \begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$ 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 $A = \begin{bmatrix} 3 & 0 & 0 \\ 5 & 4 & 0 \\ 3 & 6 & 1 \end{bmatrix}$

Solution. 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 $\begin{bmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{bmatrix}$ without actually finding the eigen values.

Solution.

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

Sum of the eigen values = trace of A = $3 + (-2) + 3 = 4$

Product of the eigen values = $|A|$

Now, $|A| = \begin{vmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{vmatrix}$

$= 3(-6 + 4) + 4(3 - 4) - 4(-1 + 2)$

$= -6 - 4 - 4 = -14$

∴The 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 $|A - \lambda I| = 0$

$\begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ -\sin\theta & \cos\theta - \lambda \end{vmatrix} = 0$

$(\cos\theta - \lambda)^2 - sin^2\theta = 0$

$(\cos\theta - \lambda - sin\theta)(\cos\theta - \lambda + sin\theta) = 0$

$[\lambda - (\cos\theta - sin\theta)][\lambda - (\cos\theta + sin\theta)] = 0$

The two characteristic roots (the two eigen values of the matrix are $(\cos\theta - sin\theta)$ and $(\cos\theta + sin\quad)$

Problem 6.

Find the characteristic roots of the matrix $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$

Solution.

Let $A = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

The characteristic equation of A is given by $|A - \lambda I| = 0$

$\begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ -\sin\theta & -\cos\theta - \lambda \end{vmatrix} = 0$

$-(\cos^2\theta - \lambda^2) - sin^2\theta = 0$

$\lambda^2 - (\cos^2\theta + sin^2\theta) = 0$

$\lambda^2 - 1 = 0$

The characteristic roots 1 and -1

Problem 7.

Find the sum and product of the eigen values of the matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ 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

$$A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$$

Solution

The characteristic equation of A is $|A - \lambda I| = 0$

(i.e) $\begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & 0-\lambda \end{vmatrix} = 0$

(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

The sum of the eigen valued $= -\left(\frac{coefficiient\ of\ \lambda^2}{coefficiient\ of\ \lambda^3}\right) = -1$

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 of $|A|$

(i.e.) $\lambda_1\lambda_2\lambda_3 = \begin{vmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{vmatrix}$

(i.e.) $16\lambda_3 = 6(9-1) + 2(-6+2) + 2(2-6)$

$= 48 - 8 - 8$

$= 32$

$\therefore \lambda_3 = 2$

∴The third eigen value is 2

Also we know that the sum of the eigen vales of

$A = trace\ of\ A = 6 + 3 + 3 = 12$

Problem 10.

The product of the two eigen values of the matrix $A = \begin{bmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{bmatrix}$ is -12. Find 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 -12

$\therefore \lambda_1, \lambda_2 = -12$…(1)

We know that the product of the eigen values of $|A|$

(i.e.) $\lambda_1\lambda_2\lambda_3 = \begin{vmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{vmatrix}$

(i.e.) $12\lambda_3 = -12$

$\therefore \lambda_3 = 1$...(2)

∴The third eigen value is 1

Also we know that the sum of the eigen vales $= tra \quad of\ A$

$\lambda_1 + \lambda_2 + \lambda_3 = 2 + 1 - 3 = 0$

$\lambda_1 + \lambda_2 = -1$ (using (2)) … (3)

Using (3) in (1) we get $\lambda_1(-1 - \lambda_1) = -12$

${\lambda_1}^2 + \lambda_1 - 12 = 0$

$(\lambda_1 + 4)(\lambda_1 - 3) = 0$

$\lambda_1 = 3$or-4

Putting $\lambda_1 = 3$ in (1) we get $\lambda_2 = -4$. Or putting $\lambda_1 = -4$ in (4) 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

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

Solution.

The characteristic equation of A is $|A - \lambda I| = 0$

$$\therefore \begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix} = 0$$

$\therefore (1-\lambda)[(5-\lambda)(1-\lambda) - 1] - [(1-\lambda) - 3] + 3[1 - 3(5-\lambda)] = 0$

$(1-\lambda)(\lambda^2 - 6\lambda + 4) + (\lambda + 2) + 3(3\lambda - 14) = 0$

$\lambda^2 - 6\lambda + 4 + 6\lambda^2 - 4\lambda + \lambda + 2 + 9\lambda - 42 = 0$

$\therefore -\lambda^3 + 7\lambda^2 - 36 = 0$. 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 $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ be an eigen vector corresponding to $\lambda = -2$

Hence $AX = -2X$

(i.e.) $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_1 \\ -2x_2 \\ -2x_3 \end{bmatrix}$

$\therefore x_1 + x_2 + 3x_3 = -2x_1$

$x_1 + 5x_2 + x_3 = -2x_2$

$3x_1 + x_2 + x_3 = -2x_3$

$\therefore 3x_1 + x_2 + 3x_3 = 0$

$x_1 + 7x_2 + x_3 = 0$

$3x_1 + x_2 + 3x_3 = 0$

Clearly this system of three equations reduces to two equations only from (1) and (2) we get

$\therefore x_1 = -2k; \;\; x_2 = 0; \;\; x_3 = 2k$

$\therefore$It has only one independent solution and can be obtained by giving any value to k say $k = 1$

$\therefore(-2,0,2)$ is an eigen vector corresponding to $\lambda = -2$

Case (ii)

Eigen vector corresponding to $\lambda = 3$.

Then $AX \quad 3X$ gives

$$-2x_1 + x_2 + 3x_3 = 0$$

$x_1 + 2x_2 + x_3 = 0$

$3x_1 + x_2 \pm 2 = 0$

Taking the first 2 equations we get

$\frac{x_1}{-5} = \frac{x_2}{5} = \frac{x_3}{-5} = k(sa$

$\therefore x_1 = -k;\ x_2 = k;\ x_3 = -k$

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 $AX \quad 6X$

Hence $-5x_1 + x_2 + 3x_3 = 0$

$x_1 - x_2 + x_3 = 0$

$3x_1 + x_2 - 5x_3 = 0$

Taking the first two equation we get

$\frac{x_1}{4} = \frac{x_2}{8} = \frac{x_3}{4} = k$

$\therefore x_1 = k;\ x_2 = 2k;\ x_3 = k$. 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

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$$

Solution.

The characteristic equation of A is $|A - \lambda I| = 0$

$$\therefore \begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0$$

$\therefore (6-\lambda)[(3-\lambda)^2 - 1] + 2[(2\lambda - 6) + 2] + 2[2 - 6 + 2\lambda] = 0$

$(6-\lambda)(8 + \lambda^2 - 6\lambda) + 4\lambda - 8 + 4\lambda - 8 = 0$

$48 + 6\lambda^2 - 36\lambda - 8\lambda - \lambda^3 + 6\lambda^2 + 8\lambda - 16 = 0$

$\therefore -\lambda^3 + 12\lambda^2 - 36\lambda + 32 = 0$. Hence $\lambda^3 - 12\lambda^2 + 36\lambda - 32 = 0$

Hence $(\lambda - 2)(\lambda - 2)(\lambda - 8) = 0$

$\therefore \lambda = -2,2,8$ are the three eigen values

Case (i)

Eigen vector corresponding to $\lambda = 2$

Let $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ be an eigen vector corresponding to $\lambda = 2$

Hence $AX = 2X$

$\therefore 6x_1 - 2x_2 + 2x_3 = 2x_1$

$-2x_1 + 3x_2 - x_3 = 2x_2$

$2x_1 - x_2 + 3x_3 = 2x_3$

$\therefore 4x_1 - 2x_2 + 2x_3 = 0$

$-2x_1 + x_2 - x_3 = 0$

$2x_1 - x_2 + x_3 = 0$

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 $x_3 = -2$

Two independent vectors corresponding to $\lambda = 2$ are $(1,2,0)$ and $(3,4,-2)$

Case (ii)

Eigen vector corresponding to $\lambda = 8$.

The eigen vector $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is got from $AX \quad 8X$ gives

$-2x_1 - 2x_2 + 2x_3 = 0$...(1)

$-2x_1 - 5x_2 - x_3 = 0$....(2)

$2x_1 - x_2 - 5x_3 = 0$...(3)

From (1) and (2) we get

$\frac{x_1}{12} = \frac{x_2}{-6} = \frac{x_3}{5} = k(say)$

$\therefore x_1 = 2k;\ x_2 = -k;\ x_3 = k$

Giveing$k = 1$(say) $(-1,1,-1)$ is an eigen vector corresponding to 8 as $(2,-1,1)$

Problem 14.

Find the eigen values and eigen vectors of the matrix

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

Solution.

The characteristic equation of A is $|A - \lambda I| = 0$

$$\therefore \begin{vmatrix} 2-\lambda & -2 & 2 \\ 1 & 1-\lambda & 1 \\ 1 & 3 & -1-\lambda \end{vmatrix} = 0$$

$\therefore (2-\lambda)[-(1-\lambda)(1+\lambda) - 3] + 2[-(1+\lambda) - 1] + 2[3 - (1-\lambda)] = 0$

$(2-\lambda)(\lambda^2 - 4) - 2(2+\lambda) + 2(2+\lambda) = 0$

$-\lambda^3 + 2\lambda^2 + 4\lambda - 8 = 0$

$\therefore -\lambda^3 + 2\lambda^2 + 4\lambda - 8 = 0$.Hence $\lambda^3 - 2\lambda^2 - 4\lambda + 8 = 0$

Hence $(\lambda - 2)(\lambda - 2)(\lambda + 2) = 0$

$\therefore \lambda = 2, 2, -2$ are the three eigen values

Case (i)

Eigen vector corresponding to $\lambda = 2$

Let $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ be an eigen vector corresponding to $\lambda = 2$

Hence $AX = 2X$

$$\begin{bmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{bmatrix}$$

The eigen vector corresponding to $\lambda = 2$ is given by the equations

$\therefore 2x_1 - 2x_2 + 2x_3 = 2x_1$

$x_1 + x_2 + x_3 = 2x_2$

$x_1 + 3x_2 - x_3 = 2x_3$

$\therefore -x_2 + 2x_3 = 0$…(1)

$x_1 - x_2 + x_3 = 0$...(2)

$x_1 + 3x_2 - 3x_3 = 0$...(3)

From (1) and (2) we get

$$\frac{x_1}{0} = \frac{x_2}{1} = \frac{x_3}{1} = k(say)$$

$\therefore x_1 = 0;\ x_2 = k;\ x_3 = k$

Giveing$k = 1$(say) $(0,1,1)$ is an eigen vector corresponding to $\lambda = 2$

Case (ii)

Eigen vector corresponding to $\lambda = -2$.

The eigen vector $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is got from $AX \quad 8X$ gives

$$2x_1 - 2x_2 + 2x_3 = -2x_1$$

$x_1 + x_2 + x_3 = -2x_1$

$x_1 + 3x_2 - x_3 = -2x_1$

$2x_1 - x_2 + x_3 = 0$

$x_1 + 3x_2 + x_3 = 0$

$x_1 + 3x_2 + x_3 = 0$

Taking the first two equations we get

$$\frac{x_1}{-4} = \frac{x_2}{-1} = \frac{x_3}{7} = k(say)$$

$\therefore x_1 = -4k;\ x_2 = -k;\ x_3 = 7k$

Giveing$k = 1$we get $(-4, -1, 7)$ as an eigen vector corresponding to the eigen value $\lambda = -2$.

8. Bilinear forms

8.0. Introduction

Consider a finite dimensional, inner product space V over the field $\mathbf{R}$ of real numbers. The inner product is a function from $V \times V$ to $\mathbf{R}$ satisfying the following conditions

(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 a linear function in each of the two variables. This type of scalar valued functions are called *bilinear forms*. In this chapter we study bilinear forms on finite dimensional vector spaces.

8.1. Bilinear forms

Definition. Let V be a vector space over a field F. A bilinear form on V is a function $f : V \times V \to F$ such 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 of the two variables when the other is fixed.

Examples

1. Let V be a vector space over $\mathbf{R}$. 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,

$$\begin{aligned}\hat{0}(\alpha u_1 + \beta u_2, v) &= 0\\ &= \alpha 0 + \beta 0\\ &= \alpha\hat{0}(u_1, v) + \beta\hat{0}(u_2, v)\end{aligned}$$

Similarly

$$\hat{0}(u, \alpha v_1 + \beta v_2) = \alpha \hat{0}(u, v_1) + \beta \hat{0}(u, v_2)$$

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 $f(x, y) = x_1 y_1 + x_2 y_2 + \dots\dots\dots + x_n y_n$ where $x = (x_1, x_2, \dots\dots, x_n)$ and $y = (y_1, y_2, \dots\dots\dots, y_n)$ 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 8.1. 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

$$(f+g)(u, v) = f(u, v) + g(u, v) \quad \text{and}$$
$$(\alpha f)(u, v) = \alpha f(u, v).$$

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
$$\begin{aligned}(\alpha_1 f)(\alpha u_1 + \beta u_2, v) &= \alpha_1 f(\alpha u_1 + \beta u_2, v)\\ &= \alpha_1[\alpha f(u_1, v) + \beta f(u_2, v)]\\ &= \alpha_1 \alpha f(u_1, v) + \alpha_1 \beta f(u_2, v)\\ &= \alpha[(\alpha_1 f)(u_1, v)] + \beta[(\alpha_1 f)(u_2, v)]\end{aligned}$$

Similarly

$$(\alpha_1 f)(u, \alpha v_1 + \beta v_2) = \alpha[(\alpha_1 f(u, v_1)] + \beta[(\alpha_1 f)(u, v_2)]$$

$\therefore \quad \alpha_1 f \in L(V, V, F)$.

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 $\{v_1, v_2, \ldots\ldots, v_n\}$ for V.

Let $u = \alpha_1 v_1 + \ldots\ldots + \alpha_n v_n$ and $v = \beta_1 v_1 + \ldots\ldots + \beta_n v_n$.

Then $f(u, v)$

$$= f(\alpha_1 v_1 + \ldots\ldots + \alpha_n v_n, \beta_1 v_1 + \ldots\ldots + \beta_n v_n)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \beta_j f(v_i, v_j)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\alpha_i \beta_j \text{ where } f(v_i, v_j) = a_{ij}$$

$$= (\alpha_1, \ldots\ldots, \alpha_n) \begin{pmatrix} a_{11} \ldots a_{1n} \\ \ldots\ldots\ldots \\ a_{n1} \ldots a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \ldots \\ \beta_n \end{pmatrix}$$

$\therefore \quad f(u, v) = XAY^T$ 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 $f : V \times V \to F$ 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) $f(e_1, e_1) = f((1, 0), (1, 0))$

$\qquad = 1 \times 1 + 0 \times 0 = 1$

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

(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 $\begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$

Exercises

1. Find the matrix of the bilinear form $f(x, y) = x_1 y_2 - x_2 y_1$ 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 $f(x, y) = x_1 y_1 + x_3 y_3$ 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.

1. $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

2. (a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

Theorem 8.2. Let V be a vector space of dimension n over a field F. Fix a basis $\{v_1, v_2, \ldots\ldots, v_n\}$ 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 $(f + g)(v_i, v_j) = f(v_i, v_j) + g(v_i, v_j)$

$= a_{ij} + b_{ij}$

$\therefore \quad \varphi(f+g) = (a_{ij}+b_{ij}) = (a_{ij}) + (b_{ij})$

$= \varphi(f) + \varphi(g).$

Also $(\alpha f)(v_i, v_j) = \alpha f(v_i, v_j) = \alpha a_{ij}$

$\therefore \quad \varphi(\alpha f) = (\alpha a_{ij}) = \alpha(a_{ij}) = \alpha\varphi(f).$

Thus φ is an isomorphism.

Corollary. $L(V, V, F)$ is a vector space of dimension n^2.

8.2. 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 $\mathbf{R}$. Then any inner product defined on V is a symmetric bilinear form.

(ii) The bilinear form $\hat{0}$ defined in example 2 of 8.1 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 8.3. A bilinear form f defined on V is symmetric iff its matrix (a_{ij}) w.r.t any one basis $\{v_1, v_2, \ldots\ldots\ldots, v_n\}$ is symmetric.

Proof. Let f be a symmetric bilinear form.

Now, $a_{ij} = f(v_i, v_j)$

$= f(v_j, v_i)$ (since f is symmetric)

$= a_{ji}$

$\therefore$ (a_{ij}) is a symmetric matrix.

Conversely, let (a_{ij}) be a symmetric matrix.

Hence $A = A^T$ (by theorem 7.5)

Then

$f(u, v) = XAY^T$

$= (XAY^T)^T$ (since XAY^T is a 1×1 matrix)

$$= YA^T X^T$$
$$= YAX^T$$
$$= f(v, u)$$

$\therefore$ 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.

4

Examples

1. Consider the bilinear form f defined on $V_n(F)$ by $f(u, v) = x_1y_1 + x_2y_2 + \ldots\ldots + x_ny_n$; $u = (x_1, \ldots\ldots, x_n)$, $v = (y_1, \ldots, y_n)$. Then the quadratic form q associated with f is given by

$$q(u) = f(u, u) = x_1^2 + \ldots\ldots + x_n^2.$$

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

$$q(X) = XAX^T = \sum_{i,j=1}^{n} a_{ij}x_ix_j$$

Ex

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 $V_3(\mathbf{R})$ is given by

$$q(v) = (x_1, x_2, x_3 \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$
$$= x_1^2 + 4x_2^2 + 6x_3^2 + 4x_1x_2 + 14x_2x_3 + 6x_1x_3.$$

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.t. the standard basis for $V_3(\mathbf{R})$ is given by

$$q(v) = (x_1, x_2, x_3) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
$$= x_1^2 + 2x_2^2 + 3x_3^2.$$

m
d
y

We say that this quadratic form q is in the diagonal form.

4. Consider the quadratic form defined on $V_2(\mathbf{R})$ 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

$$2x_1^2 + x_1x_2 + x_2^2 = (x_1, x_2) \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$= ax_1^2 + 2bx_1x_2 + cx_2^2.$$

$$\therefore \quad a = 2;\ b = \frac{1}{2};\ c = 1.$$

$$\therefore \quad A = \begin{bmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$$

Exercises

1. Find the quadratic forms associated with the following matrices w.r.t. the standard basis.

(a) $\begin{bmatrix} 2 & -3 & 1 \\ -3 & 2 & 4 \\ 1 & 4 & -5 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & -2 & -4 \\ 3 & -4 & -3 \end{bmatrix}$

2. Find the matrices for the following quadratic forms.

(a) $x_1^2 + 4x_1x_2 + 3x_2^2$ in $V_2(\mathbf{R})$

(b) $2x_1^2 + x_2^2 + 3x_1x_2$ in $V_2(\mathbf{R})$

(c) $2x_1^2 + x_3^2 - 6x_1x_2$ in $V_3(\mathbf{R})$

(d) x_1x_2 in $V_4(\mathbf{R})$

(e) $x_1x_3 + x_4^2$ in $V_4(\mathbf{R})$.

Answers.

1. (a) $2x_1^2 + 2x_2^2 - 5x_3^2 - 6x_1x_2 + 8x_2x_3 + 2x_1x_3$.

   (b) $x_1^2 - 2x_2^2 + 3x_3^2 + 4x_1x_2 - 8x_2x_3 + 6x_1x_3$.

2. (a) $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{bmatrix}$

   (c) $\begin{bmatrix} 2 & -3 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

   (e) $\begin{bmatrix} 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Theorem 8.4. 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) $\frac{1}{4}\{q(u+v) - q(u-v)\}$

$q(u) = f(u,u)$

$$= \frac{1}{4}\{f(u+v, u+v) - f(u-v, u-v)\}$$

$$= \frac{1}{4}\{f(u,u) + f(u,v) + f(v,u) + f(v,v)$$
$$- f(u,u) + f(u,v) + f(v,u) - f(v,v)\}$$

$$= \frac{1}{4}\{4f(u,v)\}$$

$$= f(u, v).$$

Proof of **(ii)** is similar.

Note. the above theorem shows that if f is a symmetric bilinear form and q the associated 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.

Reduction of a quadratic form to the diagonal form

In example 3 of the quadratic form in section 8.2 we have seen that a quadratic form associated with a diagonal matrix of order n is of the form

$$a_1x_1^2+a_2x_2^2+\ldots\ldots+a_nx_n^2$$

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

$$\varphi=\varphi(x_1,x_2,\ldots\ldots,x_n)=\sum_{i,j=1}^{n}a_{ij}x_ix_j$$

$$=a_{11}x_1^2+\ldots\ldots+a_{nn}x_n^2+2a_{12}x_1x_2$$

$$+\ldots\ldots+2a_{n(n-1)}x_nx_{n-1}.$$

Case (i) Suppose at least one of $a_{11},\ldots,a_{nn}$ is not zero. We assume, without loss of generality, that $a_{11}\neq 0$.

Then

$$\varphi=(a_{11}x_1^2+2a_{12}x_1x_2+\ldots\ldots+2a_{1n}x_1x_n)$$

$$+\sum_{i,j=2}^{n}a_{ij}x_ix_j$$

$$=a_{11}\left(x_1^2+2\frac{a_{12}}{a_{11}}x_1x_2+\ldots\ldots+2\frac{a_{1n}}{a_{11}}x_1x_n\right)$$

$$+\varphi_1(x_2,\ldots\ldots,x_n)\text{ (say)}$$

$$= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2$$
$$+ \varphi_2(x_2, \ldots\ldots\ldots, x_n) \text{ (say)}$$

Now, putting $y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n$, $y_2 = x_2, \ldots, y_n = x_n$, φ reduces to

$$\varphi = \alpha_1 y_1^2 + \varphi_2(y_2, \ldots, y_n) \qquad \ldots (1)$$

where $\alpha_1 = a_{11}$

Case (ii) Suppose $a_{11} = a_{22} = \ldots\ldots = a_{nn} = 0$. 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

$$\varphi_2 = \alpha_2 z_2^2 + \varphi_3(z_2, \ldots, z_n) \text{ so that}$$
$$\varphi = \alpha_1 z_1^2 + \alpha_2 z_2^2 + \varphi_3(z_2, \ldots, z_n).$$

Continuing this process of reduction we obtain φ in the form $\varphi = \alpha_1 w_1^2 + \ldots\ldots + \alpha_r w_r^2$.

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

$$\varphi = y_1^2 + 8y_2^2 + 8y_2y_3$$
$$= y_1^2 + 8(y_2 + \frac{1}{2}y_3)^2 - 2y_3^2.$$

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 $\varphi = 2w_1^2 - \frac{1}{2}w_2^2 - \frac{1}{2}w_3^2$

where

$$w_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{1}{2}x_4$$

$$w_2 = -x_1 + x_2;\ w_3 = x_3 + x_4;\ w_4 = x_4.$$

Exercises Reduce the following quadratic forms to diagonal 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. $(x_1x_2x_3)\begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & -1 \\ 2 & -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$