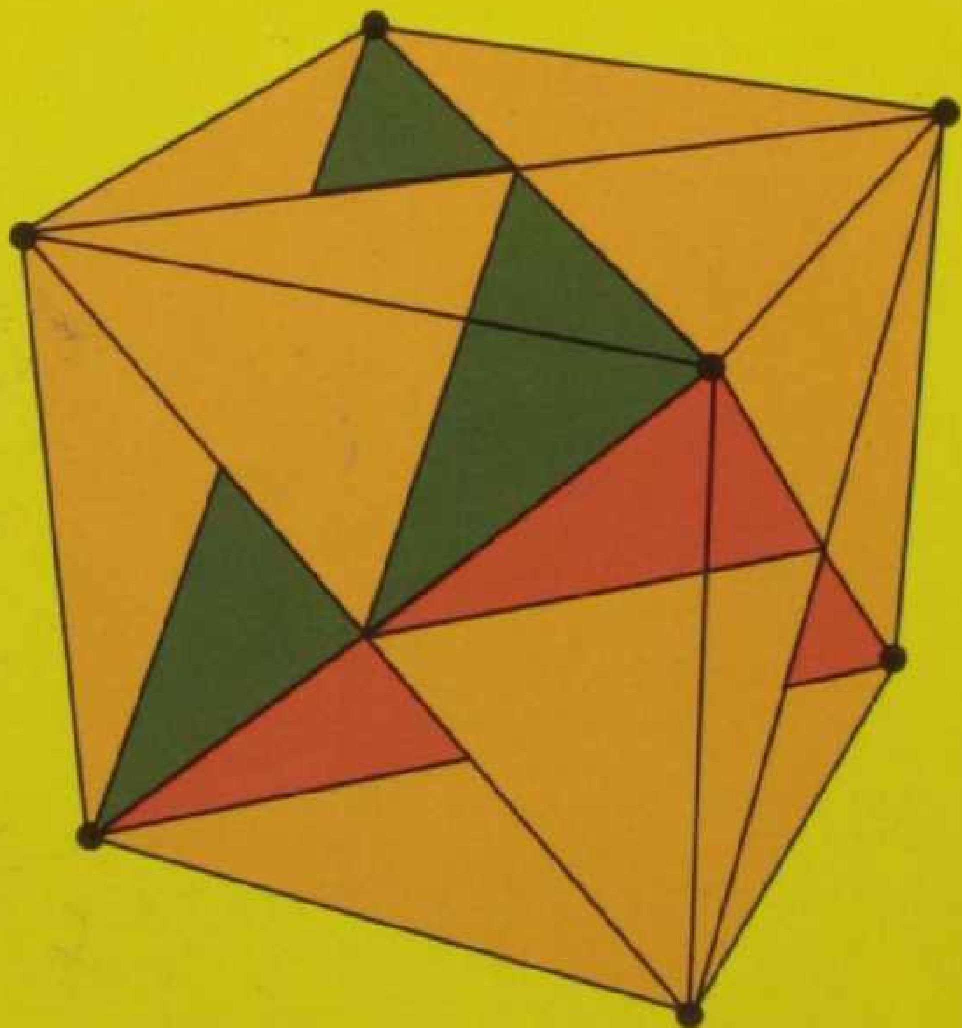
12th Edition

HIGHER ALGEBRA

Linear



S K Mapa

4. Vector Spaces

1. Matrices (I)

1.1. Introduction.

A rectangular array of mn elements a_{ij} into m rows and n columns, where the elements a_{ij} belong to a field F, is said to be a *matrix* of order $m\times n$ (or an $m\times n$ matrix) over the field F. An $m\times n$ matrix is exhibited in the form

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \text{ or in the form } \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

F is said to be the *field of scalars*. If, in particular, F be the field of real (complex) numbers, the matrix is said to be a *real (complex) matrix*. The element a_{ij} appearing in the ith row and jth column of the matrix is said to be the ijth element. The matrix is also denoted by the symbol $(a_{ij})_{m,n}$.

In an $m \times n$ matrix, if $m = 1$, the matrix is said to be a *row matrix*, and if $n = 1$, the matrix is said to be a *column matrix*.

If each element of an $m \times n$ matrix be 0, the null element of F, the matrix is said to be the *null matrix* or the *zero matrix* of order $m \times n$ and it is denoted by $O_{m,n}$. It is also denoted by O, when no confusion regarding its order arises.

An $n\times n$ matrix is said to be a *square matrix* of order n. The diagonal through the left hand top corner element of a square matrix is said to be the *principal diagonal* of the matrix and the elements in the principal diagonal are said to be the *diagonal elements* of the square matrix.

Very often capital letters $A, B, C, \ldots$ are used to denote a matrix.

Definitions.

1. **Equal matrices.** Two matrices A and B are said to be *equal* if A and B have the same order and their corresponding elements be equal. Thus if $A = (a_{ij})_{m,n}$ and $B = (b_{ij})_{m,n}$, then $A = B$ if and only if $a_{ij} = b_{ij}$ for $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$.

2. **Diagonal matrix.** A square matrix is said to be a *diagonal matrix* if the elements other than the diagonal elements be all zero.

Examples of a real diagonal matrix are

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The diagonal matrix $(d_{ij})_{n,n}$ is denoted by diag $(d_{11}, d_{22}, \ldots, d_{nn})$.

3. Scalar matrix. A diagonal matrix is said to be a *scalar matrix* if all the diagonal elements be the same scalar.

4. Identity (or unit) matrix. A scalar matrix whose diagonal elements are all 1, the identity element of the ground field F, is said to be an *identity* matrix (or a *unit* matrix). The identity matrix of order n is denoted by I_n.

Thus $I_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \ldots & & \ldots & \\ 0 & 0 & \ldots & 1 \end{pmatrix} = (\delta_{ij})_{n,n}$, where $\delta_{ij} = 1$, if $i = j$,

$\delta_{ij} = 0$, if $i \neq j$.

5. Triangular matrix.

A square matrix (a_{ij}) is said to be an *upper triangular* matrix if all the elements below the diagonal are 0. That is, $a_{ij} = 0$ if $i > j$.

A square matrix (a_{ij}) is said to be a *lower triangular* matrix if all the elements above the diagonal are 0. That is, $a_{ij} = 0$ if $i < j$.

A square matrix is said to be a *triangular* matrix if it is either upper triangular or lower triangular.

A triangular matrix $(a_{ij})_{n,n}$ is said to be *strictly triangular* if $a_{ii} = 0$ for $i = 1, 2, \ldots, n$.

Examples of a real triangular matrix are

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 4 & 1 & 6 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

A diagonal matrix is both upper triangular and lower triangular.

1.2. Algebraic operations on matrices.

We consider matrices over the same scalar field F.

1. Multiplication by a scalar. The product of an $m \times n$ matrix $A = (a_{ij})_{m,n}$ by a scalar c where $c \in F$, the field of scalars, is a matrix $B = (b_{ij})_{m,n}$ defined by $b_{ij} = ca_{ij}$, $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$ and is written as cA. Thus we have $c(a_{ij})_{m,n} = (ca_{ij})_{m,n}$.

Let A be an $m \times n$ matrix and c, d are scalars. Then the following results are obvious.

(i) $c(dA) = (cd)A$,

(ii) $0A = O_{m,n}$, 0 being the zero element of F,

(iii) $cO_{m,n} = O_{m,n}$,

(iv) $cI_n = \begin{pmatrix} c & 0 & \ldots & 0 \\ 0 & c & \ldots & 0 \\ \ldots & & \ldots & \\ 0 & 0 & \ldots & c \end{pmatrix}$,

(v) $1A = A$, 1 being the identity element of F.

The scalar matrix of order n whose diagonal elements are all c can be expressed as cI_n.

2. Addition. Two matrices A and B are said to be *conformable for addition* if they have the same order.

If $A = (a_{ij})_{m,n}$ and $B = (b_{ij})_{m,n}$, then their sum $A + B$ is the matrix $C = (c_{ij})_{m,n}$, where $c_{ij} = a_{ij} + b_{ij}$, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$.

If A and B be matrices of different orders, then $A + B$ is not defined.

Let A, B be $m \times n$ matrices and c, d are scalars. Then the following results are obvious.

(i) $c(A + B) = cA + cB$, (ii) $(c + d)A = cA + dA$.

Theorem 1.2.1. Matrix addition is commutative.

This says that if A and B be two matrices such that $A + B$ is defined, then $A + B = B + A$.

Proof. Let $A = (a_{ij})_{m,n}$ $B = (b_{ij})_{m,n}$.

Let $A + B = C = (c_{ij})_{m,n}$ and $B + A = D = (d_{ij})_{m,n}$.

$$\begin{aligned} \text{Then } c_{ij} &= a_{ij} + b_{ij} \\ &= b_{ij} + a_{ij}, \text{ since } a_{ij}, b_{ij} \in F, \text{ the ground field} \\ &= d_{ij}. \end{aligned}$$

Since C and D are of the same order and $c_{ij} = d_{ij}$, $C = D$.

That is, $A + B = B + A$. This completes the proof.

Theorem 1.2.2. Matrix addition is associative.

This says that if A, B, C be matrices such that the matrices $B+C$, $A+(B+C)$, $A+B$, $(A+B)+C$ are defined, then $A+(B+C) = (A+B)+C$.

Proof. Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{m,n}$, $C = (c_{ij})_{m,n}$.

Let $B + C = D = (d_{ij})_{m,n}$, $A + B = E = (e_{ij})_{m,n}$,

$A+D=P=(p_{ij})_{m,n}$, $E+C=Q=(q_{ij})_{m,n}$.
Then $d_{ij}=b_{ij}+c_{ij}$, $e_{ij}=a_{ij}+b_{ij}$, $p_{ij}=a_{ij}+d_{ij}$, $q_{ij}=e_{ij}+c_{ij}$.
$A+(B+C)=A+D=P=(p_{ij})_{m,n}$
and $(A+B)+C=E+C=Q=(q_{ij})_{m,n}$.

Therefore P and Q are matrices of the same order and

$$\begin{aligned} p_{ij}=a_{ij}+d_{ij} &= a_{ij}+(b_{ij}+c_{ij}) \\ &= (a_{ij}+b_{ij})+c_{ij}, \text{ since } a_{ij},b_{ij},c_{ij}\in F \\ &= e_{ij}+c_{ij}=q_{ij}. \end{aligned}$$

Hence $P=Q$, i.e., $A+(B+C)=(A+B)+C$.
This completes the proof.

Definition. Let $A=(a_{ij})_{m,n}$. Then the *negative* of A is defined to be a matrix B such that $A+B=O_{m,n}$. B is denoted by $-A$.

Thus $A+(-A)=O_{m,n}$ and $-A=(-a_{ij})_{m,n}$.

If A and B be matrices of the same order, then $A-B$ is defined by $A+(-B)$. If A and B be matrices of the same order and c, d are scalars, then the following results are obvious.

(i) $c(A-B)=cA-cB$, (ii) $(c-d)A=cA-dA$.

3. Multiplication of Matrices.

Two matrices A and B are said to be *conformable for the product* AB if the number of columns of A be equal to the number of rows of B. If $A=(a_{ij})_{m,n}$, $B=(b_{ij})_{n,p}$ then the product AB is a matrix of order $m\times p$ and $AB=C=(c_{ij})_{m,p}$ where $c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}$, $i=1,2,\ldots,m; j=1,2,\ldots,p$.

The ijth element of the product AB is obtained by multiplying the corresponding elements of the ith row of A and the jth column of B and adding the products.

If the number of columns of A be not equal to number of rows of B, then AB is not defined.

It is obvious that the products AB and BA are two distinct entities. Indeed, one of them may exist while the other may not.

For an $m\times n$ matrix A, in order that both AB and BA should exist, B must be of order $n\times m$. In this case, however, AB and BA are matrices of different orders. In order that both AB and BA should exist as matrices of the same order, both A and B must be square matrices of the same order.

In the product AB, A is said to be a *pre-multiplier* and B is said to be a *post-multiplier*.

Note. Matrix multiplication is not commutative. That is, for two matrices A and B, $AB\neq BA$, in general.

First of all, if we choose the orders of A and B to be $m\times n$ and $n\times m$ respectively so that the conformability conditions for both the products AB and BA are satisfied, then we observe that the orders of AB and BA are $m\times m$ and $n\times n$ respectively and therefore AB cannot be equal to BA.

In order that AB and BA may be equal, both of them must be of the same order and this requires that A and B must be square matrices of the same order. However if we choose the orders of A and B to be $n\times n$ and $n\times n$, then although AB and BA become matrices of the same order, they may not be equal, in general.

This can be shown by taking at random

$A=\begin{pmatrix}1&2\\2&3\end{pmatrix}$, $B=\begin{pmatrix}5&6\\4&2\end{pmatrix}$.

Here $AB=\begin{pmatrix}13&10\\22&18\end{pmatrix}$, $BA=\begin{pmatrix}17&28\\8&14\end{pmatrix}$.

In some special cases, however, $AB=BA$.

For example, let $A=\begin{pmatrix}1&2\\2&3\end{pmatrix}$, $B=\begin{pmatrix}1&0\\0&1\end{pmatrix}$.

Then $AB=\begin{pmatrix}1&2\\2&3\end{pmatrix}$, $BA=\begin{pmatrix}1&2\\2&3\end{pmatrix}$.

Definition. Two matrices A and B are said to *commute* with each other if $AB=BA$. Since $AB=BA$, A and B must be square matrices of the same order.

Some Examples of commuting matrices.

1. Let A be a square matrix. Then A commutes with A itself.

2. Let A be a square matrix of order n. Then A commutes with I_n, because $A.I_n=I_n.A=A$.

3. Let A be a square matrix of order n. Then A commutes with $O_{n,n}$, because $A.O_{n,n}=O_{n,n}.A=O_{n,n}$.

4. Let A be a square matrix of order n. Then A commutes with the scalar matrix cI_n, because $A.cI_n=cI_n.A=cA$.

Definition. Divisor of zero. A non-zero matrix A of order $m\times n$ is said to be a *divisor of zero* if there exists a non-zero matrix B of order $n\times p$ such that $AB=O_{m,p}$, or if there exists a non-zero matrix C of order $p \times m$ such that $CA = O_{p,n}$.

When $AB = O$, A is said to be a *left divisor of zero* and B is said to be a *right divisor of zero*.

Example.

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 6 & -4 \\ -3 & 2 \end{pmatrix}$,

$C = \begin{pmatrix} 4 & 2 & 6 \\ 6 & 3 & 9 \end{pmatrix}$, $D = \begin{pmatrix} 6 & 0 & 8 \\ 5 & 4 & 8 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. So A is a left divisor of zero and B is a right divisor of zero. $BA \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Also $BC = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

So B is a left divisor of zero and C is a right divisor of zero. CB is not defined.

$AC = \begin{pmatrix} 16 & 8 & 24 \\ 32 & 16 & 48 \end{pmatrix}$, $AD = \begin{pmatrix} 16 & 8 & 24 \\ 32 & 16 & 48 \end{pmatrix}$. $AC = AD$, but $C \neq D$. This happens because $A(C-D) = O$ does not imply $C - D = 0$.

Theorem 1.2.3. Matrix multiplication is associative.

This says that if A, B, C be three matrices such that the products $BC, A(BC), AB, (AB)C$ are defined, then $A(BC) = (AB)C$.

Proof. Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{n,p}$, $C = (c_{ij})_{p,q}$. Then $A(BC)$, $(AB)C$ are both defined.

Let $BC = D = (d_{ij})_{n,q}$, $AB = E = (e_{ij})_{m,p}$,

$AD = F = (f_{ij})_{m,q}$, $EC = G = (g_{ij})_{m,q}$.

Then $d_{ij} = \sum\limits_{r=1}^{p} b_{ir}c_{rj}$, $e_{ij} = \sum\limits_{s=1}^{n} a_{is}b_{sj}$, $f_{ij} = \sum\limits_{k=1}^{n} a_{ik}d_{kj}$, $g_{ij} = \sum\limits_{t=1}^{p} e_{it}c_{tj}$.

So $A(BC) = F = (f_{ij})_{m,q}$ and $(AB)C = G = (g_{ij})_{m,q}$.

$$f_{ij} = \sum_{k=1}^{n} a_{ik}d_{kj} = \sum_{k=1}^{n} a_{ik}\left(\sum_{r=1}^{p} b_{kr}c_{rj}\right) = \sum_{k=1}^{n}\sum_{r=1}^{p} a_{ik}b_{kr}c_{rj};$$

$$\text{and } g_{ij} = \sum_{t=1}^{p} e_{it}c_{tj} = \sum_{t=1}^{p}\left(\sum_{s=1}^{n} a_{is}b_{st}\right)c_{tj} = \sum_{t=1}^{p}\sum_{s=1}^{n} a_{is}b_{st}c_{tj}$$

$$= \sum_{s=1}^{n}\sum_{t=1}^{p} a_{is}b_{st}c_{tj}.$$

Therefore $f_{ij} = g_{ij}$.

Since F and G are matrices of the same order and $f_{ij} = g_{ij}$, $F = G$. That is, $A(BC) = (AB)C$. This completes the proof.

Since the products $A(BC)$ and $(AB)C$ are equal, each can be unambiguously written as ABC.

The product of four matrices A, B, C, D, taken in this order, can be written as $ABCD$, where A, B, C, D are of suitable orders so that $ABCD$ is defined.

In general, the product of n matrices $A_1, A_2, ..., A_n$, taken in this order, can be written as $A_1A_2...A_n$, where the order of each A_i is suitably chosen so that the product $A_1A_2...A_n$ is defined.

If A be a square matrix, then the products $AA, AAA, ...$ are each defined. They are denoted by $A^2, A^3, ...$ respectively.

If A be a square matrix, then the index laws

(i) $A^{m+n} = A^m.A^n$ and (ii) $(A^m)^n = A^{mn}$

hold, where m, n are positive integers.

Theorem 1.2.4. Matrix multiplication is distributive with respect to matrix addition.

(i) $A(B+C) = AB + AC$, provided either side is defined

(ii) $(B+C)A = BA + CA$, provided either side is defined.

Proof. (i) This says that if A, B, C be three matrices such that $B + C, A(B+C), AB, AC, AB + AC$ are defined then $A(B+C) = AB + AC$.

Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{n,p}$, $C = (c_{ij})_{n,p}$. Then $A(B+C)$ and $AB + AC$ are both defined.

Let $B + C = D = (d_{ij})_{n,p}$, $AD = E = (e_{ij})_{m,p}$, $AB = F = (f_{ij})_{m,p}$, $AC = G = (g_{ij})_{m,p}$, $F + G = H = (h_{ij})_{m,p}$.

Then $d_{ij} = b_{ij} + c_{ij}$, $e_{ij} = \sum\limits_{k=1}^{n} a_{ik}d_{kj}$, $f_{ij} = \sum\limits_{k=1}^{n} a_{ik}b_{kj}$,

$g_{ij} = \sum\limits_{k=1}^{n} a_{ik}c_{kj}$, $h_{ij} = f_{ij} + g_{ij}$.

$A(B+C) = AD = E = (e_{ij})_{m,p}$ and $AB + AC = F + G = H = (h_{ij})_{m,p}$.

$$e_{ij} = \sum_{k=1}^{n} a_{ik}d_{kj} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) \text{ and}$$

$$h_{ij} = f_{ij} + g_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}).$$

Hence $e_{ij} = h_{ij}$ and since E and H are of the same order, $E = H$.

That is, $A(B+C) = AB + AC$ and this completes the proof.

(ii) Proof left to the reader.

Worked Example.

1. If A and B be two square matrices of the same order, examine if $(A+B)^2 = A^2+2AB+B^2$.

Since A and B are square matrices, A^2 and B^2 are both defined. Since A and B are square matrices of the same order $A+B$ and AB are both defined.

$$\begin{aligned}(A+B)^2 &= (A+B)(A+B)\\ &= A(A+B)+B(A+B)\\ &= A^2+AB+BA+B^2.\end{aligned}$$

Since $AB \neq BA$, in general, $(A+B)^2 \neq A^2+AB+AB+B^2$, in general. That is, $(A+B)^2 \neq A^2+2AB+B^2$, in general.

However, if A and B are commuting matrices then $(A+B)^2 = A^2 + 2AB+B^2$.

For example, let

$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix}$. Then

$AB = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $BA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $AC = \begin{pmatrix} 10 & 5 \\ 24 & 11 \end{pmatrix}$, $CA = \begin{pmatrix} 8 & 19 \\ 6 & 13 \end{pmatrix}$.

Here $AB = BA$ and $(A+B)^2 = A^2+2AB+B^2$; $AC \neq CA$ and $(A+C)^2 \neq A^2+2AC+C^2$.

1.3. Transpose of a matrix.

Let A be an $m \times n$ matrix. Then the $n \times m$ matrix obtained by interchanging rows and columns of A is said to be the *transpose* of A and is denoted by A^t (or A^T).

Thus if $A = (a_{ij})_{m,n}$ then $A^t = B = (b_{ij})_{n,m}$, where $b_{ij} = a_{ji}$, $i = 1,2,...,n$; $j = 1,2,...,m$.

Theorem 1.3.1. $(A^t)^t = A$.

The proof is obvious.

Theorem 1.3.2. If A and B be two matrices such that $A+B$ is defined, then $(A+B)^t = A^t + B^t$.

Proof. Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{m,n}$. Then $A+B$ is defined.

The order of $A+B$ is $m \times n$ and the order of $(A+B)^t$ is $n \times m$.

The order of A^t is $n \times m$ and the order of B^t is $n \times m$.

Therefore the order of A^t+B^t is $n \times m$.

Thus the order of $(A+B)^t$ = the order of (A^t+B^t) ...... (i)

The ijth element of $(A+B)^t$

= the jith element of $(A+B)$

= the jith element of A + jith element of B

= the ijth element of A^t + ijth element of B^t

= the ijth element of (A^t+B^t) ... (ii)

From (i) and (ii) it follows that $(A+B)^t = A^t+B^t$.

Theorem 1.3.3. If c be a scalar, $(cA)^t = cA^t$.

Proof. Let $A = (a_{ij})_{m,n}$. Then the order of $(cA)^t$ is $n \times m$ and the order of cA^t is $n \times m$.

Thus the order of $(cA)^t$ = the order of cA^t ... (i)

The ijth element of $(cA)^t$

= the jith element of cA

= c(jith element of A)

= c(ijth element of A^t)

= the ijth element of cA^t ... (ii)

From(i) and (ii) it follows that $(cA)^t = cA^t$.

Corollary. If A and B be two matrices of the same order $(cA+dB)^t = cA^t+dB^t$, where c, d are scalars.

Theorem 1.3.4. If A and B be two matrices such that AB is defined, then $(AB)^t = B^tA^t$.

Proof. Let $A = (a_{ij})_{m,n}$, $B = (b_{ij})_{n,p}$. Then AB is defined.

The order of AB is $m \times p$. So the order of $(AB)^t$ is $p \times m$.

The order of B^t is $p \times n$, the order of A^t is $n \times m$.

So the order of B^tA^t is $p \times m$.

Thus the order of $(AB)^t$ = the order of B^tA^t ... (i)

The ijth element of $(AB)^t$

= the jith element of AB

= the sum of the products of corresponding elements of the jth row of A and the ith column of B

= the sum of the products of corresponding elements of the jth column of A^t and the ith row of B^t

= the sum of the products of corresponding elements of the ith row of B^t and the jth column of A^t

= the ijth element of B^tA^t ... (ii)

From (i) and (ii) it follows that $(AB)^t = B^tA^t$.

Note. If A, B, C be three matrices such that ABC is defined, then $(ABC)^t = C^t B^t A^t$.

In general, if $A_1, A_2, ..., A_n$ be n matrices such that the product $A_1 A_2 ... A_n$ is defined, then $(A_1 A_2 ... A_n)^t = A_n^t ... A_2^t A_1^t$.

1.4. Symmetric and skew symmetric matrices.

A square matrix A is said to be *symmetric* if $A = A^t$. Therefore $A = (a_{ij})$ is symmetric if $a_{ij} = a_{ji}$.

A square matrix A is said to be *skew symmetric* if $A = -A^t$. Therefore $A = (a_{ij})$ is skew symmetric if $a_{ij} = -a_{ji}$.

Examples of a symmetric matrix are

$$\begin{pmatrix} 1 & 3 & 5 \\ 3 & 0 & 7 \\ 5 & 7 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2+i & 3 \\ 2+i & i & 1-i \\ 3 & 1-i & 0 \end{pmatrix}, \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}.$$

Examples of a skew symmetric matrix are

$$\begin{pmatrix} 0 & 1 & -8 \\ -1 & 0 & 2 \\ 8 & -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2+i & 3i \\ -2-i & 0 & 2 \\ -3i & -2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Note. An $n \times n$ null matrix is both symmetric and skew symmetric.

Theorem 1.4.1. If A and B be two symmetric matrices of the same order then $A + B$ is symmetric.

Proof. $(A + B)^t = A^t + B^t = A + B$, since $A^t = A$, $B^t = B$.
This proves that $A + B$ is symmetric.

Theorem 1.4.2. If A and B be two symmetric matrices of the same order then AB is symmetric if and only if $AB = BA$.

Proof. Let AB be symmetric.

Then $AB = (AB)^t$
$= B^t A^t = BA$, since $B^t = B$, $A^t = A$.

Conversely, let $AB = BA$.

Then $(AB)^t = B^t A^t = BA$, since $B^t = B$, $A^t = A$
$= AB$, by the assumed condition.

Therefore AB is symmetric.
This completes the proof.

Theorem 1.4.3. If A be an $m \times n$ matrix, then the matrices AA^t and $A^t A$ are both symmetric.

AA^t and $A^t A$ are square matrices of order m and n respectively.

$(AA^t)^t = (A^t)^t A^t = AA^t$ and $(A^t A)^t = A^t (A^t)^t = A^t A$.
This shows that AA^t and $A^t A$ are both symmetric matrices.

Theorem 1.4.4. The diagonal elements of a skew symmetric matrix are all zero, provided the ground field F is of characteristic $\neq 2$.

Proof. Let $A = (a_{ij})$ be skew symmetric. Then $a_{ij} = -a_{ji}$.
In particular, $a_{ii} = -a_{ii}$ or, $2a_{ii} = 0$.

This implies $a_{ii} = 0$, provided the characteristic of the ground field F is not 2.

Note. The diagonal elements of a real (or complex) skew symmetric matrix are all zero.

Example. If the ground field F be a field of characteristic 2, then

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ is a skew symmetric matrix.}$$

Theorem 1.4.5. If A be a square matrix, then $A + A^t$ is symmetric and $A - A^t$ is skew symmetric.

Proof. $(A + A^t)^t = A^t + (A^t)^t = A^t + A = A + A^t$.
This shows that $A + A^t$ is symmetric.

$(A - A^t)^t = A^t - (A^t)^t = A^t - A = -(A - A^t)$.
This shows that $A - A^t$ is skew symmetric.

Theorem 1.4.6. A real (or complex) square matrix can be uniquely expressed as the sum of a symmetric matrix and a skew symmetric matrix.

Proof. Let A be a given matrix. Then A can be expressed as

$A = \frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$.

$[\frac{1}{2}(A + A^t)]^t = \frac{1}{2}[A^t + (A^t)^t] = \frac{1}{2}[A + A^t]$ and

$[\frac{1}{2}(A - A^t)]^t = \frac{1}{2}[A^t - (A^t)^t] = \frac{1}{2}[A^t - A] = -\frac{1}{2}[A - A^t]$.

This shows that $\frac{1}{2}(A + A^t)$ is a symmetric matrix and $\frac{1}{2}(A - A^t)$ is a skew symmetric mayrix.

Therefore A is expressed as the sum of a symmetric matrix and a skew symmetric matrix.

We now show that this decomposition is unique.

Let $A = P + Q$ where P is symmetric and Q is skew symmetric.
Then $A^t = P^t + Q^t = P - Q$.
We have $A + A^t = 2P$, $A - A^t = 2Q$.
So $P = \frac{1}{2}(A + A^t)$, $Q = \frac{1}{2}(A - A^t)$ and this proves the theorem.

Note. The theorem does not hold if the ground field F be of characteristic 2.

Worked Example (continued).

2. Express $A = \begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 1 & 6 & 8 \end{pmatrix}$ as the sum of a symmetric matrix and a skew symmetric matrix.

Let $A = P + Q$ where P is symmetric and Q is skew symmetric.
Then $A^t = P^t + Q^t = P - Q$.
We have $P = \frac{1}{2}(A + A^t)$, $Q = \frac{1}{2}(A - A^t)$.

$$P = \tfrac{1}{2}\left[\begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 1 & 6 & 8 \end{pmatrix} + \begin{pmatrix} 4 & 3 & 1 \\ 5 & 7 & 6 \\ 1 & 2 & 8 \end{pmatrix}\right] = \begin{pmatrix} 4 & 4 & 1 \\ 4 & 7 & 4 \\ 1 & 4 & 8 \end{pmatrix},$$

$$Q = \tfrac{1}{2}\left[\begin{pmatrix} 4 & 5 & 1 \\ 3 & 7 & 2 \\ 1 & 6 & 8 \end{pmatrix} - \begin{pmatrix} 4 & 3 & 1 \\ 5 & 7 & 6 \\ 1 & 2 & 8 \end{pmatrix}\right] = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

$$\text{Therefore } A = \begin{pmatrix} 4 & 4 & 1 \\ 4 & 7 & 4 \\ 1 & 4 & 8 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -2 \\ 0 & 2 & 0 \end{pmatrix}.$$

1.5. Block mutiplication of matrices.

If a natural number n be expressed as the sum of $k(\leq n)$ natural numbers $n_1, n_2, \ldots, n_k$ (repetitions allowed) as $n = n_1 + n_2 + \cdots + n_k$, then $(n_1, n_2, \ldots, n_k)$ is said to be a *partition* of n.

For example, $(1,1,3), (2,3), (1,4), (1,1,1,1,1)$ are different partitions of 5. $(1,2,2), (2,1,2), (2,2,1)$ are regarded as the same partition of 5.

Let $A = (a_{ij})_{m,n}, B = (b_{ij})_{n,p}$ be matrices over a field F. Let $A_1, A_2, \ldots, A_m$ be the m row matrices of A and $B_1, B_2, \ldots, B_p$ be the p column matrices of B. Then the product matrix AB can be expressed as

$$AB = \begin{pmatrix} A_1B_1 & A_1B_2 & \ldots & A_1B_p \\ A_2B_1 & A_2B_2 & \ldots & A_2B_p \\ \\ A_mB_1 & A_mB_2 & \ldots & A_mB_p \end{pmatrix}.$$

Block multiplication of matrices is a generalisation of this method.

Let $(m_1, m_2, \ldots, m_r)$ $(r \leq m)$, $(n_1, n_2, \ldots, n_k)$ $(k \leq n)$, $(p_1, p_2, \ldots, p_s)$ $(s \leq p)$ be arbitrary partitions of m, n, p respectively.

Then $m = m_1 + m_2 + \cdots + m_r$, $n = n_1 + n_2 + \cdots + n_k$, $p = p_1 + p_2 + \cdots + p_s$.

In the matrix A, let us draw $r-1$ horizontal lines after $m_1, m_1 + m_2, \ldots m_1 + m_2 + \cdots + m_{r-1}$ rows and $k-1$ vertical lines after $n_1, n_1 + n_2, \ldots n_1 + n_2 + \cdots + n_{k-1}$ columns.

Then A is divided into rk submatrices (rk blocks) $A_{11} = (a_{ij})_{m_1,n_1}$, $A_{12} = (a_{ij})_{m_1,n_2}, \ldots, A_{1k} = (a_{ij})_{m_1,n_k}, \ldots, A_{r1} = (a_{ij})_{m_r,n_1}, \ldots, A_{rk} = (a_{ij})_{m_r,n_k}$. A takes the form

$$A = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1k} \\ A_{21} & A_{22} & \ldots & A_{2k} \\ \\ A_{r1} & A_{r2} & \ldots & A_{rk} \end{pmatrix}. \ A_{ij} \text{ is an } m_i \times n_j \text{ matrix.}$$

In the matrix B, let us draw $k-1$ horizontal lines after $n_1, n_1 + n_2, \ldots n_1 + n_2 + \cdots + n_{k-1}$ rows and $s-1$ vertical lines after $p_1, p_1 + p_2, \ldots p_1 + p_2 + \cdots + p_{s-1}$ columns.

Then B is divided into ks submatrices (ks blocks) $B_{11} = (b_{ij})_{n_1,p_1}$, $B_{12} = (b_{ij})_{n_1,p_2}, \ldots, B_{1s} = (b_{ij})_{n_1,p_s}, \ldots, B_{k1} = (b_{ij})_{n_k,p_1}, \ldots, B_{ks} = (b_{ij})_{n_k,p_s}$. B takes the form

$$B = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1s} \\ B_{21} & B_{22} & \ldots & B_{2s} \\ \\ B_{k1} & B_{k2} & \ldots & B_{ks} \end{pmatrix}. \ B_{ij} \text{ is an } n_i \times p_j \text{ matrix.}$$

Clearly, the product $A_{it}B_{tj}$ is defined for $t = 1, 2, \ldots, k$.

This is to emphasise that we are to draw the vertical lines in A and the horizontal lines in B by using the same partition of n maintaining the same order.

Then the product matrix AB can be expressed as

$$AB = \begin{pmatrix} C_{11} & C_{12} & \ldots & C_{1s} \\ C_{21} & C_{22} & \ldots & C_{2s} \\ \\ C_{r1} & C_{r2} & \ldots & C_{rs} \end{pmatrix}, \text{ where } C_{ij} = \sum_{t=1}^{n} A_{it}B_{tj}.$$

Note. In particular, if we take $m = 1 + 1 + \cdots + 1$, $n = 1 + 1 + \cdots + 1$, $p = 1 + 1 + \cdots + 1$, then the block multiplication reduces to ordinary (element by element) multiplication of matrices.

Some particular cases of special interest.

Let $A = (a_{ij})_{m,n}, B = (b_{ij})_{n,p}$.

(i) $m = 1+1+\cdots+1$. n and p are not partitioned.

Then $A = \begin{pmatrix} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{pmatrix}$, where $A_i = (a_{i1}a_{i2}\ldots a_{in})$ = the ith row of A.

B is not partitioned. $AB = \begin{pmatrix} A_1B \\ A_2B \\ \cdot \\ \cdot \\ A_mB \end{pmatrix}$.

The rows of the product AB are $A_1B, A_2B, \ldots, A_mB$.

(ii) $p = 1+1+\cdots+1$. m and n are not partitioned.

Then A is not partitioned.

$B = (B_1\ B_2\ \ldots\ B_p)$, where $B_i = \begin{pmatrix} b_{1i} \\ b_{2i} \\ \cdot \\ \cdot \\ b_{ni} \end{pmatrix}$, the ith column of B.

$AB = (AB_1\ AB_2\ \ldots\ AB_p)$.

The columns of the product AB are $AB_1, AB_2, \ldots, AB_p$.

Example. Let $A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 4 \\ 3 & 1 \\ 0 & 0 \end{pmatrix}$.

Here $m = 3, n = 3, p = 2$. Let us take partitions of m, n as $m = 1+2, n = 2+1$.

Then $A = \begin{pmatrix} P & Q \\ I_2 & R \end{pmatrix}$, $B = \begin{pmatrix} S \\ O \end{pmatrix}$ in block form, where $P = \begin{pmatrix} 2 & 1 \end{pmatrix}, Q = \begin{pmatrix} 2 \end{pmatrix}, R = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, S = \begin{pmatrix} 1 & 4 \\ 3 & 1 \end{pmatrix}$.

By the method discussed above,

$$AB = \begin{pmatrix} P.S + Q.O \\ I_2.S + R.O \end{pmatrix} = \begin{pmatrix} PS \\ S \end{pmatrix} = \begin{pmatrix} 5 & 9 \\ 1 & 4 \\ 3 & 1 \end{pmatrix}.$$

Let us take partitions of m, n, p as $m = 1+2, n = 2+1, p = 1+1$.

Then $A = \begin{pmatrix} P & Q \\ I_2 & R \end{pmatrix}$, $B = \begin{pmatrix} S & T \\ O & O \end{pmatrix}$ in block form, where $P = \begin{pmatrix} 2 & 1 \end{pmatrix}, Q = \begin{pmatrix} 2 \end{pmatrix}, R = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, S = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, T = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.

By the method discussed above,

$$AB = \begin{pmatrix} P.S + Q.O & P.T + Q.O \\ I_2.S + R.O & I_2.T + R.O \end{pmatrix} = \begin{pmatrix} PS & PT \\ S & T \end{pmatrix} = \begin{pmatrix} 5 & 9 \\ 1 & 4 \\ 3 & 1 \end{pmatrix}.$$

1.6. Some special matrices.

Let $E_{ij} (i = 1, 2, \ldots, n; j = 1, 2, \ldots, n)$ be the real $n \times n$ matrix with the ij-th element 1 and all other elements 0.

For example, the 2×2 matrices $E_{11}, E_{12}, E_{21}, E_{22}$ are $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ respectively.

Clearly, an $n \times n$ matrix $A = (a_{ij})_{nn}$ can be expressed as the sum of n^2 matrices $E_{11}, E_{12}, \ldots, E_{nn}$ as

$$\begin{aligned} A = {} & a_{11}E_{11} + a_{12}E_{12} + \cdots + a_{1n}E_{1n} \\ & + a_{21}E_{21} + a_{22}E_{22} + \cdots + a_{2n}E_{2n} \\ & \ldots \quad \ldots \quad \ldots \\ & + a_{n1}E_{n1} + a_{n2}E_{n2} + \cdots + a_{nn}E_{nn}. \end{aligned}$$

For example, the 2×2 matrix $A = \begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$ can be expressed as $A = 2E_{11} + 3E_{12} + 5E_{21} + 4E_{22}$.

Theorem 1.6.1. Let $A = (a_{ij})_{nn}$. Then $E_{pq}AE_{rs} = a_{qr}E_{ps}$, where E_{ij} is an $n \times n$ matrix with ij-th element 1 and all other elements 0.

Proof. Each row of E_{pq} other than the p-th is a zero row. Therefore every row in the product matrix $E_{pq}A$ other than the p-th is a zero row.

Let the p-th row of E_{pq} be $(e_{p1}\ e_{p2}\ldots e_{pn})$. Then $e_{pq} = 1$ and $e_{pi} = 0$, for $i = 1, 2, \ldots, q-1, q+1, \ldots, n$.

Let $B = E_{pq}A = (b_{ij})_{nn}$. Then $b_{pk} = e_{p1}a_{1k} + e_{p2}a_{2k} + \cdots + e_{pn}a_{nk} = a_{qk}$.

Therefore the p-th row of $B (= E_{pq}A)$ is $(a_{q1}\ a_{q2}\cdots, a_{qn})$.

Each column of E_{rs} other than the s-th is a zero column. Therefore every column in the product matrix $BE_{rs} (= E_{pq}AE_{rs})$ other than the s-th is a zero column.

Let the s-th column of E_{rs} be $\begin{pmatrix} e_{1s} \\ \cdot \\ e_{ns} \end{pmatrix}$. Then $e_{rs} = 1$, $e_{is} = 0$ for $i = 1, 2, \ldots, r-1, r+1, \ldots, n$.

Then $e_{pq} = 1$ and $e_{pi} = 0$, for $i = 1, 2, \ldots, q-1, q+1, \ldots, n$.

Let $C = BE_{rs} = (c_{ij})_{nn}$. Since the only non-zero row of B is the p-th row and the only non-zero column of E_{rs} is the s-th column, the only non-zero element of the matrix $C(= E_{pq}AE_{rs})$ is the ps-th element.

$c_{ps} = b_{p1}e_{1s} + b_{p2}e_{2s} + \cdots + b_{pn}e_{ns} = a_{q1}e_{1s} + a_{q2}e_{2s} + \cdots + a_{qn}e_{ns} = a_{qr}$.

All other elements of the matrix C are 0.

Therefore $C(= E_{pq}AE_{rs}) = a_{qr}E_{ps}$.

Corollary. $E_{pq}E_{rs} = E_{ps}$ if $q = r$
$= 0$ if $q \neq r$.

Proof. $E_{pq}E_{rs} = E_{pq}I_nE_{rs}$.

But $I_n = (\delta_{ij})_{nn}$, where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

Therefore $E_{pq}E_{rs} = E_{pq}I_nE_{rs} = \delta_{qr}E_{ps} = E_{ps}$ if $q = r$
$= 0$ if $q \neq r$.

Deductions.

(i) $E_{pq}E_{qs} = E_{ps}$. (ii) $E_{pp}E_{pp} = E_{pp}$. (iii) $E_{pp}E_{qq} = 0$ if $p \neq q$.

Definition. A square matrix A is said to be *idempotent* if $A^2 = A$.

From deduction (ii) above it follows that the matrices $E_{11}, E_{22}, \ldots, E_{nn}$ are idempotent matrices.

Exercises 1

1. Find the matrices A and B if

(i) $2A + 3B = I_2$, $A + B = 2A^t$;

(ii) $2A + 3B = \begin{pmatrix} 8 & 3 \\ 7 & 6 \end{pmatrix}$, $A + B^t = \begin{pmatrix} 3 & 1 \\ 3 & 3 \end{pmatrix}$;

(iii) $2A + B^t = \begin{pmatrix} 2 & 5 \\ 10 & 2 \end{pmatrix}$, $2B + A^t = \begin{pmatrix} 1 & 8 \\ 4 & 1 \end{pmatrix}$.

2. If A and B be commuting matrices prove that A^t and B^t commute.

3. Three matrices A, B, C are such that $A \neq 0$ and $AB = AC$. Can you conclude $B = C$?

4. Three $n \times n$ matrices A, B, C are such that $AB = I_n$ and $BC = I_n$. Prove that $A = C$.

5. If A and B are square matrices of the same order, does the equality $(A + B)(A - B) = A^2 - B^2$ hold good? Give reasons.

6. Let A and B be $n \times n$ matrices and $BB^t = I_n$. Prove that $(B^tAB)^p = B^tA^pB$ where p is a positive integer.

7. Prove that the equality $AB - BA = I_2$ cannot hold, whatever the real 2×2 matrices A and B may be.

8. If $A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ show that $A^2 - 2A + I_2 = O$. Hence find A^{50}.

9. If $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ prove that $A^n = \begin{pmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$ for all $n \in \mathbb{N}$.

10. (i) If $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ show that $A^2 - 4A - 5I_3 = O$. Hence obtain a matrix B such that $AB = I_3$.

(ii) If $A = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix}$ show that $A^3 - 6A - 9I_3 = O$. Hence obtain a matrix B such that $BA = I_3$.

11. Find all non-null real matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $A^2 = O$.

12. Find all real matrices $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $A^2 = I_2$.

13. Find all 2×2 matrices that commute with the real matrix

(i) $\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$, (ii) $\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$,

(iii) $\begin{pmatrix} a & b \\ c & a \end{pmatrix}$, $bc \neq 0$, (iv) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $bc \neq 0$.

14. Find a 2×2 diagonal matrix D such that $D^2 - 5D + 6I_2 = O$.

15. If A be an $n \times n$ matrix and $D = \text{diag}(d_1, d_2, \ldots, d_n)$, show that

(i) the ith row of DA is d_iR_i, where R_i is the ith row matrix of A;

(ii) the ith column of AD is d_iC_i, where C_i is the ith column matrix of A.

Deduce that every square matrix commutes with a scalar matrix of the same order.

16. (i) If A be a symmetric matrix of order m and P be an $m \times n$ matrix, prove that P^tAP is a symmetric matrix.

(ii) If A be a real skew symmetric matrix of order n and P be a real $n \times 1$ matrix, prove that $P^tAP = O$.

17. Express A as $P+Q$ where P is a symmetric matrix and Q is a skew symmetric matrix.

(i) $A = \begin{pmatrix} 1 & 3 & 4 \\ 7 & 2 & 6 \\ 2 & 8 & 1 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 5 & 3 & 2 & 8 \\ 4 & 2 & 1 & 5 \\ 1 & 6 & 7 & 0 \end{pmatrix}$.

18. (i) If A be an idempotent matrix of order n, show that the matrix $I_n - A$ is also idempotent.

(ii) If $AB = B$ and $BA = A$, show that the matrices A and B are both idempotent.

(iii) If A and B are matrices of order n such that $A+B = I_n$ and $AB = 0$, show that A and B are both idempotent matrices.

19. If A be an idempotent matrix of order p, prove that
$$(I_p + A)^n = I_p + (2^n - 1)A \quad \text{for all } n \in \mathbb{N}.$$

20. The *trace* of a square matrix A, denoted by $tr\ A$, is the sum of the principal diagonal elements of A.

If A and B are square matrices of the same order, prove that

(i) $tr\ A + tr\ B = tr\ (A+B)$, (ii) $tr\ A^t = tr\ A$, (iii) $tr\ (BA) = tr\ (AB)$.

Establish that there cannot exist real 3×3 matrices A and B such that
$$AB - BA = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

21. A is an $m \times n$ real matrix. Prove that $tr\ (AA^t) \geq 0$, the equality occurs if A is a null matrix.

22. If $A = (a_{ij})$ and $B = (b_{ij})$ be upper triangular matrices of the same order, prove that AB is upper triangular with $a_{ii}b_{ii}$ as the diagonal elements.

Deduce that the product of two lower triangular matrices of the same order is lower triangular.

[**Hint.** Let $AB = (c_{ij})$ and let $i > j$. $c_{ij} = \sum_{r=1}^{n} a_{ir}b_{rj} = \sum_{r=1}^{i-1} a_{ir}b_{rj} + \sum_{r=i}^{n} a_{ir}b_{rj} = 0$, since $a_{ir} = 0$ for $r = 1, 2, ..., i-1$ and $b_{rj} = 0$ for $r = i, i+1, ..., n$.

$c_{ii} = \sum_{r=1}^{n} a_{ir}b_{ri} = \sum_{r=1}^{i-1} a_{ir}b_{ri} + a_{ii}b_{ii} + \sum_{r=i+1}^{n} a_{ir}b_{ri} = a_{ii}b_{ii}$, since $a_{ir} = 0$ for $r = 1, 2, ..., i-1$ and $b_{ri} = 0$ for $r = i+1, ..., n$.]

23. If A be the real matrix $\begin{pmatrix} 0 & c \\ 0 & 0 \end{pmatrix}$, show that $A^2 = O$. Deduce that if B be the real matrix $\begin{pmatrix} 0 & 0 & a & b \\ 0 & 0 & c & d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, then $B^2 = O$.

2. Determinants

2.1. Introduction.

Let us consider the set S of all square matrices of order n whose elements belong to a field F. F is said to be the field of scalars.

A mapping $f : S \to F$ which assigns to each matrix A in S a scalar c in F is to be a *scalar function* on S.

We now introduce a scalar function, called the *determinant function*, on the set S.

Definition. A determinant function $f : S \to F$ is a scalar function on the set S of all $n \times n$ matrices over the field F such that if $A = (a_{ij}) \in S$, then $f(A)$, or det A, or det (a_{ij}) is a scalar belonging to F and is defined by

det A = det $(a_{ij}) = \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)} \cdots a_{n\phi(n)}$, where ϕ is a permutation on $\{1, 2, \ldots, n\}$ and $sgn\ \phi = 1$ or -1 according as the permutation $\phi = \begin{pmatrix} 1 & 2 & \ldots & n \\ \phi(1) & \phi(2) & \ldots & \phi(n) \end{pmatrix}$ is even or odd.

det A is said to be a determinant of order n and is denoted by the symbol $\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$, or shortly by $|a_{ij}|_n$.

The summation $\sum_{\phi}$ is said to be the *expansion* of det A. It contains $n!$ terms as there are $n!$ permutations on the set $\{1, 2, \ldots, n\}$. As there are $\frac{1}{2}n!$ even and $\frac{1}{2}n!$ odd permutations on the set $\{1, 2, \ldots, n\}$, $\sum_{\phi}$ contains $\frac{1}{2}n!$ positive terms and $\frac{1}{2}n!$ negative terms.

Each term is a product of n elements. In each term the first suffices (row suffices) of the elements appear in their natural order and the second suffices appear in a permutation of $1, 2, \ldots, n$. So each term contains one element from each row and one element from each column of A.

The expansion contains the term $+a_{11}a_{22}\ldots a_{nn}$ comprising all the diagonal elements of A. This term is called the *leading term* in the expansion of det A.

Note. If (d_{ij}) be a diagonal matrix of order n, then det $(d_{ij}) = d_{11}d_{22}\ldots d_{nn}$.

Unless otherwise stated we shall always take the ground field F to be the field $\mathbb{R}$ (the field of all real numbers), or the field $\mathbb{C}$ (the field of all complex numbers). The theorems and properties of determinants of matrices over $\mathbb{R}$ (or $\mathbb{C}$) will hold good equally for matrices over the general field with some possible exceptions. The exceptional cases will be mentioned in proper places.

Examples.

1. Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

Then det $A = \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}$, where ϕ is a permutation on the set $\{1, 2\}$.

There are two such permutations.

$\phi_1 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$, $\phi_2 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. ϕ_1 is even and ϕ_2 is odd.

So det $A = sgn\ \phi_1\ a_{11}a_{22} + sgn\ \phi_2\ a_{12}a_{21} = a_{11}a_{22} - a_{12}a_{21}$.

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

det $A = \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}a_{3\phi(3)}$, where ϕ is a permutation on the set $\{1, 2, 3\}$. There are six such permutations.

$\phi_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$, $\phi_2 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$, $\phi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$,

$\phi_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, $\phi_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, $\phi_6 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$.

ϕ_1 is even, ϕ_2 is odd, ϕ_3 is odd, ϕ_4 is even, ϕ_5 is even, ϕ_6 is odd.

So det $A = sgn\ \phi_1\ a_{11}a_{22}a_{33} + sgn\ \phi_2\ a_{11}a_{23}a_{32} + sgn\ \phi_3\ a_{12}a_{21}a_{33} + sgn\ \phi_4\ a_{12}a_{23}a_{31} + sgn\ \phi_5\ a_{13}a_{21}a_{32} + sgn\ \phi_6\ a_{13}a_{22}a_{31} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$.

3. The expansion of $\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$ contains terms like $a_{11}a_{23}a_{34}a_{42}, a_{12}a_{21}a_{34}a_{43}, a_{14}a_{23}a_{31}a_{42}$ with proper sign.

The permutation $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}$ is a cycle of length 3 and so it is even. Therefore the sign of the term $a_{11}a_{23}a_{34}a_{42}$ is $+$.

The permutation $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$ is the product of two cycles $(1, 2)$, $(3, 4)$ and so it is even. Therefore the sign of the term $a_{12}a_{21}a_{34}a_{43}$ is $+$.

The permutation $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$ is a cycle of length 4 and so it is odd. Therefore the sign of the term $a_{14}a_{23}a_{31}a_{42}$ is $-$.

Equivalent Expression for the expansion of det (a_{ij}).

We have det $(a_{ij}) = \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)} \cdots a_{n\phi(n)}$.

Let ψ be the permutation on the set $\{1, 2, \ldots, n\}$ such that $\psi\phi = I$, I being the the identity permutation. Then $\psi = \phi^{-1}$.

As ϕ changes through all permutations on the set $\{1, 2, \ldots, n\}$, so does ψ.

Let $\phi(i) = j$. Then $i = \phi^{-1}(j) = \psi(j)$. So $a_{i\phi(i)} = a_{\psi(j)j}$.

As i runs through the set $\{1, 2, \ldots, n\}$, j does so.

Therefore $a_{1\phi(1)}a_{2\phi(2)} \cdots a_{n\phi(n)} = a_{\psi(1)1}a_{\psi(2)2} \cdots a_{\psi(n)n}$.

Since $\psi\phi = I$ and I is even, ψ and ϕ are both odd or both even. So we have $sgn\ \phi = sgn\ \psi$

Hence det $(a_{ij}) = \sum_{\psi} sgn\ \psi\ a_{\psi(1)1}a_{\psi(2)2} \cdots a_{\psi(n)n}$.

Note. In each term in the expansion the second suffices (column suffices) appear in their natural order.

Theorem 2.1.1. det A = det A^t, where A is an $n \times n$ matrix.

Proof. Let $A = (a_{ij})$ and let $A^t = B = (b_{ij})$. Then $b_{ij} = a_{ji}$.

$$\begin{aligned} \text{det } B &= \sum_{\phi} sgn\ \phi\ b_{1\phi(1)}b_{2\phi(2)} \cdots b_{n\phi(n)} \\ &= \sum_{\phi} sgn\ \phi\ a_{\phi(1)1}a_{\phi(2)2} \cdots a_{\phi(n)n}. \end{aligned}$$

Let ψ be a permutation on the set $\{1,2,\ldots,n\}$ such that $\psi\phi = I$, I being the identity permutation. Then $\psi = \phi^{-1}$. As ϕ changes through all permutations on the set $\{1,2,\ldots,n\}$, so does ψ.

Let $\phi(i) = j$. Then $i = \phi^{-1}(j) = \psi(j)$. Therefore $a_{\phi(i)i} = a_{j\psi(j)}$.

As i runs through the set $\{1,2,\ldots,n\}$, j does so.

Hence $a_{\phi(1)1}a_{\phi(2)2}\ldots a_{\phi(n)n} = a_{1\psi(1)}a_{2\psi(2)}\ldots a_{n\psi(n)}$.

Since $\psi\phi = I$ and I is even, ψ and ϕ are both odd or both even.

So we have $sgn\ \phi = sgn\ \psi$ and

$$\begin{aligned} \det B &= \sum_{\phi} sgn\ \phi\ a_{\phi(1)1}a_{\phi(2)2}\ldots a_{\phi(n)n} \\ &= \sum_{\psi} sgn\ \psi\ a_{1\psi(1)}a_{2\psi(2)}\ldots a_{n\psi(n)} \\ &= \det A. \end{aligned}$$

Note. As a result of this theorem, we can say that a theorem which holds for some row operations on A, also holds equally well when corresponding column operations are made on A.

Theorem 2.1.2. The interchange of two rows (columns) of an $n \times n$ matrix A changes the sign of det A.

Proof. Let $A = (a_{ij})$ and let $B = (b_{ij})$ is obtained from A by interchanging the rth and the sth row of A. Without loss of generality, we assume $r < s$.

Then $b_{ij} = a_{ij},\ i \neq r,\ i \neq s$

$b_{rj} = a_{sj},\ b_{sj} = a_{rj}$.

$$\begin{aligned} \det B &= \sum_{\phi} sgn\ \phi\ b_{1\phi(1)}b_{2\phi(2)}\ldots b_{r\phi(r)}\ldots b_{s\phi(s)}\ldots b_{n\phi(n)} \\ &= \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}\ldots a_{s\phi(r)}\ldots a_{r\phi(s)}\ldots a_{n\phi(n)}. \end{aligned}$$

Let $\lambda = \begin{pmatrix} 1 & 2 & \ldots & r & \ldots & s & \ldots & n \\ 1 & 2 & \ldots & s & \ldots & r & \ldots & n \end{pmatrix}$. Then λ is a transposition interchanging r and s and $Sgn\ \lambda = -1$.

Let $\phi\lambda = \psi$. As ϕ changes through all permutations on the set $\{1,2,\ldots,n\}$, ψ does so, because $\phi_1\lambda = \phi_2\lambda \Rightarrow \phi_1 = \phi_2$.

$$\begin{aligned} \psi = \phi\lambda &= \begin{pmatrix} 1 & \ldots & r & \ldots & s & \ldots & n \\ \phi(1) & \ldots & \phi(r) & \ldots & \phi(s) & \ldots & \phi(n) \end{pmatrix} \begin{pmatrix} 1 & \ldots & r & \ldots & s & \ldots & n \\ 1 & \ldots & s & \ldots & r & \ldots & n \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & \ldots & r & \ldots & s & \ldots & n \\ \phi(1) & \phi(2) & \ldots & \phi(s) & \ldots & \phi(r) & \ldots & \phi(n) \end{pmatrix}. \end{aligned}$$

So $\psi(i) = \phi(i),\ i \neq r,\ i \neq s$

$\psi(r) = \phi(s),\ \psi(s) = \phi(r)$.

Since λ is odd, ψ is even or odd according as ϕ is odd or even.

So we have $sgn\ \psi = -sgn\ \phi$.

$$\begin{aligned} \text{Hence } \det B &= \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}\ldots a_{s\phi(r)}\ldots a_{r\phi(s)}\ldots a_{n\ ph(n)}. \\ &= -\sum_{\psi} sgn\ \psi\ a_{1\psi(1)}a_{2\psi(2)}\ldots a_{r\psi(r)}\ldots a_{s\psi(s)}\ldots a_{n\psi(n)} \\ &= -\det A. \end{aligned}$$

Note. When two columns of A are interchanged, the theorem can be proved independently by considering the equivalent form of expansion of det A.

Corollary. If a permutation of rows (columns) be applied on A, det A does not or does change its sign according as the permutation is even or odd. This follows from the fact that an even (odd) permutation is the product of an even (odd) number of transpositions and a transposition effects an interchange of two rows (columns) and thereby changes the sign of det A.

Theorem 2.1.3. In an $n \times n$ matrix A, if two rows (columns) be identical then det $A = 0$.

Proof. Let the rth and the sth row of A be identical rows and let $\Delta =$ det A. If now the rth and the sth rows of A be interchanged then det A changes its sign, by Theorem 2.1.2, although A remains the same matrix.

Consequently, $-\Delta =$ det A.

Thus $\Delta = -\Delta$, giving $2\Delta = 0$.

This implies $\Delta = 0$, i.e., det $A = 0$.

Note. If the ground field F be of characteristic 2, Δ being a scalar belonging to F, $2\Delta = 0$ does not necessarily imply $\Delta = 0$. So the argument given in the proof breaks down if F be a field of characteristic 2. However, the theorem can be proved independently and it remains valid in all cases.

Independent proof.

Let the rth row and the sth row of $A = (a_{ij})$ be identical.

Then $a_{rj} = a_{sj}$.

det $A = \sum_{\mu} sgn\ \mu\ a_{1\mu(1)}a_{2\mu(2)}\ldots a_{n\mu(n)}$ where μ is a permutation on the set $\{1,2,\ldots,n\}$.

Let $\lambda = \begin{pmatrix} 1 & 2 & \dots & r & \dots & s & \dots & n \\ 1 & 2 & \dots & s & \dots & r & \dots & n \end{pmatrix}$. Then λ is a transposition interchanging r and s. Therefore λ is odd.

Let ϕ be an even permutation on $\{1, 2, \dots, n\}$. Then $\phi\lambda$ is odd since λ is odd. As ϕ changes through all even permutations, $\phi\lambda$ changes through all odd permutations. Let $\phi\lambda = \psi$.

Then $\sum\limits_{\mu}$ can be expressed as $\sum\limits_{\phi} + \sum\limits_{\psi}$.

$\sum\limits_{\phi}$ contains $\frac{1}{2}n!$ positive terms and $\sum\limits_{\psi}$ contains $\frac{1}{2}n!$ negative terms in the expansion of det A.

Let ϕ_1 be an even permutation and let $\phi_1\lambda = \psi_1$.

The sum of two terms in $\sum\limits_{\mu}$ corresponding to ϕ_1 and ψ_1 is

$$sgn\ \phi_1 a_{1\phi_1(1)} a_{2\phi_1(2)} \cdots a_{r\phi_1(r)} \cdots a_{s\phi_1(s)} \cdots a_{n\phi_1(n)} + sgn\ \psi_1 a_{1\psi_1(1)} a_{2\psi_1(2)} \cdots a_{r\psi_1(r)} \cdots a_{s\psi_1(s)} \cdots a_{n\psi_1(n)} \quad \dots \dots \text{(i)}$$

$\phi_1(i) = \psi(i)$, $i \neq r$, $i \neq s$.

$\phi_1(r) = \psi_1(s)$, $\phi_1(s) = \psi_1(r)$; $sgn\ \psi_1 = -sgn\ \phi_1$.

The sum (i) reduces to

$$sgn\ \phi_1 a_{1\phi_1(1)} a_{2\phi_1(2)} \cdots a_{r\phi_1(r)} \cdots a_{s\phi_1(s)} \cdots a_{n\phi_1(n)} - sgn\ \phi_1 a_{1\phi_1(1)} a_{2\phi_1(2)} \cdots a_{r\phi_1(s)} \cdots a_{s\phi_1(r)} \cdots a_{n\phi_1(n)} = 0, \text{ since } a_{rj} = a_{sj}.$$

As ϕ_1 changes through the whole set of even permutations, ψ_1 changes through the whole set of odd permutations.

If the terms $\sum\limits_{\mu}$ are paired as above, the sum of each pair is zero and this proves det $A = 0$.

An $n \times n$ matrix $A = (a_{ij})$ has n rows and n columns. Each row can be considered as a row matrix containing n elements.

Let $R_1, R_2, \dots, R_n$ be the rows of A. Then

$$\begin{aligned} R_1 &= (a_{11}\ a_{12} \dots a_{1n}) \\ R_2 &= (a_{21}\ a_{22} \dots a_{2n}) \\ &\dots \quad \dots \\ R_n &= (a_{n1}\ a_{n2} \dots a_{nn}). \end{aligned}$$

Two rows R_i and R_j are said to be *equal* if the row matrices R_i and R_j be equal, i.e., if $a_{i1} = a_{j1}$, $a_{i2} = a_{j2}, \dots, a_{in} = a_{jn}$.

Addition of two rows R_i and R_j is defined by

$$R_i + R_j = (a_{i1} + a_{j1} \quad a_{i2} + a_{j2} \quad \dots \quad a_{in} + a_{jn}).$$

Multiplication of the row R_i by a scalar c is defined by

$$cR_i = (ca_{i1} \quad ca_{i2} \dots \quad ca_{in}).$$

The sum $c_1R_1 + c_2R_2 + \dots + c_rR_r$, where $c_1, c_2, \dots, c_r$ are scalars, is said to be a *linear combination* of the rows $R_1, R_2, \dots, R_r$.

In a similar manner the operations can be defined on the columns $C_1, C_2, \dots, C_n$ of the matrix A.

Theorem 2.1.4. In an $n \times n$ matrix A, if a row (column) be multiplied by a scalar c then det A is multiplied by c.

Proof. ~~Case 1.~~ Let $c = 0$. Then the theorem is trivially true since det $A = 0$ when A has a zero row.

Case 2. Let $A = (a_{ij})$ and let $B = (b_{ij})$ be obtained from A by multiplying its rth row by a non-zero scalar c.

Then $b_{ij} = a_{ij}$, $i \neq r$; $b_{rj} = ca_{rj}$; and

$$\begin{aligned} \det B &= \sum_{\phi} sgn\ \phi\, b_{1\phi(1)} b_{2\phi(2)} \dots b_{r\phi(r)} \dots b_{n\phi(n)} \\ &= \sum_{\phi} sgn\ \phi\, a_{1\phi(1)} a_{2\phi(2)} \dots (ca_{r\phi(r)}) \dots a_{n\phi(n)} \\ &= c\sum_{\phi} sgn\ \phi\, a_{1\phi(1)} a_{2\phi(2)} \dots a_{r\phi(r)} \dots a_{n\phi(n)} \\ &= c \det A. \end{aligned}$$

Corollary. If a row (column) of A be a scalar multiple of another row (column) then det $A = 0$.

Theorem 2.1.5. In an $n \times n$ matrix (a_{ij}) if each element in a particular row, say the rth, be expressed as the sum of two terms as $a_{rj} = a'_{rj} + a''_{rj}$ then det (a_{ij}) = det (b_{ij})+ det (c_{ij}), where

$b_{ij} = a_{ij}$, $i \neq r$; $b_{rj} = a'_{rj}$ and $c_{ij} = a_{ij}$, $i \neq r$; $c_{rj} = a''_{rj}$.

i.e., $$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & & \dots & \\ a'_{r1} + a''_{r1} & a'_{r2} + a''_{r2} & \dots & a'_{rn} + a''_{rn} \\ \dots & & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & \dots & \\ a'_{r1} & a'_{r2} & \dots & a'_{rn} \\ & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & \dots & \dots & \\ a''_{r1} & a''_{r2} & \dots & a''_{rn} \\ & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Proof. We have

$$\begin{aligned}
\det(a_{ij}) &= \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}\cdots a_{r\phi(r)}\cdots a_{n\phi(n)} \\
&= \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}\cdots(a'_{r\phi(r)} + a''_{r\phi(r)})\cdots a_{n\phi(n)} \\
&= \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}\cdots a'_{r\phi(r)}\cdots a_{n\phi(n)} + \\
&\quad \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)}\cdots a''_{r\phi(r)}\cdots a_{n\phi(n)} \\
&= \sum_{\phi} sgn\ \phi\ b_{1\phi(1)}b_{2\phi(2)}\cdots b_{r\phi(r)}\cdots b_{n\phi(n)} + \\
&\quad \sum_{\phi} sgn\ \phi\ c_{1\phi(1)}c_{2\phi(2)}\cdots c_{r\phi(r)}\cdots c_{n\phi(n)} \\
&= \det(b_{ij}) + \det(c_{ij}).
\end{aligned}$$

This completes the proof.

Note. The theorem also holds if the statement be made for a particular column.

Corollary. If a_{rj} be expressed as the sum of n terms

$$a_{rj}^{(1)} + a_{rj}^{(2)} + \cdots + a_{rj}^{(n)},$$

then $\det(a_{ij})$ can be expressed as the sum of the determinants of n matrices.

Theorem 2.1.6. In an $n \times n$ matrix A, if a scalar multiple of one row (column) be added to another row (column) then $\det A$ remains unchanged.

Proof. Let $A = (a_{ij})$ and c be a scalar. Let c times the sth row of A be added to its rth row and let $B = (b_{ij})$ be the resulting matrix.

Then $b_{ij} = a_{ij}$, $i \neq r$; $b_{rj} = a_{rj} + ca_{sj}$; and

$$\det B = \begin{vmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & & \\ b_{r1} & b_{r2} & \cdots & b_{rn} \\ \cdots & \cdots & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \\ a_{r1} + ca_{s1} & a_{r2} + ca_{s2} & \cdots & a_{rn} + ca_{sn} \\ \cdots & \cdots & \cdots & \\ a_{s1} & a_{s2} & \cdots & a_{sn} \\ \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \cdots & \cdots & & \\ a_{s1} & a_{s2} & \cdots & a_{sn} \\ \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + c \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \\ a_{s1} & a_{s2} & \cdots & a_{sn} \\ \cdots & \cdots & & \\ a_{s1} & a_{s2} & \cdots & a_{sn} \\ \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$= \det A + 0$, since the determinant in the right hand side has two identical rows

$= \det A$.

Theorem 2.1.7. In an $n \times n$ matrix A, if one row (column) be expressed as a linear combination of the remaining rows (columns) then $\det A = 0$.

Proof. Without loss of generality, we may assume that the last row is a linear combination of the first $n-1$ rows. Because if any other row, say the rth, be such then it can be brought to the last by interchange of rows with a possible change of sign of $\det A$ but this will not affect the theorem.

Let $A = (a_{ij})$ and $a_{nj} = c_1a_{1j} + c_2a_{2j} + \cdots + c_{n-1}a_{n-1j}$, where $c_1, c_2, \ldots, c_{n-1}$ are scalars, $j = 1, 2, \ldots, n$.

$$\text{Then } \det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ c_1a_{11} & c_1a_{12} & \cdots & c_1a_{1n} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & & \cdots \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ c_2a_{21} & c_2a_{22} & \cdots & c_2a_{2n} \end{vmatrix}$$

$$+ \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ c_{n-1}a_{n-11} & c_{n-1}a_{n-12} & \cdots & c_{n-1}a_{n-1n} \end{vmatrix}$$

by the corollary of the Theorem 2.1.5

$$= c_1 \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ a_{11} & a_{12} & \cdots & a_{1n} \end{vmatrix} + c_2 \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \end{vmatrix}$$

$$+ \dots + c_{n-1} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n} \end{vmatrix}$$

$= c_1.0 + c_2.0 + \cdots + c_{n-1}.0$, by Theorem 2.1.3

$= 0.$

Worked Examples.

1. Without expanding the determinant, prove that

$$\begin{vmatrix} 0 & -a & -b \\ a & 0 & c \\ b & -c & 0 \end{vmatrix} = 0.$$

Let $\Delta = \begin{vmatrix} 0 & -a & -b \\ a & 0 & c \\ b & -c & 0 \end{vmatrix}$.

Then $\Delta = -\begin{vmatrix} 0 & a & b \\ -a & 0 & -c \\ -b & c & 0 \end{vmatrix}$ [multiplying each row by -1]

$= -\begin{vmatrix} 0 & -a & -b \\ a & 0 & c \\ b & -c & 0 \end{vmatrix}$ [interchanging rows and columns]

$= -\Delta.$

This gives $\Delta = 0$.

2. Without expanding the determinant, prove that

$$\begin{vmatrix} 1 & a & a^2 & a^3 + bcd \\ 1 & b & b^2 & b^3 + cda \\ 1 & c & c^2 & c^3 + dab \\ 1 & d & d^2 & d^3 + abc \end{vmatrix} = 0.$$

Let $\Delta = \begin{vmatrix} 1 & a & a^2 & a^3 + bcd \\ 1 & b & b^2 & b^3 + cda \\ 1 & c & c^2 & c^3 + dab \\ 1 & d & d^2 & d^3 + abc \end{vmatrix}$.

Then $\Delta = \begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix} + \begin{vmatrix} 1 & a & a^2 & bcd \\ 1 & b & b^2 & cda \\ 1 & c & c^2 & dab \\ 1 & d & d^2 & abc \end{vmatrix} = \Delta_1 + \Delta_2$, say.

Now $\Delta_2 = \frac{1}{abcd} \begin{vmatrix} a & a^2 & a^3 & abcd \\ b & b^2 & b^3 & abcd \\ c & c^2 & c^3 & abcd \\ d & d^2 & d^3 & abcd \end{vmatrix}$ [multiplying R_1 by a, R_2 by b, R_3 by c, R_4 by d]

$= \begin{vmatrix} a & a^2 & a^3 & 1 \\ b & b^2 & b^3 & 1 \\ c & c^2 & c^3 & 1 \\ d & d^2 & d^3 & 1 \end{vmatrix}$ [multiplyng c_4 by $\frac{1}{abcd}$]

$= -\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix}$ [applying three successive interchanges to bring C_4 to C_1]

$= -\Delta_1$. Therefore $\Delta = 0$.

We have defined determinant of a square matrix whose elements belong to a field F. If, however, the elements of the matrix belong to some commutative ring with unity then also the determinant of such a matrix can be defined in like manner and all the theorems established so far will hold good in such case.

Now we consider square matrices whose elements belong to the commutative ring of all polynomials in x with real (or complex) coefficients.

Theorem 2.1.8. If the elements of an $n \times n$ matrix A are real (or complex) polynomials in x and if two rows (columns) of A become identical when $x = a$, then $x - a$ is a factor of det A.

Proof. Let $A = \begin{pmatrix} f_{11}(x) & f_{12}(x) & \dots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \dots & f_{2n}(x) \\ \dots & & \dots & \\ f_{n1}(x) & f_{n2}(x) & \dots & f_{nn}(x) \end{pmatrix}$,

where $f_{ij}(x)$ is a polynomial in x. Let the rth row and the sth row of A become identical when $x = a$.

Then $f_{rj}(a) = f_{sj}(a)$ for $j = 1, 2, \dots, n$. ... ... (i)

Since the elements of A are polynomials in x, det A is a polynomial in x. Let det $A = f(x)$.

Then $f(x) = \begin{vmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ \cdots & & \cdots & \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{vmatrix}$.

By (i) it follows that $f(a) = 0$.

Therefore $x - a$ is a factor of $f(x)$, by the Factor theorem.

This completes the proof.

Theorem 2.1.9. If the elements of an $n \times n$ matrix A are real (or complex) polynomials in x and r rows (columns) of A become identical when $x = a$, then $(x-a)^{r-1}$ is a factor of det A.

Proof. Let $A = \begin{pmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ \cdots & & \cdots & \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{pmatrix}$,

where $f_{ij}(x)$ is a polynomial in x.

Without loss of generality, we can assume that first r rows of A become identical when $x = a$, because the rows which become identical when $x = a$ can be made first r rows by interchange of rows and this operation may alter the sign of the determinant but will not affect the theorem.

Since the elements of A are polynomials in x, det A is a polynomial in x. Let det $A = f(x)$.

Then $f(x) = \begin{vmatrix} f_{11}(x) & f_{12}(x) & \cdots & f_{1n}(x) \\ f_{21}(x) & f_{22}(x) & \cdots & f_{2n}(x) \\ \cdots & & \cdots & \\ f_{r1}(x) & f_{r2}(x) & \cdots & f_{rn}(x) \\ \cdots & & \cdots & \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{vmatrix}$

$= \begin{vmatrix} f_{11}(x) - f_{r1}(x) & f_{12}(x) - f_{r2}(x) & \cdots & f_{1n}(x) - f_{rn}(x) \\ f_{21}(x) - f_{r1}(x) & f_{22}(x) - f_{r2}(x) & \cdots & f_{2n}(x) - f_{rn}(x) \\ \cdots & & \cdots & \\ f_{r-11}(x) - f_{r1}(x) & f_{r-12}(x) - f_{r2}(x) & \cdots & f_{r-1n}(x) - f_{rn}(x) \\ f_{r1}(x) & f_{r2}(x) & \cdots & f_{rn}(x) \\ \cdots & & \cdots & \\ f_{n1}(x) & f_{n2}(x) & \cdots & f_{nn}(x) \end{vmatrix}$

[subtracting the rth row from each of the first $r-1$ rows]

But $f_{1j}(a) - f_{rj}(a) = 0$ for $j = 1, 2, \ldots, n$.

By the Factor theorem, $x - a$ is a factor of $f_{1j}(x) - f_{rj}(x)$, $j = 1, 2, \ldots, n$. That is, $x - a$ is a factor of each element of the first row.

By similar arguments, $x - a$ is a factor of each element of the 2nd row, each element of the third row, ..., each element of the $(r-1)$th row. Hence $(x-a)^{r-1}$ is a factor of $f(x)$.

This completes the proof.

Worked Examples (continued).

3. Prove that $\begin{vmatrix} a^2 & a & 1 \\ b^2 & b & 1 \\ c^2 & c & 1 \end{vmatrix} = -(a-b)(b-c)(c-a)$.

Let $A = \begin{pmatrix} a^2 & a & 1 \\ b^2 & b & 1 \\ c^2 & c & 1 \end{pmatrix}$.

Let us consider the elements of A as polynomials in a. Two rows of A become identical when $a = b$. Therefore $a - b$ is a factor of det A.

Let us consider the elements of A as polynomials in b. Two rows of A become identical when $b = c$. Therefore $b - c$ is a factor of det A.

Similarly $c - a$ is a factor of det A.

The expansion of det A is a polynomial in a, b, c of degree 3.

Therefore det $A = k(a-b)(b-c)(c-a)$, ... (i)

where k is independent of a, b, c.

The leading term in the expansion of det A is a^2b. No other term in the expansion of det A is a^2b.

Equating coefficients of a^2b from both sides of the equality (i), we have $1 = k.1.1 - 1$ and this gives $k = -1$.

Therefore det $A = -(b-c)(c-a)(a-b)$.

4. Prove that $\begin{vmatrix} 1+x & 2 & 2 \\ 2 & 1+x & 2 \\ 2 & 2 & 1+x \end{vmatrix} = (x-1)^2(x+5)$.

Let $A = \begin{pmatrix} 1+x & 2 & 2 \\ 2 & 1+x & 2 \\ 2 & 2 & 1+x \end{pmatrix}$.

If $x = 1$, three rows of A become identical. Therefore $(x-1)^2$ is a factor of det A.

Again $\det A = \begin{vmatrix} 1+x & 2 & 2 \\ 2 & 1+x & 2 \\ 2 & 2 & 1+x \end{vmatrix} = \begin{vmatrix} 5+x & 2 & 2 \\ 5+x & 1+x & 2 \\ 5+x & 2 & 1+x \end{vmatrix}$ $[C_1' = C_1 + C_2 + C_3]$

$= (5+x) \begin{vmatrix} 1 & 2 & 2 \\ 1 & 1+x & 2 \\ 1 & 2 & 1+x \end{vmatrix}$.

This shows that $x+5$ is a factor of det A.

The expansion of det A is a polynomial in x of degree 3.

Therefore $\det A = k(x-1)^2(x+5)$, ...(i)

where k is independent of x.

Equating the coefficient of x^3 from both sides of (i), we have $k = 1$.

Therefore $\det A = (x-1)^2(x+5)$.

5. Vandermonde Determinant.

$$\begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \dots & & \dots & & \\ x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{vmatrix} = \prod_{1 \le i < j \le n} (x_i - x_j).$$

Let $A = \begin{pmatrix} x_1^{n-1} & x_1^{n-2} & \dots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \dots & x_2 & 1 \\ \dots & & \dots & & \\ x_n^{n-1} & x_n^{n-2} & \dots & x_n & 1 \end{pmatrix}$.

If $x_i = x_j$ where $1 \le i < j \le n$, two rows of A become identical. Therefore $x_i - x_j$ is a factor of det A.

Thus $(x_1 - x_2)(x_1 - x_3)\dots(x_1 - x_n)$, $(x_2 - x_3)(x_2 - x_4)\dots(x_2 - x_n)$, $\dots$, $(x_{n-1} - x_n)$ are factors of det A. These are in all $(n-1)+(n-2)+\dots+1$ linear factors of det A.

But the expansion of det A is a polynomial in $x_1, x_2, \dots, x_n$ of degree $1+2+\dots+(n-1)$.

Therefore $\det A = k[(x_1 - x_2)(x_1 - x_3)\dots(x_1 - x_n)]$ $[(x_2 - x_3)(x_2 - x_4)\dots(x_2 - x_n)]\dots[(x_{n-1} - x_n)]$, where k is independent of $x_1, x_2, \dots, x_n$.

The coefficient of $x_1^{n-1}x_2^{n-2}\dots x_{n-1}$ (the leading term) in the expansion of det A is 1 and the coefficient of $x_1^{n-1}x_2^{n-2}\dots x_{n-1}$ in the R.H.S expression is k. Therefore $k = 1$ and

$$\det A = \prod_{1 \le i < j \le n} (x_i - x_j).$$

2.2. Cofactors and Minors.

Let $A = (a_{ij})$ be an $n \times n$ matrix. Then

$$\det A = \sum_{\phi} sgn\ \phi\ a_{1\phi(1)} a_{2\phi(2)} \dots a_{n\phi(n)},$$

where ϕ is a permutation on $\{1, 2, \dots, n\}$ and $sgn\ \phi = 1$ or -1 according as the permutation $\phi = \begin{pmatrix} 1 & 2 & \dots & n \\ \phi(1) & \phi(2) & \dots & \phi(n) \end{pmatrix}$ is even or odd.

There are $n!$ terms in the expansion. Each term contains one and only one element from each row and also one and only one element from each column.

Let us consider, in particular, the ith row. Each term of the expansion of det A contains one and only one of $a_{i1}, a_{i2}, \dots, a_{in}$. Therefore the expansion of det A can be exhibited as

$$\det A = a_{i1}(* * *) + a_{i2}(* * *) + \dots + a_{in}(* * *).$$

The companion factor $(* * *)$ of a_{ij} is called the *cofactor* of a_{ij} in det A and is denoted by A_{ij}.

Thus $\det A = a_{i1}A_{i1} + a_{i2}A_{i2} + \dots + a_{in}A_{in}$.

This is called the expansion of det A in terms of the elements of the ith row. Taking $i = 1, 2, \dots, n$, there are n different expansions for det A.

$$\begin{aligned} \det A &= a_{11}A_{11} + a_{12}A_{12} + \dots + a_{1n}A_{1n} \\ &= a_{21}A_{21} + a_{22}A_{22} + \dots + a_{2n}A_{2n} \\ &\dots \quad \dots \quad \dots \\ &= a_{n1}A_{n1} + a_{n2}A_{n2} + \dots + a_{nn}A_{nn}. \end{aligned}$$

Proceeding with similar arguments in respect of columns of A and considering, in particular, the elements of the ith column, the expansion of det A can be exhibited as

$$\det A = a_{1i}A_{1i} + a_{2i}A_{2i} + \dots + a_{ni}A_{ni},$$

where A_{ki} is the cofactor of a_{ki} in det A.

This is called the expansion of det A in terms of the elements of the ith column.

Taking $i = 1, 2, \dots, n$, there are n different expansions for det A.

$$\begin{aligned} \det A &= a_{11}A_{11} + a_{21}A_{21} + \dots + a_{n1}A_{n1} \\ &= a_{12}A_{12} + a_{22}A_{22} + \dots + a_{n2}A_{n2} \\ &\dots \quad \dots \quad \dots \\ &= a_{1n}A_{1n} + a_{2n}A_{2n} + \dots + a_{nn}A_{nn}. \end{aligned}$$

Note. A_{ij} contains $(n-1)!$ terms each of which contains one and only one element from each row of A excepting the ith and one and only one element from each column of A excepting the jth.

Example.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

The expansion can be expressed as $a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$.

Therefore $A_{11} = a_{22}a_{33} - a_{23}a_{32}$, $A_{12} = a_{23}a_{31} - a_{21}a_{33}$, $A_{13} = a_{21}a_{32} - a_{22}a_{31}$.

The expansion can also be expressed as $a_{21}(a_{13}a_{32} - a_{12}a_{33}) + a_{22}(a_{11}a_{33} - a_{13}a_{31}) + a_{23}(a_{12}a_{31} - a_{11}a_{32})$.

Therefore $A_{21} = a_{13}a_{32} - a_{12}a_{33}$, $A_{22} = a_{11}a_{33} - a_{13}a_{31}$, $A_{23} = a_{12}a_{31} - a_{11}a_{32}$.

Similarly, $A_{31} = a_{12}a_{23} - a_{13}a_{22}$, $A_{32} = a_{13}a_{21} - a_{11}a_{23}$, $A_{33} = a_{11}a_{22} - a_{12}a_{21}$.

If one row and one column be deleted from an $n \times n$ matrix (a_{ij}) the determinant of the remaining $(n-1) \times (n-1)$ matrix is said to be a *minor* of order $n-1$ of A. The minor of order $n-1$ obtained by deleting the ith row and the jth column is denoted by M_{ij} and it is said to be the *minor* of the element a_{ij} in det A.

In the determinant $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$, $M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$,

$M_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$, $M_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$ etc.

Theorem 2.2.1. The cofactor of a_{ij} in det A is equal to $(-)^{i+j}$.(the minor of a_{ij}),

i.e., $A_{ij} = (-1)^{i+j} M_{ij}$.

Proof. We have det $A = \sum_{\phi} sgn\ \phi\ a_{1\phi(1)}a_{2\phi(2)} \cdots a_{n\phi(n)}$, where ϕ is a permutation on $\{1, 2, \ldots, n\}$.

In the summation, there are some terms containing a_{11}, some containing $a_{12}, \ldots$, and some containing a_{1n}. The terms containing a_{11} are those corresponding to the permutations ϕ_1 for which $\phi_1(1) = 1$. The terms containing a_{12} are those corresponding to the permutations ϕ_2 for which $\phi_2(1) = 2, \ldots$, the terms containing a_{1n} are those corresponding to the permutations ϕ_n for which $\phi_n(1) = n$.

Hence det $A = a_{11} \sum_{\phi_1} sgn\ \phi_1\ a_{2\phi_1(2)}a_{3\phi_1(3)} \cdots a_{n\phi_1(n)}$
$+ a_{12} \sum_{\phi_2} sgn\ \phi_2\ a_{2\phi_2(2)}a_{3\phi_2(3)} \cdots a_{n\phi_2(n)}$
$+ \cdots + a_{1n} \sum_{\phi_n} sgn\ \phi_n\ a_{2\phi_n(2)}a_{3\phi_n(3)} \cdots a_{n\phi_n(n)}$

Therefore $A_{11} = \sum_{\phi_1} sgn\ \phi_1\ a_{2\phi_1(2)}a_{3\phi_1(3)} \cdots a_{n\phi_1(n)}$.

There are $(n-1)!$ permutations on the set $\{1, 2, \ldots, n\}$ for which $1 \to 1$. Therefore ϕ_1 can be considered as a permutation of $\{2, 3, \ldots, n\}$ and the summation $\sum_{\phi_1}$ contains $(n-1)!$ terms.

$$\text{So } A_{11} = \begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ a_{32} & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & & \\ a_{n2} & a_{n3} & \cdots & a_{nn} \end{vmatrix} = M_{11} = (-1)^{1+1} M_{11}.$$

Therefore the theorem is true for A_{11}.

$$\text{Now det } A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & & & \cdots & \cdots & \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \cdots & & & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

$$= (-1)^{i-1} \begin{vmatrix} a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \cdots & & & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

[applying $i-1$ successive interchanges of rows to bring the ith row to the first row]

$$= (-1)^{i-1+j-1} \begin{vmatrix} a_{ij} & a_{i1} & \cdots & a_{in} \\ a_{1j} & a_{11} & \cdots & a_{1n} \\ a_{2j} & a_{21} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{nj} & a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

[applying $j-1$ successive interchanges of columns to bring the jth column to the first column]

$= (-1)^{i+j-2}$ det B, say, where

$$B = \begin{bmatrix} a_{ij} & a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{1j} & a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & & \cdots & & \\ a_{nj} & a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

But in det B, the cofactor of a_{ij}

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & a_{2j+1} & \cdots & a_{2n} \\ \cdots & & \cdots & & \cdots & & \cdots \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \cdots & & \cdots & & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{vmatrix} = M_{ij}.$$

Therefore the cofactor of a_{ij} in det A
$=$ the cofactor of a_{ij} in $(-1)^{i+j-2}$ det $B = (-1)^{i+j-2}M_{ij} = (-1)^{i+j}M_{ij}$.

That is, $A_{ij} = (-1)^{i+j}M_{ij}$. This completes the proof.

The theorem gives a working rule for the computation of det A.

We have det $A = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}$.

In particular, det A
$$= a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}$$
$$= a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n}.$$

This is the expansion of det A in terms of the elements of the first row and their corresponding minors.

Also det A
$$= a_{11}A_{11} + a_{21}A_{21} + \cdots + a_{n1}A_{n1}$$
$$= a_{11}M_{11} - a_{21}M_{21} + \cdots + (-1)^{n+1}a_{n1}M_{n1}.$$

This is the expansion of det A in terms of the elements of the first column and their corresponding minors.

In a similar manner the computation of det A can be effected in terms of the elements of any chosen row (column) of A.

Example. Compute $\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$.

Expanding in terms of the elements of the first row, we have

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = a\begin{vmatrix} b & f \\ f & c \end{vmatrix} - h\begin{vmatrix} h & f \\ g & c \end{vmatrix} + g\begin{vmatrix} h & b \\ g & f \end{vmatrix}$$
$$= a(bc - f^2) - h(hc - fg) + g(hf - bg)$$
$$= abc + 2fgh - af^2 - bg^2 - ch^2.$$

Expanding in terms of the elements of the second row, we have

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = -h\begin{vmatrix} h & g \\ f & c \end{vmatrix} + b\begin{vmatrix} a & g \\ g & c \end{vmatrix} - f\begin{vmatrix} a & h \\ g & f \end{vmatrix}$$
$$= -h(ch - fg) + b(ac - g^2) - f(af - gh)$$
$$= abc + 2fgh - af^2 - bg^2 - ch^2.$$

Expanding in terms of the elements of the first column, we have

$$\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix} = a\begin{vmatrix} b & f \\ f & c \end{vmatrix} - h\begin{vmatrix} h & g \\ f & c \end{vmatrix} + g\begin{vmatrix} h & g \\ b & f \end{vmatrix}$$
$$= a(bc - f^2) - h(ch - fg) + g(hf - bg)$$
$$= abc + 2fgh - af^2 - bg^2 - ch^2.$$

Theorem 2.2.2. Let (a_{ij}) be an $n \times n$ matrix and A_{ij} be the cofactor of a_{ij} in det (a_{ij}). Then

(i) $a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = 0,\ i \neq k;$

(ii) $a_{1i}A_{1k} + a_{2i}A_{2k} + \cdots + a_{ni}A_{nk} = 0,\ i \neq k.$

Lemma. $c_1A_{k1} + c_2A_{k2} + \cdots + c_nA_{kn} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \\ a_{k-11} & a_{k-12} & \cdots & a_{k-1n} \\ c_1 & c_2 & \cdots & c_n \\ a_{k+11} & a_{k+12} & \cdots & a_{k+1n} \\ \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$

Let $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \\ a_{k-11} & a_{k-12} & \cdots & a_{k-1n} \\ c_1 & c_2 & \cdots & c_n \\ a_{k+11} & a_{k+12} & \cdots & a_{k+1n} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = B = (b_{ij}).$

Then $b_{ij} = a_{ij}$ $i \neq k$; $b_{kj} = c_j$. Let M_{rs} be the minors of A, M'_{rs} be those of B. Then $M_{kj} = M'_{kj}, j = 1, 2, \ldots, n$.

det $B = b_{k1}B_{k1} + b_{k2}B_{k2} + \cdots + b_{kn}B_{kn}$
$= c_1(-1)^{k+1}M'_{k1} + c_2(-1)^{k+2}M'_{k2} + \cdots + c_n(-1)^{k+n}M'_{kn}$
$= c_1(-1)^{k+1}M_{k1} + c_2(-1)^{k+2}M_{k2} + \cdots + c_n(-1)^{k+n}M_{kn}$
$= c_1A_{k1} + c_2A_{k2} + \cdots + c_nA_{kn}.$

This proves the Lemma.

Let the kth row of B be replaced by ith row of A, where $i \neq k$ and let the new matrix be B'.

Then $\det B' = a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn}$, by the Lemma.

But $\det B' = 0$, since two rows of B' are identical.

Therefore $a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = 0$, if $i \neq k$.

The second part of the theorem follows from the consideration of columns.

Worked Examples.

1. Prove that

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \binom{m}{1} & \binom{m+1}{1} & \binom{m+2}{1} & \cdots & \binom{m+n-1}{1} \\ \binom{m}{2} & \binom{m+1}{2} & \binom{m+2}{2} & \cdots & \binom{m+n-1}{2} \\ \cdots & & \cdots & & \\ \binom{m}{n-1} & \binom{m+1}{n-1} & \binom{m+2}{n-1} & \cdots & \binom{m+n-1}{n-1} \end{vmatrix} = 1,$$

m, n being positive integers and $m \geq n-1 \geq 1$ and $\binom{m}{r} = m_{C_r}$.

$$\text{Let } \Delta_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \binom{m}{1} & \binom{m+1}{1} & \binom{m+2}{1} & \cdots & \binom{m+n-1}{1} \\ \binom{m}{2} & \binom{m+1}{2} & \binom{m+2}{2} & \cdots & \binom{m+n-1}{2} \\ \cdots & & \cdots & & \\ \binom{m}{n-1} & \binom{m+1}{n-1} & \binom{m+2}{n-1} & \cdots & \binom{m+n-1}{n-1} \end{vmatrix}.$$

Subtracting the preceding column from each column beginning with the second, we have

$$\Delta_n = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ \binom{m}{1} & 1 & 1 & \cdots & 1 \\ \binom{m}{2} & \binom{m}{1} & \binom{m+1}{1} & \cdots & \binom{m+n-2}{1} \\ \cdots & & \cdots & & \\ \binom{m}{n-1} & \binom{m}{n-1} & \binom{m+1}{n-1} & \cdots & \binom{m+n-2}{n-1} \end{vmatrix}.$$

Expanding in terms of the elements of the first row, we have

$$\Delta_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \binom{m}{1} & \binom{m+1}{1} & \cdots & \binom{m+n-2}{1} \\ \cdots & \cdots & \cdots & \cdots \\ \binom{m}{n-2} & \binom{m+1}{n-2} & \cdots & \binom{m+n-2}{n-2} \end{vmatrix}$$

$= \Delta_{n-1}$.

Therefore $\Delta_n = \Delta_{n-1} = \Delta_{n-2} = \cdots = \Delta_2$.

$$\text{But } \Delta_2 = \begin{vmatrix} 1 & 1 \\ \binom{m}{1} & \binom{m+1}{1} \end{vmatrix} = 1.$$

Consequently, $\Delta_n = 1$.

2. Prove that
$$\begin{vmatrix} x^3 & x^2 & x & 1 \\ \alpha^3 & \alpha^2 & \alpha & 1 \\ \beta^3 & \beta^2 & \beta & 1 \\ \gamma^3 & \gamma^2 & \gamma & 1 \end{vmatrix} = -(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)$$
$$(x-\alpha)(x-\beta)(x-\gamma).$$

Deduce that
$$\begin{vmatrix} \alpha^3 & \alpha^2 & 1 \\ \beta^3 & \beta^2 & 1 \\ \gamma^3 & \gamma^2 & 1 \end{vmatrix} = -(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)(\alpha\beta+\beta\gamma+\gamma\alpha).$$

$$\text{Let } A = \begin{pmatrix} x^3 & x^2 & x & 1 \\ \alpha^3 & \alpha^2 & \alpha & 1 \\ \beta^3 & \beta^2 & \beta & 1 \\ \gamma^3 & \gamma^2 & \gamma & 1 \end{pmatrix}.$$

$x-\alpha$, $x-\beta$, $x-\gamma$, $\alpha-\beta$, $\beta-\gamma$, $\gamma-\alpha$ are factors of $\det A$ and $\det A$ is a polynomial in x, α, β, γ of degree 6.

Therefore $\det A = k(x-\alpha)(x-\beta)(x-\gamma)(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)$, where k is independent of x, α, β, γ. The leading term in the expansion of $\det A$ is $\alpha^2\beta x^3$.

Equating coefficient of $\alpha^2\beta x^3$, we have $1 = k.1.1.1.1.1.-1$. Therefore $k = -1$.

$$\begin{aligned} \det A &= -(x-\alpha)(x-\beta)(x-\gamma)(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha) \\ &= -(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha) \\ &\quad [x^3 - (\alpha+\beta+\gamma)x^2 + (\alpha\beta+\beta\gamma+\gamma\alpha)x - \alpha\beta\gamma]. \end{aligned}$$

Again expanding $\det A$ in terms of the first row, $\det A =$

$$x^3\begin{vmatrix}\alpha^2 & \alpha & 1\\ \beta^2 & \beta & 1\\ \gamma^2 & \gamma & 1\end{vmatrix} - x^2\begin{vmatrix}\alpha^3 & \alpha & 1\\ \beta^3 & \beta & 1\\ \gamma^3 & \gamma & 1\end{vmatrix} + x\begin{vmatrix}\alpha^3 & \alpha^2 & 1\\ \beta^3 & \beta^2 & 1\\ \gamma^3 & \gamma^2 & 1\end{vmatrix} - 1\begin{vmatrix}\alpha^3 & \alpha^2 & 1\\ \beta^3 & \beta^2 & 1\\ \gamma^3 & \gamma^2 & 1\end{vmatrix}.$$

Equating coefficient of x, we have

$$\begin{vmatrix}\alpha^3 & \alpha^2 & 1\\ \beta^3 & \beta^2 & 1\\ \gamma^3 & \gamma^2 & 1\end{vmatrix} = -(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)(\alpha\beta+\beta\gamma+\gamma\alpha).$$

3. Let $D_n = \begin{vmatrix} 1 & x & 0 & 0 & \cdots & 0 & 0\\ x & 1 & x & 0 & \cdots & 0 & 0\\ 0 & x & 1 & x & \cdots & 0 & 0\\ & \cdots & \cdots & \cdots & \cdots & & \\ 0 & 0 & 0 & 0 & \cdots & x & 1\end{vmatrix}$

be a determinant of order n in which the diagonal elements are 1 and the elements just above and just below the diagonal elements are x and all other elements zero.

Prove that $D_n - D_{n-1} + x^2 D_{n-2} = 0$.

Hence prove that $\begin{vmatrix} 1 & \frac{1}{2} & 0 & 0 & \cdots & 0\\ \frac{1}{2} & 1 & \frac{1}{2} & 0 & \cdots & 0\\ 0 & \frac{1}{2} & 1 & \frac{1}{2} & \cdots & 0\\ & \cdots & \cdots & \cdots & & \\ 0 & 0 & 0 & 0 & \cdots & 1\end{vmatrix} = \frac{n+1}{2^n}$.

Expanding D_n in terms of the elements of first row,

$$D_n = D_{n-1} - x\begin{vmatrix} x & x & 0 & \cdots & 0\\ 0 & 1 & x & \cdots & 0\\ 0 & 0 & 0 & \cdots & 1\end{vmatrix} = D_{n-1} - x^2 D_{n-2}$$

[Expanding in terms of the elements of the first column]

or, $D_n - D_{n-1} + x^2 D_{n-2} = 0$.

When $x = \frac{1}{2}$, $D_n - D_{n-1} + \frac{1}{4}D_{n-2} = 0$.

This is a second order linear difference equation. The roots of the auxiliary equation are $\frac{1}{2}, \frac{1}{2}$.

So $D_n = (A + nB)(\frac{1}{2})^n$, where A and B are constants.

But $D_1 = 1$, $D_2 = \begin{vmatrix} 1 & \frac{1}{2}\\ \frac{1}{2} & 1\end{vmatrix} = \frac{3}{4}$.

Therefore $A + B = 2$ and $A + 2B = 3$.

This gives $A = B = 1$ and hence $D_n = \frac{n+1}{2^n}$.

2.3. Complementary minor and Algebraic complement.

Let $A = (a_{ij})$ be an $n \times n$ matrix.

If r rows and r columns be deleted from A, the determinant of the remaining $(n-r) \times (n-r)$ matrix is said to be a *minor of order* $n - r$ of the matrix A.

If $i_1, i_2, \ldots, i_r$ be r integers chosen from $1, 2, \ldots, n$ and arranged in the natural order (i.e., $i_1 < i_2 < \cdots < i_r$) and $j_1, j_2, \ldots, j_r$ be r integers chosen from $1, 2, \ldots, n$ and arranged in the natural order, then the minor of order $n - r$ obtained by deleting r rows i_1th, i_2th, $\ldots$, i_rth; and r columns j_1th, j_2th, $\ldots$, j_rth is denoted by

$$M_{i_1 i_2 \ldots i_r, j_1 j_2 \ldots j_r}.$$

For example, in the determinant $|a_{ij}|_4$,

$$M_{12,13} = \begin{vmatrix} a_{32} & a_{34}\\ a_{42} & a_{44}\end{vmatrix}, \quad M_{13,23} = \begin{vmatrix} a_{21} & a_{24}\\ a_{41} & a_{44}\end{vmatrix}, \text{ etc.}$$

If $i_1, i_2, \ldots, i_r$ be integers chosen from $1, 2, \ldots, n$ and arranged in the natural order and $p_1, p_2, \ldots, p_{n-r}$ be the remaining $n - r$ integers and also arranged in the natural order, then $i_1, i_2, \ldots, i_r$ and $p_1, p_2, \ldots, p_{n-r}$ are said to be *complementary* sets.

If $i_1, i_2, \ldots, i_r$ and $p_1, p_2, \ldots, p_{n-r}$ be complementary sets of row indices and $j_1, j_2, \ldots, j_r$ and $q_1, q_2, \ldots, q_{n-r}$ be complementary sets of column indices then the minors

$M_{i_1 i_2 \ldots i_r, j_1 j_2 \ldots j_r}$ and $M_{p_1 p_2 \ldots p_{n-r}, q_1 q_2 \ldots q_{n-r}}$ are said to be complementary minors.

For example, in the determinant $|a_{ij}|_4$,

$\begin{vmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}$ and $\begin{vmatrix} a_{33} & a_{34}\\ a_{43} & a_{44}\end{vmatrix}$ are complementary minors;

$\begin{vmatrix} a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}$ and $\begin{vmatrix} a_{13} & a_{14}\\ a_{43} & a_{44}\end{vmatrix}$ are complementary minors;

a_{11} and M_{11} are complementary minors.

Let M be a minor of order r obtained from r rows i_1th, i_2th, $\ldots$, i_rth $(i_1 < i_2 < \cdots < i_r)$ and r columns j_1th, j_2th, $\ldots$, j_rth $(j_1 < j_1 < \cdots < j_r)$ and M' be the complementary minor of M. Then the *algebraic complement* of M is defined as $(-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r} M'$.

In particular, if $M = a_{ij}$, then $M' = M_{ij}$ and the algebraic complement of a_{ij} is $(-1)^{i+j} M_{ij}$. In other words, the algebraic complement of an element a_{ij} is the cofactor of a_{ij} in det (a_{ij}).

For example, in $|a_{ij}|_4$, the algebraic complement of

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \text{ is } (-1)^{1+2+1+2}\begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix}$$

and the algebraic complement of

$$\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \text{ is } (-1)^{2+3+1+2}\begin{vmatrix} a_{13} & a_{14} \\ a_{43} & a_{44} \end{vmatrix}.$$

We have seen that det A can be expanded in terms of the elements of a particular row (or column) and their corresponding minors. Treating the elements of a row as minors of order 1, the expansion of det A can be considered as the sum of products of all minors of order 1 appearing in a chosen row and their respective algebraic complements. Laplace's method of expansion of det A gives a generalisation of this concept.

2.3.1. Laplace's method of expansion.

In an $n \times n$ matrix A if any r rows be selected, det A can be expressed as the sum of the products of all minors of order r formed from those r selected rows and their respective algebraic complements.

The method can be applied to columns of A in an analogous manner.

As an illustration of the method, the determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}$$

can be expanded in terms of minors of order 2 formed from the first two rows as

$$(-1)^{1+2+1+2}\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} + (-1)^{1+2+1+3}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}\begin{vmatrix} a_{32} & a_{34} \\ a_{42} & a_{44} \end{vmatrix}$$

$$+(-1)^{1+2+1+4}\begin{vmatrix} a_{11} & a_{14} \\ a_{21} & a_{24} \end{vmatrix}\begin{vmatrix} a_{32} & a_{33} \\ a_{42} & a_{43} \end{vmatrix} + (-1)^{1+2+2+3}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\begin{vmatrix} a_{31} & a_{34} \\ a_{41} & a_{44} \end{vmatrix}$$

$$+(-1)^{1+2+2+4}\begin{vmatrix} a_{12} & a_{14} \\ a_{22} & a_{24} \end{vmatrix}\begin{vmatrix} a_{31} & a_{33} \\ a_{41} & a_{43} \end{vmatrix} + (-1)^{1+2+3+4}\begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix}\begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} - \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}\begin{vmatrix} a_{32} & a_{34} \\ a_{42} & a_{44} \end{vmatrix}$$

$$+ \begin{vmatrix} a_{11} & a_{14} \\ a_{21} & a_{24} \end{vmatrix}\begin{vmatrix} a_{32} & a_{33} \\ a_{42} & a_{43} \end{vmatrix} + \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\begin{vmatrix} a_{31} & a_{34} \\ a_{41} & a_{44} \end{vmatrix}$$

$$- \begin{vmatrix} a_{12} & a_{14} \\ a_{22} & a_{24} \end{vmatrix}\begin{vmatrix} a_{31} & a_{33} \\ a_{41} & a_{43} \end{vmatrix} + \begin{vmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{vmatrix}\begin{vmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{vmatrix}.$$

Worked Examples.

1. Prove that $\begin{vmatrix} a & -b & -a & b \\ b & a & -b & -a \\ c & -d & c & -d \\ d & c & d & c \end{vmatrix} = 4(a^2+b^2)(c^2+d^2).$

Expanding the determinant by Laplace's method in terms of minors of order 2 formed from the first two rows, we have

$$\Delta = (-1)^{1+2+1+2}\begin{vmatrix} a & -b \\ b & a \end{vmatrix}\begin{vmatrix} c & -d \\ d & c \end{vmatrix}$$

$$+(-1)^{1+2+1+4}\begin{vmatrix} a & b \\ b & -a \end{vmatrix}\begin{vmatrix} -d & c \\ c & d \end{vmatrix}$$

$$+(-1)^{1+2+2+3}\begin{vmatrix} -b & -a \\ a & -b \end{vmatrix}\begin{vmatrix} c & -d \\ d & c \end{vmatrix}$$

$$+(-1)^{1+2+3+4}\begin{vmatrix} -a & b \\ -b & -a \end{vmatrix}\begin{vmatrix} c & -d \\ d & c \end{vmatrix}, \text{ other minors being zero.}$$

$$= (a^2+b^2)(c^2+d^2) + (-a^2-b^2)(-c^2-d^2) + (a^2+b^2)(c^2+d^2)$$
$$+(a^2+b^2)(c^2+d^2)$$
$$= 4(a^2+b^2)(c^2+d^2).$$

2. Prove that $\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = (af-be+cd)^2.$

Expanding the determinant by Laplace's method in terms of minors of order 2 formed from the first two rows, we have

$$\Delta = (-1)^{1+2+1+2}\begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix}\begin{vmatrix} 0 & f \\ -f & 0 \end{vmatrix}$$

$$+(-1)^{1+2+1+3}\begin{vmatrix} 0 & b \\ -a & d \end{vmatrix}\begin{vmatrix} -d & f \\ -e & 0 \end{vmatrix}$$

$$+(-1)^{1+2+1+4}\begin{vmatrix} 0 & c \\ -a & e \end{vmatrix}\begin{vmatrix} -d & 0 \\ -e & -f \end{vmatrix}$$

$$+(-1)^{1+2+2+3}\begin{vmatrix} a & b \\ 0 & d \end{vmatrix}\begin{vmatrix} -b & f \\ -c & 0 \end{vmatrix}$$

$$+(-1)^{1+2+2+4}\begin{vmatrix} a & c \\ 0 & e \end{vmatrix}\begin{vmatrix} -b & 0 \\ -c & -f \end{vmatrix}$$

$$+(-1)^{1+2+3+4}\begin{vmatrix} b & c \\ d & e \end{vmatrix}\begin{vmatrix} -b & -d \\ -c & -e \end{vmatrix}$$

$= a^2f^2 - abcf + acdf + adcf - aebf + (be - cd)^2$
$= a^2f^2 - 2af(be - cd) + (be - cd)^2$
$= (af - be + cd)^2.$

2.4. Multiplication of determinants.

Theorem 2.4.1. If A and B be square matrices of the same order, then det $(AB) =$ det A.det B.

Proof. Let $A = (a_{ij})_{n,n}$, $B = (b_{ij})_{n,n}$.

Then $AB = (c_{ij})_{n,n}$ where $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.

Let us consider the matrix $P = \begin{pmatrix} A & O \\ Q & B \end{pmatrix}$, where $Q = -I_n$ and O is the null matrix O_{nn}.

$$\det P = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & 0 & \cdots & 0 \\ & & \cdots & & & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ & & \cdots & & & & \cdots & \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}.$$

Expanding det P by Laplace's method in terms of minors of order n formed from the first n rows, det $P =$ det A.det B.

$$\text{Also } \det P = \begin{vmatrix} 0 & 0 & \cdots & 0 & c_{11} & c_{12} & \cdots & c_{1n} \\ 0 & 0 & \cdots & 0 & c_{21} & c_{22} & \cdots & c_{2n} \\ & & \cdots & & & \cdots & \cdots & \\ 0 & 0 & \cdots & 0 & c_{n1} & c_{n2} & \cdots & c_{2n} \\ -1 & 0 & \cdots & 0 & b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & -1 & \cdots & 0 & b_{21} & b_{22} & \cdots & b_{2n} \\ & & \cdots & & & \cdots & \cdots & \\ 0 & 0 & \cdots & -1 & b_{n1} & b_{n2} & \cdots & b_{nn} \end{vmatrix}$$

[$R_i' = R_i + a_{i1}R_{n+1} + a_{i2}R_{n+2} + \cdots + a_{in}R_{2n}$, $(i = 1, 2, \ldots, n)$; where R_i' is the ith row of the transformed matrix and R_i is that of the original matrix P]

$$= \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} \times (-1)^{1+2+\cdots+2n} \begin{vmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 \end{vmatrix}$$

[Expanding by Laplace's method in terms of minors of order n formed from the first n rows]

$$= \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} \times (-1)^{2n(2n+1)/2}.(-1)^n$$

$$= \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix} = \det(AB).$$

Therefore det A.det $B =$ det (AB).

This proves the theorem.

Rules of multiplication of determinants.

Let $A = (a_{ij})$, $B = (b_{ij})$ be $n \times n$ matrices.

Then det A.det $B =$ det (AB)

$= \det C = \det(c_{ij})$, where $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$.

For example,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{vmatrix}.$$

The rule of formation of the elements in the product is called the '*matrix rule*' or the rule of '*multiplication of rows by columns*'.

We have det A.det $B =$ det A.det B^t, since det $B^t =$ det B

$= \det(AB^t)$,

$= \det D = \det(d_{ij})$, where $d_{ij} = \sum_{k=1}^{n} a_{ik}b_{jk}$.

For example,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11}b_{11}+a_{12}b_{12}+a_{13}b_{13} & a_{11}b_{21}+a_{12}b_{22}+a_{13}b_{23} & a_{11}b_{31}+a_{12}b_{32}+a_{13}b_{33} \\ a_{21}b_{11}+a_{22}b_{12}+a_{23}b_{13} & a_{21}b_{21}+a_{22}b_{22}+a_{23}b_{23} & a_{21}b_{31}+a_{22}b_{32}+a_{23}b_{33} \\ a_{31}b_{11}+a_{32}b_{12}+a_{33}b_{13} & a_{31}b_{21}+a_{32}b_{22}+a_{33}b_{23} & a_{31}b_{13}+a_{32}b_{32}+a_{33}b_{33} \end{vmatrix}$$

The rule of formation of the elements in the product is called the rule of '*multiplication by rows*'.

We have det A.det B = det A^t.det B, since det A^t = det A

$$= \det (A^tB)$$

$$= \det E = \det (e_{ij}), \text{ where } e_{ij} = \sum_{k=1}^{n} a_{ki}b_{kj}.$$

For example,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11}b_{11}+a_{21}b_{21}+a_{31}b_{31} & a_{11}b_{12}+a_{21}b_{22}+a_{31}b_{32} & a_{11}b_{13}+a_{21}b_{23}+a_{31}b_{33} \\ a_{12}b_{11}+a_{22}b_{21}+a_{32}b_{31} & a_{12}b_{12}+a_{22}b_{22}+a_{32}b_{32} & a_{12}b_{13}+a_{22}b_{23}+a_{32}b_{33} \\ a_{13}b_{11}+a_{23}b_{21}+a_{33}b_{31} & a_{13}b_{12}+a_{23}b_{22}+a_{33}b_{32} & a_{13}b_{13}+a_{23}b_{23}+a_{33}b_{33} \end{vmatrix}$$

The rule of formation of the elements in the product is called the rule of '*multiplication by columns*'.

We have det A.det B = det A^t.det B^t

$$= \det (A^tB^t)$$

$$= \det F = \det (f_{ij}), \text{ where } f_{ij} = \sum_{k=1}^{n} a_{ki}b_{jk}.$$

For example,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11}b_{11}+a_{21}b_{12}+a_{31}b_{13} & a_{11}b_{21}+a_{21}b_{22}+a_{31}b_{23} & a_{11}b_{31}+a_{21}b_{32}+a_{31}b_{33} \\ a_{12}b_{11}+a_{22}b_{12}+a_{32}b_{13} & a_{12}b_{21}+a_{22}b_{22}+a_{32}b_{23} & a_{12}b_{31}+a_{22}b_{32}+a_{32}b_{33} \\ a_{13}b_{11}+a_{23}b_{12}+a_{33}b_{13} & a_{13}b_{21}+a_{23}b_{22}+a_{33}b_{23} & a_{13}b_{31}+a_{23}b_{32}+a_{33}b_{33} \end{vmatrix}$$

The rule of formation of the elements in the product is called the rule of '*multiplication of columns by rows*'.

Worked Examples.

1. Prove that

$$\begin{vmatrix} 2bc-a^2 & c^2 & b^2 \\ c^2 & 2ca-b^2 & a^2 \\ b^2 & a^2 & 2ab-c^2 \end{vmatrix} = (a^3+b^3+c^3-3abc)^2.$$

We have $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = -(a^3+b^3+c^3-3abc).$

$$\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \begin{vmatrix} -a & -b & -c \\ c & a & b \\ b & c & a \end{vmatrix}$$

[interchanging R_2 and R_3 in the right hand determinant]

$$= \begin{vmatrix} 2bc-a^2 & c^2 & b^2 \\ c^2 & 2ca-b^2 & a^2 \\ b^2 & a^2 & 2ab-c^2 \end{vmatrix}.$$

[multiplication of rows by columns]

Therefore

$$\begin{vmatrix} 2bc-a^2 & c^2 & b^2 \\ c^2 & 2ca-b^2 & a^2 \\ b^2 & a^2 & 2ab-c^2 \end{vmatrix} = (a^3+b^3+c^3-3abc)^2.$$

2. If $s_r = \alpha^r + \beta^r + \gamma^r$, prove that

$$\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = (\alpha-\beta)^2(\beta-\gamma)^2(\gamma-\alpha)^2.$$

We have $\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} = (\alpha-\beta)(\beta-\gamma)(\gamma-\alpha).$

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} \begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} = \begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix}.$$

[multiplication by rows]

Therefore $\begin{vmatrix} s_0 & s_1 & s_2 \\ s_1 & s_2 & s_3 \\ s_2 & s_3 & s_4 \end{vmatrix} = (\alpha-\beta)^2(\beta-\gamma)^2(\gamma-\alpha)^2.$

Theorem 2.4.2. (Jacobi)

If $A = (a_{ij})$ be an $n \times n$ matrix and A_{rs} be the cofactor of a_{rs} in det A, then det $(A_{ij}) = [\det (a_{ij})]^{n-1}$.

Proof. det (a_{ij}). det (A_{ij})

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{2n} & \cdots & a_{nn} \end{vmatrix} \begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{2n} & \cdots & A_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} \sum_{k=1}^{n} a_{1k}A_{1k} & \sum_{k=1}^{n} a_{1k}A_{2k} & \cdots & \sum_{k=1}^{n} a_{1k}A_{nk} \\ \sum_{k=1}^{n} a_{2k}A_{1k} & \sum_{k=1}^{n} a_{2k}A_{2k} & \cdots & \sum_{k=1}^{n} a_{2k}A_{nk} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{k=1}^{n} a_{nk}A_{1k} & \sum_{k=1}^{n} a_{nk}A_{2k} & \cdots & \sum_{k=1}^{n} a_{nk}A_{nk} \end{vmatrix}$$

[multiplication by rows]

$$= \begin{vmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \det A \end{vmatrix}, \text{ since } \sum_{k=1}^{n} a_{ik}A_{jk} = \det A \text{ if } i=j$$
$$= 0 \text{ if } i \neq j$$

$= (\det A)^n$.

Two cases come up for consideration.

Case 1. det $A \neq 0$. Then det $(A_{ij}) = (\det A)^{n-1}$.

Case 2. det $A = 0$.

Subcase (i). If each $a_{ij} = 0$, then each $A_{ij} = 0$. Therefore det $(A_{ij}) = 0 = (\det A)^{n-1}$.

Subcase (ii). If there exists at least one non-zero element in (a_{ij}), let us assume, without loss of generality, that $a_{11} \neq 0$.

$$a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} = \det A = 0$$
$$a_{11}A_{21} + a_{12}A_{22} + \cdots + a_{1n}A_{2n} = 0$$
$$\cdots \qquad \cdots$$
$$a_{11}A_{n1} + a_{12}A_{n2} + \cdots + a_{1n}A_{nn} = 0.$$

$$a_{11} \det (A_{ij}) = \begin{vmatrix} a_{11}A_{11} & A_{12} & \cdots & A_{1n} \\ a_{11}A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & & \cdots \\ a_{11}A_{n1} & A_{n2} & \cdots & A_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} & A_{12} & \cdots & A_{1n} \\ a_{11}A_{21} + a_{12}A_{22} + \cdots + a_{1n}A_{2n} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{11}A_{n1} + a_{12}A_{n2} + \cdots + a_{1n}A_{nn} & A_{n2} & \cdots & A_{nn} \end{vmatrix}$$

$[C_1' = C_1 + a_{12}C_2 + \cdots + a_{1n}C_n]$

$= 0$, since each element in the first column is 0.

Since $a_{11} \neq 0$, det $(A_{ij}) = 0$.

Hence det $(A_{ij}) = (\det A)^{n-1}$, in this subcase also.

Thus the theorem is proved in all cases.

Definition. If $A = (a_{ij})$ be a square matrix and A_{ij} be the cofactor of a_{ij} in det A then det (A_{ij}) is said to be the *adjoint* of det A.

Theorem 2.4.3. Let (a_{ij}) be an $n \times n$ matrix and A_{ij} be the cofactor of a_{ij} in det (a_{ij}). If det $(a_{ij}) = 0$ then any two non-zero rows (non-zero columns) of A_{ij} are proportional.

Proof. Let the rth row and the kth row of the matrix (A_{ij}) be non-zero rows.

Let us consider the system of equations

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0$$
$$\cdots \qquad \cdots$$
$$a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0.$$

It is clearly satisfied by $x_1 = A_{r1}, x_2 = A_{r2}, \ldots, x_n = A_{rn}$.

Also we have

$$x_2.M_{k1} = \begin{vmatrix} a_{12}x_2 & a_{13} & \cdots & a_{1n} \\ a_{22}x_2 & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n2}x_2 & a_{n3} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} -a_{11}x_1 & a_{13} & \cdots & a_{1n} \\ -a_{21}x_1 & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1}x_1 & a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

$[C_1' = C_1 + x_3C_2 + x_4C_3 + \cdots + x_nC_{n-1}]$

$= -x_1M_{k2}$.

Similarly, $x_3.M_{k1} = -x_1.(-1)M_{k3}, \ldots, x_n.M_{k1} = -x_1.(-1)^{n-2}M_{kn}$

or, $x_2.(-1)^{k+1}M_{k1} = x_1.(-1)^{k+2}M_{k2}$,

$x_3.(-1)^{k+1}M_{k1} = x_1.(-1)^{k+3}M_{k3}$,

$\cdots \quad \cdots \quad \cdots$

$x_n.(-1)^{k+1}M_{k1} = x_1.(-1)^{k+n}M_{kn}$

or, $x_2.A_{k1} = x_1.A_{k2}$, $x_3.A_{k1} = x_1.A_{k3}$, $\cdots \cdots$ $x_n.A_{k1} = x_1.A_{kn}$.

Therefore $A_{r2}A_{k1} = A_{r1}A_{k2}, A_{r3}A_{k1} = A_{r1}A_{k3}, \ldots, A_{rn}A_{k1} = A_{r1}A_{kn}$.

If none of $A_{r1}, A_{r2}, \ldots, A_{rn}$ be zero, then $\frac{A_{k1}}{A_{r1}} = \frac{A_{k2}}{A_{r2}} = \cdots = \frac{A_{kn}}{A_{rn}}$.

If $A_{rj} = 0$ then $x_j = 0$ and therefore $A_{kj} = 0$.

Thus the rth row and the kth row of the matrix (A_{ij}) are proportional. This completes the proof.

Theorem 2.4.4. (Jacobi)

If M be a minor of order r of a square matrix $A = (a_{ij})$ and M^* be the corresponding minor of (A_{ij}), then $M^* = (\det A)^{r-1}.\bar{M}$, where $\bar{M}$ is the algebraic complement of M in det A.

Proof. When $r = 1$, let $M = a_{ij}$. Then $\bar{M} = A_{ij}$ and $M^* = A_{ij}$. Therefore the theorem holds in this case.

When $r \geq 2$, we consider the following cases.

Case 1. det $A \neq 0$.

Let the minor M be obtained from r rows i_1th, i_2th, $\ldots, i_r$th $(i_1 < i_2 < \cdots < i_r)$ and r columns j_1th, j_2th, $\ldots, j_r$th $(j_1 < j_1 < \cdots < j_r)$.

Let M' be the complementary minor of M and it is obtained from $n - r$ rows p_1th, p_2th, $\ldots, p_{n-r}$th $(p_1 < p_2 < \cdots < p_{n-r})$ and $n-r$ columns q_1th, q_2th, $\ldots, q_{n-r}$th $(q_1 < q_1 < \cdots < q_{n-r})$.

Then $\bar{M} = (-1)^{i_1+i_2+\cdots+i_r+j_1+j_2+\cdots+j_r} M' = (-1)^{\Sigma i + \Sigma j} M'$.

$$M = \begin{vmatrix} a_{i_1j_1} & \cdots & a_{i_1j_r} \\ \cdots & \cdots & \cdots \\ a_{i_rj_1} & \cdots & a_{i_rj_r} \end{vmatrix}, \bar{M} = (-1)^{\Sigma i+\Sigma j} \begin{vmatrix} a_{p_1q_1} & \cdots & a_{p_1q_{n-r}} \\ \cdots & \cdots & \cdots \\ a_{p_{n-r}q_1} & \cdots & a_{p_{n-r}q_{n-r}} \end{vmatrix}.$$

$$\det A = (-1)^{(i_1-1)+(i_2-2)+\cdots+(i_r-r)} \begin{vmatrix} a_{i_11} & a_{i_12} & \cdots & a_{i_1n} \\ \cdots & \cdots & \cdots \\ a_{i_r1} & a_{i_r2} & \cdots & a_{i_rn} \\ a_{p_11} & a_{p_12} & \cdots & a_{p_1n} \\ \cdots & \cdots & \cdots \\ a_{p_{n-r}1} & a_{p_{n-r}2} & \cdots & a_{p_{n-r}n} \end{vmatrix}$$

[the row i_1 is brought to the row 1 by $i_1 - 1$ interchanges of rows, the row i_2 to the row 2 by $i_2 - 2$ interchanges of rows and so on]

$$= (-1)^{(i_1-1)+(i_2-2)+\cdots+(i_r-r)+(j_1-1)+(j_2-2)+\cdots+(j_r-r)}.$$
$$\begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_r} & a_{i_1q_1} & \cdots & a_{i_1q_{n-r}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_r} & a_{i_rq_1} & \cdots & a_{i_rq_{n-r}} \\ a_{p_1j_1} & a_{p_1j_2} & \cdots & a_{p_1j_r} & a_{p_1q_1} & \cdots & a_{p_1q_{n-r}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{p_{n-r}j_1} & a_{p_{n-r}j_2} & \cdots & a_{p_{n-r}j_r} & a_{p_{n-r}q_1} & \cdots & a_{p_{n-r}q_{n-r}} \end{vmatrix}$$

[the column j_1 is brought to the column 1 by $j_1 - 1$ interchanges of columns, the column j_2 to the column 2 by $j_2 - 2$ interchanges of columns and so on]

$$= (-1)^{\Sigma i+\Sigma j} \begin{vmatrix} a_{i_1j_1} & a_{i_1j_2} & \cdots & a_{i_1j_r} & a_{i_1q_1} & \cdots & a_{i_1q_{n-r}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{i_rj_1} & a_{i_rj_2} & \cdots & a_{i_rj_r} & a_{i_rq_1} & \cdots & a_{i_rq_{n-r}} \\ a_{p_1j_1} & a_{p_1j_2} & \cdots & a_{p_1j_r} & a_{p_1q_1} & \cdots & a_{p_1q_{n-r}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{p_{n-r}j_1} & a_{p_{n-r}j_2} & \cdots & a_{p_{n-r}j_r} & a_{p_{n-r}q_1} & \cdots & a_{p_{n-r}q_{n-r}} \end{vmatrix}.$$

$$M^* = \begin{vmatrix} A_{i_1j_1} & \cdots & A_{i_1j_r} \\ \cdots & \cdots & \cdots \\ A_{i_rj_1} & \cdots & A_{i_rj_r} \end{vmatrix} = \begin{vmatrix} A_{i_1j_1} & \cdots & A_{i_rj_1} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{i_1j_r} & \cdots & A_{i_rj_r} & 0 & \cdots & 0 \\ A_{i_1q_1} & \cdots & A_{i_rq_1} & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_{i_1q_{n-r}} & \cdots & A_{i_rq_{n-r}} & 0 & \cdots & 1 \end{vmatrix}.$$

Therefore $\det A . M^*$

$$= (-1)^{\Sigma i+\Sigma j} \begin{vmatrix} \det A & 0 & \cdots & 0 & a_{i_1q_1} & \cdots & a_{i_1q_{n-r}} \\ 0 & \det A & \cdots & 0 & a_{i_2q_1} & \cdots & a_{i_2q_{n-r}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \det A & a_{i_rq_1} & \cdots & a_{i_rq_{n-r}} \\ 0 & 0 & \cdots & 0 & a_{p_1q_1} & \cdots & a_{p_1q_{n-r}} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & a_{p_{n-r}q_1} & \cdots & a_{p_{n-r}q_{n-r}} \end{vmatrix}$$

$$= (-1)^{\Sigma i+\Sigma j} (\det A)^r \begin{vmatrix} a_{p_1q_1} & \cdots & a_{p_1q_{n-r}} \\ \cdots & \cdots & \cdots \\ a_{p_{n-r}q_1} & \cdots & a_{p_{n-r}q_{n-r}} \end{vmatrix}$$

$= (\det A)^r \bar{M}$. Therefore $M^* = (\det A)^{r-1}\bar{M}$, since $\det A \neq 0$.

Case 2. det $A = 0$.

In this case every second order minor of (A_{ij}) vanishes by Theorem 2.4.3. Therefore $M^* = 0$ for $r \geq 2$ and therefore $M^* = (\det A)^{r-1}\bar{M}$.

This completes the proof.

Worked Examples (continued).

3. If $A, B, C, \ldots$ be the cofactors of $a, b, c, \ldots$ in

$$\Delta = \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}, \text{ prove that } BC - F^2 = a\Delta.$$

$$\text{The adjoint of } \Delta = \Delta' = \begin{vmatrix} A & H & G \\ H & B & F \\ G & F & C \end{vmatrix}.$$

Let us consider the minor $\begin{vmatrix} B & F \\ F & C \end{vmatrix}$ in Δ' and the corresponding minor $\begin{vmatrix} b & f \\ f & c \end{vmatrix}$ in Δ. The algebraic complement of $\begin{vmatrix} b & f \\ f & c \end{vmatrix}$ in Δ is a.

By the theorem, $\begin{vmatrix} B & F \\ F & C \end{vmatrix} = a\Delta$, i.e., $BC - F^2 = a\Delta$.

Note. Similarly, $CA - G^2 = b\Delta$, $AB - H^2 = c\Delta$, $GH - AF = f\Delta$, $HF - BG = g\Delta$ and $FG - CH = h\Delta$.

In particular, if $\Delta = 0$ then $BC = F^2$, $CA = G^2$, $AB = H^2$, $GH = AF$, $HF = BG$ and $FG = CH$.

4. Prove that

$$\begin{vmatrix} bc - a^2 & ca - b^2 & ab - c^2 \\ ca - b^2 & ab - c^2 & bc - a^2 \\ ab - c^2 & bc - a^2 & ca - b^2 \end{vmatrix} = (a^3 + b^3 + c^3 - 3abc)^2.$$

$$\text{We have } \Delta = \begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} = -(a^3 + b^3 + c^3 - 3abc).$$

$$\text{The adjoint of } \Delta = \begin{vmatrix} +\begin{vmatrix} c & a \\ a & b \end{vmatrix} & -\begin{vmatrix} b & a \\ c & b \end{vmatrix} & +\begin{vmatrix} b & c \\ c & a \end{vmatrix} \\ -\begin{vmatrix} b & c \\ a & b \end{vmatrix} & +\begin{vmatrix} a & c \\ c & b \end{vmatrix} & -\begin{vmatrix} a & b \\ c & a \end{vmatrix} \\ +\begin{vmatrix} b & c \\ c & a \end{vmatrix} & -\begin{vmatrix} a & c \\ b & a \end{vmatrix} & +\begin{vmatrix} a & b \\ b & c \end{vmatrix} \end{vmatrix}$$

$$= \begin{vmatrix} bc - a^2 & ca - b^2 & ab - c^2 \\ ca - b^2 & ab - c^2 & bc - a^2 \\ ab - c^2 & bc - a^2 & ca - b^2 \end{vmatrix}$$

But adjoint of $\Delta = \Delta^2$. Therefore L.H.S. = R.H.S.

2.5. Cramer's Rule.

Let

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \quad \cdots \quad \cdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

be a system of n linear equations in n unknowns $x_1, x_2, \ldots, x_n$, where $\det A = \det\,(a_{ij}) \neq 0$. Then there exists a *unique solution* of the system given by

$$x_1 = \frac{\det A_1}{\det A},\; x_2 = \frac{\det A_2}{\det A}, \ldots,\; x_n = \frac{\det A_n}{\det A},$$

where A_i is the $n \times n$ matrix obtained from A by replacing its ith column by the column $\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$, $i = 1, 2, \ldots, n$.

Proof. $x_1 \det A = \begin{vmatrix} x_1a_{11} & a_{12} & \cdots & a_{1n} \\ x_1a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \\ x_1a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$

$$= \begin{vmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} & a_{12} & \cdots & a_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \\ x_1a_{n1} + x_2a_{n2} + \cdots + x_na_{nn} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$[\,C_1' = C_1 + x_2C_2 + \cdots + x_nC_n\,]$$

$$= \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \det A_1.$$

$$\cdots \quad \cdots \quad \cdots$$

$$\cdots \quad \cdots \quad \cdots$$

$$\cdots \quad \cdots \quad \cdots$$

$$x_n \det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & x_na_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & x_na_{2n} \\ \cdots & \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & x_na_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & x_1a_{11}+x_2a_{12}+\cdots+x_na_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n-1} & x_1a_{21}+x_2a_{22}+\cdots+x_na_{2n} \\ \cdots & & \cdots & & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & x_1a_{n2}+x_2a_{n2}+\cdots+x_na_{nn} \end{vmatrix}$$

$$[C'_n = x_1C_1 + x_2C_2 + \cdots + x_{n-1}C_{n-1} + C_n]$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n-1} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n-1} & b_2 \\ \cdots & \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn-1} & b_n \end{vmatrix} = \det A_n.$$

Since $\det A \neq 0$, $x_1 = \frac{\det A_1}{\det A}$, $x_2 = \frac{\det A_2}{\det A}, \ldots, x_n = \frac{\det A_n}{\det A}$.

Note. A is said to be the *coefficient matrix* of the system.

Worked Example.

1. Solve the system of equations by Cramer's rule.

$$\begin{aligned} x+2y-3z &= 1 \\ 2x-y+z &= 4 \\ x+3y &= 5. \end{aligned}$$

Here the coefficient matrix is $A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ 1 & 3 & 0 \end{pmatrix}$

and $\det A = \begin{vmatrix} 1 & 2 & -3 \\ 2 & -1 & 1 \\ 1 & 3 & 0 \end{vmatrix} = -22 \neq 0.$

By Cramer's rule, there exists a unique solution for x, y, z and

$$x = \frac{\begin{vmatrix} 1 & 2 & -3 \\ 4 & -1 & 1 \\ 5 & 3 & 0 \end{vmatrix}}{\det A} = \frac{-44}{-22} = 2,$$

$$y = \frac{\begin{vmatrix} 1 & 1 & -3 \\ 2 & 4 & 1 \\ 1 & 5 & 0 \end{vmatrix}}{\det A} = \frac{-22}{-22} = 1,$$

$$z = \frac{\begin{vmatrix} 1 & 2 & 1 \\ 2 & -1 & 4 \\ 1 & 3 & 5 \end{vmatrix}}{\det A} = \frac{-22}{-22} = 1.$$

The solution is given by $x = 2$, $y = 1$, $z = 1$.

If $b_1 = b_2 = \cdots = b_n = 0$, the system of equations is said to be a *homogeneous* system. $x_1 = 0, x_2 = 0, \ldots, x_n = 0$ is an obvious solution of the system. This solution is said to be the *zero solution* of the system. It may be that a homogeneous system, in addition to its zero solution, admits of a non-zero solution.

For example, the system of equations

$$\begin{aligned} x+y+2z &= 0 \\ x-2y-z &= 0 \\ 2x-y+z &= 0 \end{aligned}$$

has a non-zero solution $x = 1, y = 1, z = -1$.

It is also to be noted that Cramer's rule is not applicable to the system, because $\begin{vmatrix} 1 & 1 & 2 \\ 1 & -2 & -1 \\ 2 & -1 & 1 \end{vmatrix} = 0.$

If, however, $\det(a_{ij}) \neq 0$, then the homogeneous system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots \quad \cdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

admits of a unique solution by Cramer's rule, and the unique solution is the zero solution.

Theorem 2.5.1. If the homogeneous system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots \quad \cdots & \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

has a non-zero solution, then $\det(a_{ij}) = 0$.

Proof. Let $x_1 = t_1$, $x_2 = t_2, \ldots, x_n = t_n$ be a non-zero solution of the system.

Then
$$\begin{aligned} a_{11}t_1 + a_{12}t_2 + \cdots + a_{1n}t_n &= 0 \\ a_{21}t_1 + a_{22}t_2 + \cdots + a_{2n}t_n &= 0 \\ \cdots \quad \cdots & \\ a_{n1}t_1 + a_{n2}t_2 + \cdots + a_{nn}t_n &= 0. \end{aligned}$$

At least one of the t's is non-zero. Let $t_1 \neq 0$.

Then $t_1 \det(a_{ij}) = \begin{vmatrix} t_1a_{11} & a_{12} & \cdots & a_{1n} \\ t_1a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \\ t_1a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$

$$= \begin{vmatrix} t_1a_{11}+t_2a_{12}+\cdots+t_na_{1n} & a_{12} & \cdots & a_{1n} \\ t_1a_{21}+t_2a_{22}+\cdots+t_na_{2n} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \\ t_1a_{n1}+t_2a_{n2}+\cdots+t_na_{nn} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0.$$

$$[C_1' = C_1 + t_2C_2 + \cdots + t_nC_n]$$

Therefore det $(a_{ij}) = 0$, since $t_1 \neq 0$.

Note. The theorem says that a *necessary condition* for the existence of a *non-zero solution* of the homogeneous system is det $(a_{ij}) = 0$.

We shall see later that det $(a_{ij}) = 0$ is also a *sufficient condition* for the existence of a *non-zero solution* of the homogeneous system.

2.6. Symmetric and Skew symmetric determinants.

If A be a symmetric matrix, then det A is said to be a *symmetric determinant.*

If A be a skew symmetric matrix, then det A is said to be a *skew symmetric determinant.*

Theorem 2.6.1. The adjoint of a symmetric determinant is a symmetric determinant.

Proof. Let $A = (a_{ij})$ be a symmetric matrix of order n. Then det A is a symmetric determinant of order n.

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, a_{ij} = a_{ji}.$$

Let A_{ij} be the cofactor of a_{ji} in det A.

$$\text{Then } A_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & a_{2j+1} & \cdots & a_{2n} \\ \cdots & & & \cdots & \cdots & & \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \cdots & & & \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}$$

$$A_{ji} = (-1)^{j+i} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i-1} & a_{1i+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i-1} & a_{2i+1} & \cdots & a_{2n} \\ \cdots & & & \cdots & \cdots & & \\ a_{j-11} & a_{j-12} & \cdots & a_{j-1i-1} & a_{j-1i+1} & \cdots & a_{j-1n} \\ a_{j+11} & a_{j+12} & \cdots & a_{j+1i-1} & a_{j+1i+1} & \cdots & a_{j+1n} \\ \cdots & & & \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{ni-1} & a_{ni+1} & \cdots & a_{nn} \end{vmatrix}$$

$$= (-1)^{j+i} \begin{vmatrix} a_{11} & a_{21} & \cdots & a_{i-11} & a_{i+11} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{i-12} & a_{i+12} & \cdots & a_{n2} \\ \cdots & & & \cdots & \cdots & & \\ a_{1j-1} & a_{2j-1} & \cdots & a_{i-1j-1} & a_{i+1j-1} & \cdots & a_{nj-1} \\ a_{1j+1} & a_{2j+1} & \cdots & a_{i-1j+1} & a_{i+1j+1} & \cdots & a_{nj+1} \\ \cdots & & & \cdots & \cdots & & \\ a_{1n} & a_{2n} & \cdots & a_{i-1n} & a_{i+1n} & \cdots & a_{nn} \end{vmatrix},$$

since $a_{ij} = a_{ji}$

$= A_{ij}$, since det A = det A^t.

$$\text{The adjoint of det } A = \begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{vmatrix}.$$

Since $A_{ij} = A_{ji}$, the adjoint of det A is symmetric.

Theorem 2.6.2. The adjoint of a skew symmetric determinant of order n is symmetric if n be odd and skew symmetric if n be even.

Proof. Let $A = (a_{ij})$ be a skew symmetric matrix of order n. Then det A is a skew symmetric determinant of order n.

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}, a_{ij} = -a_{ji}.$$

Let A_{ij} be the cofactor of a_{ij} in det A. Then

$$A_{ij} = (-1)^{i+j} \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j-1} & a_{2j+1} & \cdots & a_{2n} \\ \cdots & & & \cdots & \cdots & & \\ a_{i-11} & a_{i-12} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n} \\ a_{i+11} & a_{i+12} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n} \\ \cdots & & & \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn} \end{vmatrix}.$$

$$A_{ji}=(-1)^{j+i}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1i-1} & a_{1i+1} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i-1} & a_{2i+1} & \cdots & a_{2n} \\ & \cdots & & \cdots & \cdots & & \\ a_{j-11} & a_{j-12} & \cdots & a_{j-1i-1} & a_{j-1i+1} & \cdots & a_{j-1n} \\ a_{j+11} & a_{j+12} & \cdots & a_{j+1i-1} & a_{j+1i+1} & \cdots & a_{j+1n} \\ & \cdots & & \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{ni-1} & a_{ni+1} & \cdots & a_{nn} \end{vmatrix}$$

$$=(-1)^{j+i}\begin{vmatrix} -a_{11} & -a_{21} & \cdots & -a_{i-11} & -a_{i+11} & \cdots & -a_{n1} \\ -a_{12} & -a_{22} & \cdots & -a_{i-12} & -a_{i+12} & \cdots & -a_{n2} \\ & \cdots & & \cdots & \cdots & & \\ -a_{1j-1} & -a_{2j-1} & \cdots & -a_{i-1j-1} & -a_{i+1j-1} & \cdots & -a_{nj-1} \\ -a_{1j+1} & -a_{2j+1} & \cdots & -a_{i-1j+1} & -a_{i+1j+1} & \cdots & -a_{nj+1} \\ & \cdots & & \cdots & \cdots & & \\ -a_{1n} & -a_{2n} & \cdots & -a_{i-1n} & -a_{i+1n} & \cdots & -a_{nn} \end{vmatrix},$$

since $a_{ij}=-a_{ji}$

$$=(-1)^{j+i}.(-1)^{n-1}\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{i-11} & a_{i+11} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{i-12} & a_{i+12} & \cdots & a_{n2} \\ & \cdots & & \cdots & \cdots & & \\ a_{1j-1} & a_{2j-1} & \cdots & a_{i-1j-1} & a_{i+1j-1} & \cdots & a_{nj-1} \\ a_{1j+1} & a_{2j+1} & \cdots & a_{i-1j+1} & a_{i+1j+1} & \cdots & a_{nj+1} \\ & \cdots & & \cdots & \cdots & & \\ a_{1n} & a_{2n} & \cdots & a_{i-1n} & a_{i+1n} & \cdots & a_{nn} \end{vmatrix}$$

$=(-1)^{n-1}A_{ij}$, since $\det A=\det A^t$.

The adjoint of $\det A = \begin{vmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ & \cdots & \cdots & \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{vmatrix}$.

If n be odd, $A_{ij}=A_{ji}$ and therefore the adjoint of det A is symmetric. If n be even, $A_{ij}=-A_{ji}$ and therefore the adjoint of det A is skew symmetric.

Theorem 2.6.3. A skew symmetric determinant of odd order is zero.

Proof. Let $A=(a_{ij})$ be a skew symmetric matrix of order n and n is odd. Since (a_{ij}) is skew symmetric, $a_{ij}=-a_{ji}$.

$$\det A=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\begin{vmatrix} -a_{11} & -a_{21} & \cdots & -a_{n1} \\ -a_{12} & -a_{22} & \cdots & -a_{n2} \\ & \cdots & \cdots & \\ -a_{1n} & -a_{2n} & \cdots & -a_{nn} \end{vmatrix},$$

$$=(-1)^n\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & \cdots & \cdots & \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix}=(-1)^n \det A^t$$

$=-\det A$, since n is odd.

Therefore $2\det A=0$ and this implies $\det A=0$.

Note. If the ground field of scalars be of characteristic 2, $2\det A=0$ does not necessarily imply $\det A=0$. So the theorem is not valid if the elements of A are taken from a field of characteristic 2.

In such a field, $A=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is skew symmetric but $\det A\neq 0$.

Theorem 2.6.4. A skew symmetric determinant of even order is the square of a polynomial function of its elements.

Proof. Let $A=(a_{ij})$ be a skew symmetric matrix of order n and n is even. Then $a_{ij}=-a_{ji}$.

Let $\Delta_n=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$

Let A_{ij} be the cofactor of a_{ij} in det A. Then $A_{ij}=-A_{ji}$. We have

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}\begin{vmatrix} 1 & 0 & \cdots & 0 & A_{n-11} & A_{n1} \\ 0 & 1 & \cdots & 0 & A_{n-12} & A_{n2} \\ \cdots & \cdots & & & \cdots & \\ 0 & 0 & \cdots & 1 & A_{n-1n-2} & A_{nn-2} \\ 0 & 0 & \cdots & 0 & A_{n-1n-1} & A_{nn-1} \\ 0 & 0 & \cdots & 0 & A_{n-1n} & A_{nn} \end{vmatrix}$$

$$=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n-2} & 0 & 0 \\ a_{21} & a_{22} & \cdots & a_{2n-2} & 0 & 0 \\ & & \cdots & & \cdots & \\ a_{n-21} & a_{n-22} & \cdots & a_{n-2n-2} & 0 & 0 \\ a_{n-11} & a_{n-12} & \cdots & a_{n-1n-2} & \Delta_n & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn-2} & 0 & \Delta_n \end{vmatrix}.$$

L.H.S. $=\Delta_n\begin{vmatrix} A_{n-1n-1} & A_{nn-1} \\ A_{n-1n} & A_{nn} \end{vmatrix}$ and R.H.S. $=\Delta_{n-2}.\Delta_n^2$.

But $A_{n-1n-1}=0$, $A_{nn}=0$, $A_{nn-1}=-A_{n-1n}$, since n is even.

Therefore $\Delta_n.A_{nn-1}^2=\Delta_{n-2}.\Delta_n^2$.

If $\Delta_n=0$, the theorem is trivially true.

If $\Delta_n \neq 0$, $\Delta_n.\Delta_{n-2} = {A_{nn-1}}^2$.

But $\Delta_2 = \begin{vmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{vmatrix} = {a_{12}}^2$ and it is a perfect square.

By the principle of induction, Δ_n is perfect square of a polynomial function of its elements.

We give below the proof of the particular case when the order of the matrix is 4.

$$\text{Let } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}, \; a_{ij} = -a_{ji}.$$

Two cases come up for consideration.

Case 1. det $A = 0$.

In this case the theorem holds trivially.

Case 2. det $A \neq 0$.

In this case at least one element of the first row is non-zero. Without loss of generality we assume $a_{12} \neq 0$. Because if $a_{12} = 0$ and $a_{13} \neq 0$, then a_{13} can be brought to $(1,2)$ position by effecting the operations R_{23} (interchange of R_2 and R_3) and C_{23} (interchange of C_2 and C_3) and thereby the determinant remains skew symmetric.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} \begin{vmatrix} 1 & 0 & A_{31} & A_{41} \\ 0 & 1 & A_{32} & A_{42} \\ 0 & 0 & A_{33} & A_{43} \\ 0 & 0 & A_{34} & A_{44} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & \det A & 0 \\ a_{41} & a_{42} & 0 & \det A \end{vmatrix}$$

$$\text{or, } \det A \begin{vmatrix} A_{33} & A_{43} \\ A_{34} & A_{44} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} \det A & 0 \\ 0 & \det A \end{vmatrix}$$

[expanding by Laplace's method]

or, $\det A.{A_{43}}^2 = {a_{12}}^2 (\det A)^2$, since $A_{33} = A_{44} = 0, A_{34} = -A_{43}$, $a_{11} = a_{22} = 0, a_{12} = -a_{21}$

or, ${a_{12}}^2 \det A = {A_{43}}^2$, since $\det A \neq 0$

or, $\det A = \left(\frac{A_{43}}{a_{12}}\right)^2$, since $a_{12} \neq 0$. This completes the proof.

Exercises 2

1. *Prove* without expanding

$$\text{(i)} \begin{vmatrix} b-c & c-a & a-b \\ c-a & a-b & b-c \\ a-b & b-c & c-a \end{vmatrix} = 0, \quad \text{(ii)} \begin{vmatrix} 1 & 1 & 1 & 1 \\ a & b & c & d \\ b & c & d & a \\ c+d & d+a & a+b & b+c \end{vmatrix} = 0,$$

$$\text{(iii)} \begin{vmatrix} 2a & a+b & a+c & a+d \\ b+a & 2b & b+c & b+d \\ c+a & c+b & 2c & c+d \\ d+a & d+b & d+c & 2d \end{vmatrix} = 0, \quad \text{(iv)} \begin{vmatrix} 0 & a & b & c & 0 \\ -a & 0 & c & 0 & -d \\ -b & -c & 0 & -d & -e \\ -c & 0 & d & 0 & -f \\ 0 & d & e & f & 0 \end{vmatrix} = 0.$$

2. Prove without expanding

(i) $\begin{vmatrix} 3 & 4 & 5 \\ 4 & 3 & 7 \\ 5 & 7 & 5 \end{vmatrix}$ is divisible by 23, it is given that 345, 437, 575 are each divisible by 23;

$$\text{(ii)} \begin{vmatrix} a^3 & a^2 & 1 \\ b^3 & b^2 & 1 \\ c^3 & c^2 & 1 \end{vmatrix} = (ab+bc+ca) \begin{vmatrix} a^2 & a & 1 \\ b^2 & b & 1 \\ c^2 & c & 1 \end{vmatrix};$$

$$\text{(iii)} \begin{vmatrix} 1 & bcd & b+c+d & a^2 \\ 1 & cda & c+d+a & b^2 \\ 1 & dab & d+a+b & c^2 \\ 1 & abc & a+b+c & d^2 \end{vmatrix} = \begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix};$$

$$\text{(iv)} \begin{vmatrix} a^4 & a^2 & a & 1 \\ b^4 & b^2 & b & 1 \\ c^4 & c^2 & c & 1 \\ d^4 & d^2 & d & 1 \end{vmatrix} = (a+b+c+d) \begin{vmatrix} a^3 & a^2 & a & 1 \\ b^3 & b^2 & b & 1 \\ c^3 & c^2 & c & 1 \\ d^3 & d^2 & d & 1 \end{vmatrix}.$$

[**Hint.** $a^4 - a^3\Sigma a + a^2\Sigma ab - a\Sigma abc + abcd = 0$.]

3. Prove that

(i) $\begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix} = (a+b+c)(a+b\omega+c\omega^2)(a+b\omega^2+c\omega)$, where $\omega^3 = 1$;

[The determinant of this type is called a *circulant*.]

(ii) $\begin{vmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{vmatrix} = (a+b+c+d)(a-b+c-d)(a+bi-c-di)(a-bi-c+di)$;

(iii) $\begin{vmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{vmatrix} = (a+b+c+d)(a+b-c-d)(a-b+c-d)$
$(a-b-c+d)$.

4. Prove that

(i) $\begin{vmatrix} 1+a_1 & 1 & 1 & 1 \\ 1 & 1+a_2 & 1 & 1 \\ 1 & 1 & 1+a_3 & 1 \\ 1 & 1 & 1 & 1+a_4 \end{vmatrix} = a_1a_2a_3a_4\left(1+\frac{1}{a_1}+\frac{1}{a_2}+\frac{1}{a_3}+\frac{1}{a_4}\right)$.

(ii) $\begin{vmatrix} a+1 & a & a & a \\ a & a+2 & a & a \\ a & a & a+3 & a \\ a & a & a & a+4 \end{vmatrix} = 24\left(1+\frac{a}{1}+\frac{a}{2}+\frac{a}{3}+\frac{a}{4}\right)$.

5. Prove that

(i) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ \binom{m}{1} & \binom{m+1}{1} & \binom{m+2}{1} & \binom{m+3}{1} \\ \binom{m+1}{2} & \binom{m+2}{2} & \binom{m+3}{2} & \binom{m+4}{2} \\ \binom{m+2}{3} & \binom{m+3}{3} & \binom{m+4}{3} & \binom{m+5}{3} \end{vmatrix} = 1$, where m is a positive integer and $\binom{n}{r} = n_{c_r}$;

(ii) $\begin{vmatrix} 1 & m & m^2 & m^3 \\ m^3 & 1 & m & m^2 \\ m^2 & m^3 & 1 & m \\ m & m^2 & m^3 & 1 \end{vmatrix} = (1-m^4)^3$;

(iii) $\begin{vmatrix} a^3 & 3a^2 & 3a & 1 \\ a^2 & a^2+2a & 2a+1 & 1 \\ a & 2a+1 & a+2 & 1 \\ 1 & 3 & 3 & 1 \end{vmatrix} = (a-1)^6$;

(iv) $\begin{vmatrix} 1 & 2 & 3 & x \\ 2 & 3 & 4 & y \\ 3 & 4 & 5 & z \\ x & y & z & 0 \end{vmatrix} = (x-2y+z)^2$.

6. Prove that

(i) $\begin{vmatrix} (b+c)^2 & a^2 & a^2 \\ b^2 & (c+a)^2 & b^2 \\ c^2 & c^2 & (a+b)^2 \end{vmatrix} = 2abc(a+b+c)^3$;

(ii) $\begin{vmatrix} (b+c)^2 & c^2 & b^2 \\ c^2 & (c+a)^2 & a^2 \\ b^2 & a^2 & (a+b)^2 \end{vmatrix} = 2(bc+ca+ab)^3$;

[Hint. Let $a=\frac{1}{u}, b=\frac{1}{v}, c=\frac{1}{w}$. Then proceed as in (i)]

(iii) $\begin{vmatrix} b^2c^2+a^2d^2 & bc+ad & 1 \\ c^2a^2+b^2d^2 & ca+bd & 1 \\ a^2b^2+c^2d^2 & ab+cd & 1 \end{vmatrix} = (b-c)(c-a)(a-b)(a-d)(b-d)(c-d)$;

(iv) $\begin{vmatrix} m^3 & m^2 & m & 1 \\ (m+1)^3 & (m+1)^2 & (m+1) & 1 \\ (m+2)^3 & (m+2)^2 & (m+2) & 1 \\ (m+3)^3 & (m+3)^2 & (m+3) & 1 \end{vmatrix} = 12$.

7. Prove that

(i) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \\ \alpha^4 & \beta^4 & \gamma^4 & \delta^4 \end{vmatrix} = (\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta)(\alpha+$
$\beta+\gamma+\delta)$;

(ii) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \\ \alpha^3 & \beta^3 & \gamma^3 & \delta^3 \\ \alpha^4 & \beta^4 & \gamma^4 & \delta^4 \end{vmatrix} = (\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta)(\alpha\beta\gamma+$
$\beta\gamma\delta+\gamma\delta\alpha+\alpha\beta\gamma)$.

8. Prove that

(i) $\begin{vmatrix} a^2+x & ab & ac & ad \\ ab & b^2+x & bc & bd \\ ac & bc & c^2+x & cd \\ ad & bd & cd & d^2+x \end{vmatrix} = x^3(a^2+b^2+c^2+d^2+x)$;

(ii) $\begin{vmatrix} x & a & a & a \\ a & x & a & a \\ a & a & x & a \\ a & a & a & x \end{vmatrix} = (x-a)^3(x+3a)$.

9. Prove that

(i) $\begin{vmatrix} -2a & a+b & a+c \\ b+a & -2b & b+c \\ c+a & c+b & -2c \end{vmatrix} = 4(b+c)(c+a)(a+b)$;

(ii) $\begin{vmatrix} a^2 & (s-a)^2 & (s-a)^2 \\ (s-b)^2 & b^2 & (s-b)^2 \\ (s-c)^2 & (s-c)^2 & c^2 \end{vmatrix} = 2s^3(s-a)(s-b)(s-c)$,
where $2s = a+b+c$.

10. Prove that

(i) $$\begin{vmatrix} a\alpha & b\beta & c\gamma & 0 \\ b\beta & a\alpha & 0 & c\gamma \\ c\gamma & 0 & a\alpha & b\beta \\ 0 & c\gamma & b\beta & a\alpha \end{vmatrix} = \begin{vmatrix} \alpha^2 & b^2 & c^2 & 0 \\ \beta^2 & a^2 & 0 & c^2 \\ \gamma^2 & 0 & a^2 & b^2 \\ 0 & \gamma^2 & \beta^2 & \alpha^2 \end{vmatrix} = -16s(s-a\alpha)(s-b\beta)$$
$(s-c\gamma)$, where $2s = a\alpha + b\beta + c\gamma$;

(ii) $$\begin{vmatrix} 1 & b^2 & c^2 & 0 \\ 1 & a^2 & 0 & c^2 \\ 1 & 0 & a^2 & b^2 \\ 0 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} a & b & c & 0 \\ b & a & 0 & c \\ c & 0 & a & b \\ 0 & c & b & a \end{vmatrix} = -(a+b+c)(b+c-a)(c+$$
$a-b)(a+b-c)$.

11. Expand by Laplace's method to prove that

(i) $$\begin{vmatrix} 0 & 0 & 0 & a & b \\ 0 & 0 & e & a & b \\ 0 & d & 0 & a & 0 \\ c & 0 & 0 & 0 & b \\ c & d & e & 0 & 0 \end{vmatrix} = 0;$$

(ii) $$\begin{vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{vmatrix} = (a^2+b^2+c^2+d^2)^2.$$

12. Expand $\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{vmatrix}$ by Laplace's method to prove that

$p_{12}p_{34} + p_{13}p_{42} + p_{14}p_{23} = 0$, where $p_{ij} = \begin{vmatrix} x_i & x_j \\ y_i & y_j \end{vmatrix}, i = 1,2,3,4; j = 1,2,3,4$

13. Express as the product of two determinants and hence prove that

(i) $$\begin{vmatrix} (x-a)^2 & (x-b)^2 & (x-c)^2 \\ (y-a)^2 & (y-b)^2 & (y-c)^2 \\ (z-a)^2 & (z-b)^2 & (z-c)^2 \end{vmatrix} = 2(x-y)(y-z)(z-x)(a-b)(b-$$
$c)(c-a)$;

(ii) $$\begin{vmatrix} a^2+b^2 & ac & bc \\ ac & b^2+c^2 & ab \\ bc & ab & c^2+a^2 \end{vmatrix} = 4a^2b^2c^2;$$

(iii) $$\begin{vmatrix} 2 & a+b+c+d & ab+cd \\ a+b+c+d & 2(a+b)(c+d) & ab(c+d)+cd(a+b) \\ ab+cd & ab(c+d)+cd(a+b) & 2abcd \end{vmatrix} = 0.$$

14. Express as the product of two determinants and hence prove that

(i) $$\begin{vmatrix} S_0 & S_1 & S_2 & S_3 \\ S_1 & S_2 & S_3 & S_4 \\ S_2 & S_3 & S_4 & S_5 \\ 1 & x & x^2 & x^3 \end{vmatrix} = (x-a)(x-b)(x-c)(a-b)^2(b-c)^2(c-a)^2,$$
where $S_r = a^r + b^r + c^r$;

(ii) $$\begin{vmatrix} 0 & (a-b)^3 & (a-c)^3 & (a-d)^3 \\ (b-a)^3 & 0 & (b-c)^3 & (b-d)^3 \\ (c-a)^3 & (c-b)^3 & 0 & (c-d)^3 \\ (d-a)^3 & (d-b)^3 & (d-c)^3 & 0 \end{vmatrix} = 9(a-b)^2(a-c)^2(a-d)^2$$
$(b-c)^2(b-d)^2(c-d)^2$;

(iii) $$\begin{vmatrix} 0 & (a-b)^3 & (a-c)^3 & a^3 \\ (b-a)^3 & 0 & (b-c)^3 & b^3 \\ (c-a)^3 & (c-b)^3 & 0 & c^3 \\ -a^3 & -b^3 & -c^3 & 0 \end{vmatrix} = 9(a-b)^2(b-c)^2(c-a)^2a^2b^2c^2;$$

(iv) $$\begin{vmatrix} b^2+c^2+1 & c^2+1 & b^2+1 & b+c \\ c^2+1 & c^2+a^2+1 & a^2+1 & c+a \\ b^2+1 & a^2+1 & a^2+b^2+1 & a+b \\ b+c & c+a & a+b & 3 \end{vmatrix} = (ab+bc+ca)^2;$$

(v) $$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & a^2+x^2 & ab+xy & ac+xz \\ 1 & ab+xy & b^2+y^2 & bc+yz \\ 1 & ac+xz & bc+yz & c^2+z^2 \end{vmatrix} = -\begin{vmatrix} 1 & a & x \\ 1 & b & y \\ 1 & c & z \end{vmatrix}^2.$$

15. Evaluate $\begin{vmatrix} a+ib & c+id \\ -c+id & a-ib \end{vmatrix} \begin{vmatrix} p+iq & r+is \\ -r+is & p-iq \end{vmatrix}$, where $i = \sqrt{-1}$

and hence express $(a^2+b^2+c^2+d^2)(p^2+q^2+r^2+s^2)$ as the sum of four squares.

16. Evaluate $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \begin{vmatrix} p & q & r \\ q & r & p \\ r & p & q \end{vmatrix}$ and hence express

$(a^3+b^3+c^3-3abc)(p^3+q^3+r^3-3pqr)$ in the form $l^3+m^3+n^3-3lmn$.

17. Evaluate $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \begin{vmatrix} a & b & c \\ c & a & b \\ b & c & a \end{vmatrix}$ and hence prove that

$(a^3+b^3+c^3-3abc)^2 = x^3+y^3+z^3-3xyz$, where $x = a^2+2bc$, $y = b^2+2ca$, $z = c^2+2ab$.

18. Evaluate $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} \begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}$ and hence prove that

$(a^2+b^2+c^2-ab-bc-ca)(x^2+y^2+z^2-xy-yz-zx) = p^2+q^2+r^2-pq-qr-rp$, where $p = ax+by+cz$, $q = bx+cy+az$, $r = cx+ay+bz$.

19. Solve by Cramer's rule

(i) $\begin{aligned} x+y+z &= 6 \\ x+2y+3z &= 14 \\ x-y+z &= 2, \end{aligned}$ (ii) $\begin{aligned} x+y+z &= 1 \\ ax+by+cz &= 1 \\ a^2x+b^2y+c^2z &= 1, \quad a \neq b \neq c. \end{aligned}$

20. Solve for $\cos A$, $\cos B$, $\cos C$, if $\begin{aligned} a &= b\cos C + c\cos B \\ b &= c\cos A + a\cos C \\ c &= a\cos B + b\cos A, \end{aligned}$ and $abc \neq 0$.

21. The planes $x = cy+bz$, $y = az+cx$, $z = bx+ay$ have a common point other than the origin. Prove that $a^2+b^2+c^2+2abc = 1$.

22. If the system of equations $\begin{aligned} ax+by+cz &= 0 \\ bx+cy+az &= 0 \\ cx+ay+bz &= 0 \end{aligned}$

has non-zero solutions, prove that either $a+b+c=0$, or $a=b=c$.

23. Let $D_n = \det\,(a_{ij})_{n,n}$, where $a_{ij} = a^{|i-j|}, a > 0$.

Show that (i) $D_n = (1-a^2)D_{n-1}$ for $n \geq 2$; (ii) $D_n = (1-a^2)^{n-1}$.

24. Let $D_n = \det\,(a_{ij})_{n,n}$, where $\begin{aligned} a_{ij} &= 2 \text{ if } i = j \\ &= 1 \text{ if } |i-j| = 1 \\ &= 0 \text{ if } i \neq j \text{ and } |i-j| \neq 1. \end{aligned}$

Show that (i) $D_n - 2D_{n-1} + D_{n-2} = 0$ for $n \geq 3$; (ii) $D_n = n+1$.

[**Hint.** (ii) $D_n - D_{n-1} = D_{n-1} - D_{n-2} = \cdots = D_2 - D_1 = 1$.]

25. Let $D_n = \det\,(a_{ij})_{n,n}$, where $\begin{aligned} a_{ij} &= 1 \text{ if } i = j \\ &= \tfrac{1}{2} \text{ if } |i-j| = 1 \\ &= 0 \text{ if } i \neq j \text{ and } |i-j| \neq 1. \end{aligned}$

Show that (i) $D_n - D_{n-1} + \frac{1}{4}D_{n-2} = 0$ for $n \geq 3$; (ii) $D_n = \frac{n+1}{2^n}$.

[**Hint.** (ii) $D_n - \frac{1}{2}D_{n-1} = \frac{1}{2}[D_{n-1} - \frac{1}{2}D_{n-2}] = \frac{1}{2^2}[D_{n-2} - \frac{1}{2}D_{n-3}] = \cdots = \frac{1}{2^{n-2}}[D_2 - \frac{1}{2}D_1] = \frac{1}{2^{n-2}}(\frac{1}{4}) = \frac{1}{2^n}$.]

26. Let $D_n = \det\,(a_{ij})_{n,n}$, where $\begin{aligned} a_{ij} &= 1+a^2 \text{ if } i = j \\ &= a \text{ if } |i-j| = 1 \\ &= 0 \text{ if } i \neq j \text{ and } |i-j| \neq 1. \end{aligned}$

Show that (i) $D_n - (1+a^2)D_{n-1} + a^2D_{n-2} = 0$ for $n \geq 3$;

(ii) $D_n = 1 + a^2 + a^4 + \cdots + a^{2n}$.

[**Hint.** (ii) $D_n - a^2D_{n-1} = D_{n-1} - a^2D_{n-2} = \cdots = D_2 - a^2D_1 = 1$.]

3. Matrices (II)

3.1. Adjoint of a square matrix.

Let $A = (a_{ij})$ be a square matrix. Let A_{ij} be the cofactor of a_{ij} in $\det A$. The transpose of the matrix (A_{ij}) is said to be the *adjoint* (or *adjugate*) of A and is denoted by *adj* A.

Worked Example.

1. Compute *adj* A where $A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix}$.

$$adj\,A = \begin{pmatrix} +\begin{vmatrix} 4 & 5 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 3 & 4 \end{vmatrix} & +\begin{vmatrix} 0 & 1 \\ 4 & 5 \end{vmatrix} \\ -\begin{vmatrix} 3 & 5 \\ 2 & 4 \end{vmatrix} & +\begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 3 & 5 \end{vmatrix} \\ +\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix} & -\begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix} & +\begin{vmatrix} 1 & 0 \\ 3 & 4 \end{vmatrix} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 3 & -4 \\ -2 & 2 & -2 \\ 1 & -3 & 4 \end{pmatrix}.$$

Properties.

1. $adj\,(A^t) = (adj\,A)^t$.

Proof. Let $A = (a_{ij})$, $A^t = B = (b_{ij})$.

B_{ij} = the cofactor of b_{ij} in $\det B$
= the cofactor of a_{ji} in $\det A$
$= A_{ji}$.

So $adj\,(A^t) = adj\,B = (B_{ij})^t = (A_{ji})^t = \{(A_{ij})^t\}^t = (adj\,A)^t$.

2. If A be an $n \times n$ matrix and c be a scalar, $adj\,(cA) = c^{n-1} adj\,A$.

Proof. Let $A = (a_{ij})$. Then $cA = (ca_{ij})$.

Let A_{ij} be the cofactor of a_{ij} in det A. Then the cofactor of ca_{ij} in det (cA) is $c^{n-1}A_{ij}$.
So $adj\ (cA) = (c^{n-1}A_{ij})^t = c^{n-1}(A_{ij})^t = c^{n-1} adj\ A$.

Worked Example (continued).

2. If A be an $n \times n$ matrix, prove that $adj\ (adj\ A) = (\det A)^{n-2} A$.

Let $A = (a_{ij})$ and $\bar{A} = (A_{ij})$, where A_{ij} is the cofactor of a_{ij} in det A. Then $adj\ A = (\bar{A})^t$.

Let A^*_{ij} be the cofactor of the element A_{ij} in det $\bar{A}$. Then $A^*_{ij} = (-1)^{i+j}.M^*_{ij}$, where M^*_{ij} is the minor of A_{ij} in det $\bar{A}$. Let M_{ij} be the minor of the element a_{ij} in det A. M^*_{ij} is a minor of order $n-1$ in $\bar{A}$.

By Jacobi's theorem, $M^*_{ij} = (\det A)^{n-2}$.(algebraic complement of M_{ij} in det A) $= (\det A)^{n-2}.(-1)^{i+j}.a_{ij}$. Therefore $A^*_{ij} = (\det A)^{n-2}.a_{ij}$ and $adj\ (adj\ A) = adj\ [(\bar{A})^t] = (adj\ \bar{A})^t = (A^*_{ij}) = (\det A)^{n-2}A$.

Theorem 3.1.1. Let A be a square matrix of order n. Then $A.adj\ A = adj\ A.A = (\det A)I_n$.

Proof. Let $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$.

Then $adj A = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$, where A_{ij} is the cofactor of a_{ij} in det A.

$$A.adj A = \begin{bmatrix} \sum_{k=1}^{n} a_{1k}A_{1k} & \sum_{k=1}^{n} a_{1k}A_{2k} & \cdots & \sum_{k=1}^{n} a_{1k}A_{nk} \\ \sum_{k=1}^{n} a_{2k}A_{1k} & \sum_{k=1}^{n} a_{2k}A_{2k} & \cdots & \sum_{k=1}^{n} a_{2k}A_{nk} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{k=1}^{n} a_{nk}A_{1k} & \sum_{k=1}^{n} a_{nk}A_{2k} & \cdots & \sum_{k=1}^{n} a_{nk}A_{nk} \end{bmatrix}$$

$$= \begin{pmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \det A \end{pmatrix} = (\det A)I_n,$$

$$\text{since } \sum_{k=1}^{n} a_{ik}A_{jk} = \det A, \text{ if } i = j$$
$$= 0, \text{ if } i \neq j.$$

$$adj A.A = \begin{bmatrix} \sum_{k=1}^{n} A_{k1}a_{k1} & \sum_{k=1}^{n} A_{k1}a_{k2} & \cdots & \sum_{k=1}^{n} A_{k1}a_{kn} \\ \sum_{k=1}^{n} A_{k2}a_{k1} & \sum_{k=1}^{n} A_{k2}a_{k2} & \cdots & \sum_{k=1}^{n} A_{k2}a_{kn} \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{k=1}^{n} A_{kn}a_{k1} & \sum_{k=1}^{n} A_{kn}a_{k2} & \cdots & \sum_{k=1}^{n} A_{kn}a_{kn} \end{bmatrix}$$

$$= \begin{pmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \det A \end{pmatrix} = (\det A)I_n,$$

$$\text{since } \sum_{k=1}^{n} a_{ki}A_{kj} = \det A, \text{ if } i = j$$
$$= 0, \text{ if } i \neq j.$$

Therefore $A.adj\ A = adj\ A.A = (\det A)I_n$.

Worked Example (continued).

3. Verify that $A.adj\ A = adj\ A.A$ and each is a scalar matrix, where $A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix}$.

From the previous example, $adj\ A = \begin{pmatrix} 1 & 3 & -4 \\ -2 & 2 & -2 \\ 1 & -3 & 4 \end{pmatrix}$.

$$A.adj\ A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 & -4 \\ 2 & -2 & 2 \\ 1 & -3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \text{ and}$$

$$adj\ A.A = \begin{pmatrix} 1 & 3 & -4 \\ 2 & -2 & 2 \\ 1 & -3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Therefore $A.adj\ A = adj\ A.A = 2I_3$, a scalar matrix.

3.2. Inverse of a matrix.

A square matrix A of order n is said to be *invertible* if there exists a matrix B such that $AB = BA = I_n$. B is said to be *an inverse* of A.

In order that both AB and BA should exist, B must be a square matrix of order n.

Theorem 3.2.1. An invertible matrix has a unique inverse.

Proof. Let A be an invertible matrix of order n. If possible, let B and C be two inverses of A.

Then $AB = BA = I_n$, $AC = CA = I_n$.
We have $C(AB) = (CA)B$, since multiplication is associative

$\Rightarrow CI_n = I_nB$

$\Rightarrow C = B$.

This proves that A has a unique inverse.

In view of this theorem we can now speak of *'the inverse'* of an invertible matrix. The inverse of A is denoted by A^{-1} and it satisfies the relation $AA^{-1} = A^{-1}A = I_n$.

Definition. A square matrix A is said to be ***non-singular*** if $\det A \neq 0$ and *singular* if $\det A = 0$.

Theorem 3.2.2. An $n \times n$ matrix A over a field F is invertible if and only if A is non-singular.

Proof. Let A be an $n \times n$ invertible matrix. Then there exists an $n \times n$ matrix B such that $AB = BA = I_n$.

We have $\det AB = \det I_n = 1$, or $\det A . \det B = 1$.

This implies that $\det A \neq 0$ and therefore A is non-singular.

Conversely, let A be a non-singular matrix of order n. Then $\det A \neq 0$

For an $n \times n$ matrix A, $A.adj\ A = adj\ A.A = (\det A)I_n$.

Since $\det A \neq 0$, $A.\frac{1}{\det A}(adj\ A). = \frac{1}{\det A}(adj\ A).A = I_n$.

From the definition of an inverse it follows that $\frac{1}{\det A}(adj\ A)$ is the inverse of A. Hence A is invertible.

Note 1. The theorem gives a clue to the determination of A^{-1}, when it exists. $A^{-1} = \frac{1}{\det A}(adj\ A)$.

Note 2. The inverse of cA, where c is a non-zero scalar and A is non-singular, is $c^{-1}A^{-1}$ since $(cA)(c^{-1}A^{-1}) = (c^{-1}A^{-1})(cA) = I_n$.

Note 3. Since $I_n.I_n = I_n.I_n = I_n$, it follows that $I_n^{-1} = I_n$.

It follows that the inverse of a scalar matrix cI_n, where c is a non-zero scalar, is $c^{-1}I_n$.

Worked Example.

1. If $A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix}$, find A^{-1}.

$$\det A = \begin{vmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{vmatrix} = 2 \neq 0.$$

Since A is non-singular, A^{-1} exists and $A^{-1} = \frac{1}{\det A}(adj A)$.

We have $adj A = \begin{pmatrix} 1 & 3 & -4 \\ 2 & -2 & 2 \\ 1 & -3 & 4 \end{pmatrix}$, from the previous example.

Therefore $A^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 3 & -4 \\ 2 & -2 & 2 \\ 1 & -3 & 4 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} & -2 \\ -1 & 1 & -1 \\ \frac{1}{2} & -\frac{3}{2} & 2 \end{pmatrix}$.

Theorem 3.2.3. If A and B be invertible matrices of the same order then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. Let A, B be $n \times n$ matrices.

Since A is invertible, A^{-1} exists and $AA^{-1} = A^{-1}A = I_n$ and also $\det A \neq 0$. Similarly, B^{-1} exists and $BB^{-1} = B^{-1}B = I_n$ and also $\det B \neq 0$.

We have $\det(AB) = \det A.\ \det B \neq 0$.

This shows that AB is non-singular and hence it is invertible.

Now $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI_n)A^{-1} = AA^{-1} = I_n$ and $(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = (B^{-1}I_n)B = B^{-1}B = I_n$.

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

From the definition of an inverse and its uniqueness, it follows that $B^{-1}A^{-1}$ is the inverse of AB. That is, $(AB)^{-1} = B^{-1}A^{-1}$.

Corollary 1. If $A_1, A_2, ..., A_p$ be invertible matrices of the same order, $(A_1A_2...A_p)^{-1} = A_p^{-1}A_{p-1}^{-1}...A_2^{-1}A_1^{-1}$.

2. If A be an invertible matrix and p be a positive integer, $[A.A....A(p \text{ factors})]^{-1} = A^{-1}.A^{-1}....A^{-1}(p \text{ factors})$.

That is, $(A^p)^{-1} = (A^{-1})^p$.

Theorem 3.2.4. If A be an invertible matrix then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

Proof. Let A be an $n \times n$ matrix. Since A is invertible, A^{-1} exists and $AA^{-1} = A^{-1}A = I_n$ and also $\det A \neq 0$.

We have $\det A . \det A^{-1} = \det(AA^{-1}) = \det I_n = 1$.

This shows that $\det A^{-1} \neq 0$ and hence A^{-1} is invertible.

Again $A^{-1}A = AA^{-1} = I_n$.

From the definition of an inverse and its uniqueness, it follows that A is the inverse of A^{-1}. That is, $(A^{-1})^{-1} = A$.

Integral powers of a square matrix.

We have defined in Art.1.2 positive integral powers of a square matrix A.

We define $A^o = I_n$, n being the order of the square matrix A.

We define negative integral powers of an *invertible* matrix A by $A^{-p} = A^{-1}.A^{-1}...A^{-1}$($p$ times), p being a positive integer.

With this definition of A^n in some restricted sense, the laws of indices for matrices

1. $A^m A^n = A^{m+n}$, **2.** $(A^m)^n = A^{mn}$

hold good where m, n are integers.

Theorem 3.2.5. If A be an invertible matrix, then A^t is invertible and $(A^t)^{-1} = (A^{-1})^t$.

Proof. Let A be an $n \times n$ invertible matrix.

Then A^{-1} exists and $AA^{-1} = A^{-1}A = I_n$ and also $\det A \neq 0$.

$\det A^t = \det A \neq 0$. Therefore A^t is non-singular and hence it is invertible.

We have $A^t.(A^{-1})^t = (A^{-1}A)^t = I_n{}^t = I_n$

and $(A^{-1})^t A^t = (AA^{-1})^t = I_n{}^t = I_n$.

Therefore $A^t(A^{-1})^t = (A^{-1})^t A^t = I_n$.

From the definition of an inverse and its uniqueness, it follows that $(A^{-1})^t$ is the inverse of A^t. That is, $(A^t)^{-1} = (A^{-1})^t$.

Theorem 3.2.6. If A and B be square matrices of the same order and A is non-singular, then the matrix equation

$$AX = B$$

has the unique solution $X = A^{-1}B$, and the matrix equation

$$YA = B$$

has the unique solution $Y = BA^{-1}$.

Proof. Since A is non-singular, A^{-1} exists. Since A^{-1} and B are square matrices of the same order, $A^{-1}B$ and BA^{-1} are well defined.

We have $A(A^{-1}B) = (AA^{-1}).B = B$.

This shows that $X = A^{-1}B$ is a solution of the equation $AX = B$.

Let $X = P$, $X = Q$ be two solutions of the equation $AX = B$.

Then $AP = B = AQ$.

$$\begin{aligned} AP = AQ &\Rightarrow A^{-1}(AP) = A^{-1}(AQ) \quad \text{[pre multiplication by } A^{-1}] \\ &\Rightarrow (A^{-1}A)P = (A^{-1}A)Q \\ &\Rightarrow P = Q. \end{aligned}$$

This proves the uniqueness of the solution of $AX = B$. Therefore $X = A^{-1}B$ is the unique solution of the equation $AX = B$.

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

This shows that $Y = BA^{-1}$ is a solution of the equation $YA = B$.

Arguing as in the previous steps we can prove the uniqueness of the solution.

Note 1. The theorem is also valid in a more general case . If A be a non-singular $n \times n$ matrix and B be an $n \times p$ matrix, there exists a unique solution of the equation $AX = B$ and the solution is the $n \times p$ matrix $X = A^{-1}B$.

If A be a non-singular $n \times n$ matrix and B be an $m \times n$ matrix, there exists a unique solution of the equation $YA = B$ and the solution is the $m \times n$ matrix $Y = BA^{-1}$.

Note 2. If A be a singular $n \times n$ matrix and B be an $n \times p$ matrix then there may or may not exist a solution of the matrix equation $AX = B$. The solution, when it exists, is however, not unique. For example,

(i) the equation $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} X = \begin{pmatrix} 5 \\ 10 \end{pmatrix}$ has solutions

$X = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $X = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $X = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$ and many others; and

(ii) the equation $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} X = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$ has no solution.

Theorem 3.2.7. Let A and B be square matrices of the same order.

If A be non-singular, $AB = O$ implies $B = O$.

If B be non-singular, $AB = O$ implies $A = O$.

Proof. Let A be non-singular. Then A^{-1} exists.

$$\begin{aligned} AB = O &\Rightarrow A^{-1}(AB) = A^{-1}O \quad \text{[pre multiplication by } A^{-1}] \\ &\Rightarrow (A^{-1}A)A = O \\ &\Rightarrow B = O. \end{aligned}$$

Let B be non-singular. Then B^{-1} exists.

$$\begin{aligned} AB = O &\Rightarrow (AB)B^{-1} = OB^{-1} \quad \text{[post multiplication by } B^{-1}] \\ &\Rightarrow A(BB^{-1}) = O \\ &\Rightarrow A = O. \end{aligned}$$

Corollary. Let A and B square matrices of the same order. Then $AB = O$ implies either $A = O$, or $B = O$, or both A and B are singular matrices.

3.2.8. Solution of a system of linear equations by matrix method.

Let us consider the system of n linear equations is n unknowns

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \quad & \quad \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n, \end{aligned}$$

where $a_{ij} \in \mathbb{R}$, $b_i \in \mathbb{R}$ and det $(a_{ij}) \neq 0$.

Let $A = (a_{ij})$, $X = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$.

Then the system of equations can be expressed as $AX = B$.

Since det $A \neq 0, A^{-1}$ exists. Therefore the system has the unique solution $X = A^{-1}B$.

Note. x_i = the ith row of $A^{-1}B$

$$= \frac{1}{\det A}[A_{1i}b_1 + A_{2i}b_2 + \dots + A_{ni}b_n]$$

$$= \frac{1}{\det A}\begin{vmatrix} a_{11} & a_{12} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & b_2 & \dots & a_{2n} \\ & \dots & \dots & & \dots & \\ a_{n1} & a_{n2} & \dots & b_n & \dots & a_{nn} \end{vmatrix}$$

$= \frac{\det A_i}{\det A}$, where A_i is the matrix obtained from A by replacing its ith column by B.

This is Cramer's rule.

Worked Example (continued).

2. Solve by matrix method the system of equations

$$\begin{aligned} x + z &= 0 \\ 3x + 4y + 5z &= 2 \\ 2x + 3y + 4z &= 1. \end{aligned}$$

Let $A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$.

Then $AX = B$.

From worked Ex.1, det $A = 2 \neq 0$. So A^{-1} exists.

Hence the unique solution of system is given by $X = A^{-1}B$.

Using worked Example 1,

$$X = \begin{pmatrix} \frac{1}{2} & \frac{3}{2} & -2 \\ -1 & 1 & -1 \\ \frac{1}{2} & -\frac{3}{2} & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

The solution of the system is $x = 1, y = 1, z = -1$.

3.3. Orthogonal Matrices.

Definition. A real $n \times n$ matrix A is said to be *orthogonal* if $AA^t = I_n$.

Theorem 3.3.1. If A be an orthogonal matrix then A is non-singular and det $A \pm 1$.

Proof. Let A be an orthogonal matrix of order n. Then $AA^t = I_n$.

det $(AA^t) =$ det $I_n = 1$

or, det A. det $A^t = 1$

or, (det $A)^2 = 1$.

This implies det $A = \pm 1$ and since det $A \neq 0$, A is non-singular.

Theorem 3.3.2. If A be an $n \times n$ orthogonal matrix then $A^tA = I_n$.

Proof. Since A is an $n \times n$ orthogonal matrix, $AA^t = I_n$.

$$\begin{aligned} AA^t = I_n \quad &\Rightarrow \quad (AA^t)A = I_nA \\ &\Rightarrow \quad A(A^tA) = I_nA = AI_n \\ &\Rightarrow \quad A(A^tA - I_n) = O. \end{aligned}$$

Since A is orthogonal, A is non-singular.

So $A(A^tA - I_n) = O \;\Rightarrow\; A^tA - I_n = O$, by Theorem 3.2.7

$\Rightarrow\; A^tA = I_n$.

Theorem 3.3.3. If A be an orthogonal matrix then $A^{-1} = A^t$.

Proof. Since A is orthogonal, $AA^t = I_n$, and this implies $A^tA = I_n$.

So $AA^t = A^tA = I_n$.

From the definition of an inverse it follows that $A^{-1} = A^t$.

Theorem 3.3.4. If A and B be orthogonal matrices of the same order then AB is orthogonal.

Proof. Let A, B be orthogonal matrices of order n. Then $AA^t = I_n$ and $BB^t = I_n$.

We have $(AB)(AB)^t = (AB)(B^tA^t) = A(BB^t)A^t$

$= (AI_n)A^t = AA^t = I_n$.

This proves that AB is orthogonal.

Theorem 3.3.5. If A be an orthogonal matrix, then A^{-1} is orthogonal.

Proof. Since A is orthogonal, A is non-singular and therefore A^{-1} exists.

We have
$$\begin{aligned} A^{-1}(A^{-1})^t &= A^{-1}(A^t)^{-1} \\ &= (A^tA)^{-1} \\ &= (I_n)^{-1} = I_n, \text{ since } A^tA = I_n. \end{aligned}$$

This proves that A^{-1} is orthogonal.

Note. From the definition of an orthogonal marix it follows that the identity matrix I_n is orthogonal.

The set of all real orthogonal matrices of order n forms a group with respect to matrix multiplication. This group is denoted by $O(n, \mathbb{R})$.

Worked Example.

1. Find a real orthogonal matrix of order 3 having the elements $\frac{2}{3}, \frac{1}{3}, \frac{2}{3}$ as the elements of a row.

Let $A = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ a & b & c \\ p & q & r \end{pmatrix}$ be an orthogonal matrix.

Then $AA^t = I_3$ gives
$$\begin{aligned} 2a + b + 2c &= 0 \quad \cdots \quad \text{(i)} \\ 2p + q + 2r &= 0 \quad \cdots \quad \text{(ii)} \\ ap + bq + cr &= 0 \quad \cdots \quad \text{(iii)} \\ a^2 + b^2 + c^2 &= 1 \quad \cdots \quad \text{(iv)} \\ p^2 + q^2 + r^2 &= 1 \quad \cdots \quad \text{(v)} \end{aligned}$$

There are many solutions for a, b, c satisfying (i) and (iv).

Taking $b = 0$, we have from (i) $c = -a$.

Using (iv) we have $a^2 = \frac{1}{2}$.

Taking $a = \frac{1}{\sqrt{2}}$, we have $b = 0, c = -\frac{1}{\sqrt{2}} \quad \cdots \quad$ (vi)

From (ii), (iii) and (vi) $\quad 2p + q + 2r = 0$ and $p - r = 0$. Therefore $r = p, q = -4p$.

Using (v), $p^2 = \frac{1}{18}$. Taking $p = \frac{1}{3\sqrt{2}}$, we have $q = -\frac{4}{3\sqrt{2}}, r = \frac{1}{3\sqrt{2}}$.

Therefore $A = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{3\sqrt{2}} & -\frac{4}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} \end{pmatrix}$.

Note. A can be determined in many other ways and therefore A is not unique.

3.4. Complex Matrices.

A matrix A whose elements are taken from the field $\mathbb{C}$ is said to be a *complex matrix*. A complex matrix A can be expressed as $P + iQ$ where P and Q are real matrices.

The matrix $\bar{A} = P - iQ$ is said to be the *conjugate* of A. The elements of $\bar{A}$ are the conjugates of the corresponding elements of A.

Properties.

1. $\bar{\bar{A}} = A$. Quite obvious.

2. $\overline{AB} = \bar{A}\bar{B}$ where A and B are complex matrices and AB is defined.

Proof. Let $A = P + iQ,\ B = R + iS$, where P, Q, R, S are real matrices.

Then $AB = (P + iQ)(R + iS) = (PR - QS) + i(QR + PS)$.

$\overline{AB} = (PR - QS) - i(QR + PS)$ and

$\bar{A}\bar{B} = (P - iQ)(R - iS) = (PR - QS) - i(QR + PS)$.

Therefore $\overline{AB} = \bar{A}\bar{B}$.

3. $(\bar{A})^t = \overline{(A^t)}$

Proof. Let $A = P + iQ$, where P, Q are real matrices.

Then $\bar{A} = P - iQ,\ A^t = P^t + iQ^t$.

Therefore $(\bar{A})^t = P^t - iQ^t = \overline{(A^t)}$.

Note. $(\bar{A})^t$ is said to be the *conjugate transpose* of A and is denoted by A°.

Theorem 3.4.1. Let A and B be matrices over the field $\mathbb{C}$. Then

(i) $(A^\circ)^\circ = A$;

(ii) $(cA)^\circ = \bar{c}A^\circ,\ c \in \mathbb{C}$;

(iii) $(A + B)^\circ = A^\circ + B^\circ$, provided $A + B$ is defined;

(iv) $(AB)^\circ = B^\circ A^\circ$, provided AB is defined.

Proof left as exercise.

Theorem 3.4.2. If A be a non-singular matrix over $\mathbb{C}$, then $(A^\circ)^{-1} = (A^{-1})^\circ$.

Proof. Let A be a non-singular matrix of order n.

Then A^{-1} exists and $A.A^{-1} = A^{-1}.A = I_n$.

$(I_n)^\circ = (A.A^{-1})^\circ = (A^{-1}.A)^\circ$.

or, $I_n = (A^{-1})^\circ.A^\circ = A^\circ.(A^{-1})^\circ$.

From the definition of an inverse and its uniqueness it follows that $(A^{-1})^\circ$ is the inverse of A°. That is, $(A^\circ)^{-1} = (A^{-1})^\circ$.

Hermitian and Skew Hermitian matrices.

A complex $n \times n$ matrix A is said to be *Hermitian* if $A^o = A$, and *skew Hermitian* if $A^o = -A$, where $A^o = \bar{A}^t$.

Therefore a complex $n \times n$ matrix $A = (a_{ij})$ is Hermitian if $\bar{a_{ij}} = a_{ji}$ and skew Hermitian if $\bar{a_{ij}} = -a_{ji}$ for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$.

Examples of a Hermitian matrix are

$$\begin{pmatrix} 1 & 2+i & -1 \\ 2-i & 2 & i \\ -1 & -i & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3i \\ 2 & 0 & 4 \\ -3i & 4 & 2 \end{pmatrix}.$$

Examples of a skew Hermitian matrix are

$$\begin{pmatrix} 0 & 2 & i \\ -2 & 0 & 1+i \\ i & -1+i & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 2i & 1+2i \\ 2i & i & 1 \\ -1+2i & -1 & 0 \end{pmatrix}.$$

In particular, a real Hermitian matrix is a real symmetric matrix and a real skew Hermitian matrix is a real skew symmetric matrix.

Theorem 3.4.3. If $H = P + iQ$ be a Hermitian matrix, then

(i) the diagonal elements of H are all real numbers,

(ii) P is a real symmetric matrix and Q is a real skew symmetric matrix.

Proof. (i) Let $H = (h_{ij})$. Then $\bar{h}_{ij} = h_{ji}$.

So $\bar{h}_{ii} = h_{ii}$ and this proves that h_{ii} is purely real.

(ii) Since H is Hermitian, $\bar{H}^t = H$.

Therefore $P^t - iQ^t = P + iQ$. This implies $P^t = P$ and $Q^t = -Q$, proving that P is symmetric and Q is skew symmetric.

Theorem 3.4.4. If $S = M + iN$ be a skew Hermitian matrix, then

(i) the diagonal elements of S are purely imaginary numbers or zero,

(ii) M is a real skew symmetric matrix and N is a real symmetric matrix.

Proof. (i) Let $S = (s_{ij})$. Then $\bar{s_{ij}} = -s_{ji}$.

So $\bar{s}_{ii} = -s_{ii}$ and this proves that s_{ii} is purely imaginary or zero.

(ii) Since S is skew Hermitian, $\bar{S}^t = -S$.

Therefore $M^t - iN^t = -(M + iN)$. This implies $M^t = -M$ and $N^t = N$ proving that M is skew symmetric and N is symmetric.

Theorem 3.4.5. If A be a complex square matrix, then $A + A^o$ is Hermitian and $A - A^o$ is skew Hermitian.

Proof is immediate.

Theorem 3.4.6. A complex square matrix can be uniquely expressed as the sum of a Hermitian matrix and a skew Hermitian matrix.

Proof. Let A be a given matrix. Then A can be expressed as

$A = \frac{1}{2}(A + A^o) + \frac{1}{2}(A - A^o)$.

$[\frac{1}{2}(A + A^o)]^o = \frac{1}{2}[A^o + (A^o)^o] = \frac{1}{2}(A + A^o)$, and

$[\frac{1}{2}(A - A^o)]^o = \frac{1}{2}[A^o - (A^o)^o] = -\frac{1}{2}(A - A^o)$.

This shows that $\frac{1}{2}(A + A^o)$ is Hermitian and $\frac{1}{2}(A - A^o)$ is skew Hermitian. Therefore A is expressed as the sum of a Hermitian and a skew Hermitian matrix.

We now show that this decomposition is unique.

Let $A = P + Q$ where P is Hermitian and Q is skew Hermitian.

Then $A^o = P^o + Q^o = P - Q$.

We have $A + A^o = 2P$, $A - A^o = 2Q$.

So $P = \frac{1}{2}(A + A^o)$, $Q = \frac{1}{2}(A - A^o)$ and this proves the theorem.

Unitary matrices.

A complex $n \times n$ matrix A is said to be *unitary* if $AA^o = I_n$.

Examples of a unitary matrix are

$$\begin{pmatrix} \frac{1+i}{2} & \frac{1-i}{2} \\ \frac{1-i}{2} & \frac{1+i}{2} \end{pmatrix}, \quad \frac{1}{\sqrt{3}}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix} \text{ where } \omega = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}.$$

Note. For a real unitary matrix A, $AA^o = AA^t = I_n$. Hence every real unitary matrix is a real orthogonal matrix.

Theorem 3.4.7. If A be a unitary matrix then A is non-singular and $|\det A| = 1$.

Proof. Let A be an $n \times n$ unitary matrix. Then $A\bar{A}^t = I_n$.

$\det(A\bar{A}^t) = \det I_n = 1$

or, $\det A \det \bar{A} = 1$, since $\det \bar{A}^t = \det \bar{A}$

or, $\det A . \overline{\det A} = 1$. That is, $|\det A|^2 = 1$.

This implies $|\det A| = 1$ and A is necessarily non-singular.

Theorem 3.4.8. If A be an $n \times n$ unitary matrix then $A^o A = I_n$.

Proof. Since A is an $n \times n$ unitary matrix, $AA^o = I_n$.

$$AA^\circ = I_n \Rightarrow (AA^\circ)A = I_nA \quad \Rightarrow \quad A(A^\circ A) = I_nA$$
$$\Rightarrow \quad A(A^\circ A) = AI_n$$
$$\Rightarrow \quad A(A^\circ A - I_n) = O.$$

Since A is unitary, A is non-singular. Therefore A^{-1} exists.
So $A(A^\circ A - I_n) = O \Rightarrow A^{-1}[A(A^\circ A - I_n)] = O \Rightarrow A^\circ A = I_n$.

Theorem 3.4.9. If A be a unitary matrix then $A^{-1} = A^\circ$.

Proof. Since A is unitary, $AA^\circ = I_n$, and this implies $A^\circ A = I_n$. So $AA^\circ = A^\circ A = I_n$. From the definition of an inverse it follows that $A^{-1} = A^\circ$.

The product of two unitary matrices of the same order is an unitary matrix and the inverse of an unitary matrix is unitary.

The theory of complex unitary matrices is closely similar to that of real orthogonal matrices.

Note. The set of all $n \times n$ unitary matrices forms a group with respect to matrix multiplication. This group is denoted by $U(n, \mathbb{C})$.

Worked Example.

1. A is a complex $n \times n$ matrix. Prove that $A\bar{A}^t = O \Rightarrow A = O$ and $\bar{A}^tA = O \Rightarrow A = O$.

Let $A\bar{A}^t = O$ and let $A = (a_{ij})_{n,n}$. Let $A\bar{A}^t = B = (b_{ij})_{n,n}$. Then

$$b_{11} = a_{11}\bar{a}_{11} + a_{12}\bar{a}_{12} + \cdots + a_{1n}\bar{a}_{1n}$$
$$b_{22} = a_{21}\bar{a}_{21} + a_{22}\bar{a}_{22} + \cdots + a_{2n}\bar{a}_{2n}$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots$$
$$b_{nn} = a_{n1}\bar{a}_{n1} + a_{n2}\bar{a}_{n2} + \cdots + a_{nn}\bar{a}_{nn}.$$

$A\bar{A}^t = O \Rightarrow B = O \Rightarrow b_{11} = 0, b_{22} = 0, \ldots, b_{nn} = 0.$

$$b_{11} = 0 \Rightarrow |a_{11}|^2 + |a_{12}|^2 + \cdots + |a_{1n}|^2 = 0$$
$$\Rightarrow a_{11} = a_{12} = \cdots = a_{1n} = 0.$$
$$b_{22} = 0 \Rightarrow |a_{21}|^2 + |a_{22}|^2 + \cdots + |a_{2n}|^2 = 0$$
$$\Rightarrow a_{21} = a_{22} = \cdots = a_{2n} = 0.$$
$$\ldots \quad \ldots \quad \ldots \quad \ldots$$
$$b_{nn} = 0 \Rightarrow |a_{n1}|^2 + |a_{n2}|^2 + \cdots + |a_{nn}|^2 = 0$$
$$\Rightarrow a_{n1} = a_{n2} = \cdots = a_{nn} = 0.$$

Therefore $A = O$. Similar proof for the second part.

Note. For a complex $n \times 1$ matrix X, $\bar{X}^t.X$ is a scalar ≥ 0, and $\bar{X}^t X = 0 \Rightarrow X = O$.

Exercises 3

1. If A be a non-singular matrix, prove that *adj* A is non-singular and $(adj\ A)^{-1} = adj\ (A^{-1}) = (\det A)^{-1}A$.

2. (i) If A be a non-singular matrix, prove that $\det A^{-1} = (\det A)^{-1}$.
(ii) If A be a real orthogonal matrix, prove that $\det A^{-1} = \det A$.
(iii) If $A = (a_{ij})$ be a real orthogonal matrix with $\det A = 1$, prove that each element a_{rs} of A is equal to its cofactor A_{rs} in $\det A$.

3. A square matrix A is said to be a *nilpotent* matrix of index p, if p is the least positive integer such that $A^p = O$.

If A is an $n \times n$ nilpotent matrix of index 2 $(n \geq 2)$, prove that both $I_n - A$ and $I_n + A$ are non-singular matrices.

4. Compute the adjoint and the inverse of the matrix $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

5. Find the matrix A, if $adj\ A = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ and $\det A = 2$.

6. Find the matrix A, if $adj\ A = \begin{pmatrix} 1 & 3 & -4 \\ -2 & 2 & -2 \\ 1 & -3 & 4 \end{pmatrix}$.

7. Find the matrix A, if $A^{-1} = \begin{pmatrix} 3 & -1 & 1 \\ 1 & -2 & 3 \\ 3 & -3 & 4 \end{pmatrix}$.

8. If $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$, show that $A^{-1} = A^3$.

9. If $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$, show that $A^2 - 10A + 16I_3 = O$. Hence obtain A^{-1}.

10. If $A = \begin{pmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{pmatrix}$, compute AA^t.

If a, b, c, d be all real, deduce that A is invertible if least one of a, b, c, d be non-zero. Find A^{-1} in that case.

11. A and B are real $n \times n$ matrices such that A, B and $A + B$ are invertible. Prove that $\quad A(A+B)^{-1}B = B(A+B)^{-1}A$.

12. If $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -4 & 1 \\ 3 & 0 & -3 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 1 & -1 \end{pmatrix}$, show that $AB = 6I_3$.

Utilise this result to solve

(i) $\begin{aligned} 2x + y + z &= 5 \\ x - y &= 0 \\ 2x + y - z &= 1, \end{aligned}$ (ii) $\begin{aligned} x + y + 3z &= 6 \\ 2x - 4y &= 0 \\ x + y - 3z &= 0. \end{aligned}$

13. If $A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{pmatrix}$, show that A^{-1} exists.

Without computing A^{-1} find the sum of all the elements of A^{-1}.

14. Find all real x so that the matrix

$$\begin{pmatrix} 2+x & 2 & 2 & 2 \\ 2 & 2+x & 2 & 2 \\ 2 & 2 & 2+x & 2 \\ 2 & 2 & 2 & 2+x \end{pmatrix}$$

is invertible. Assuming that A^{-1} exists, find the sum of all the elements of A^{-1} without computing A^{-1}.

15. (i) A is a non-singular matrix such that the sum of the elements in each row is k. Prove that the sum of the elements in each row of A^{-1} is k^{-1}.

(ii) A is a non-singular matrix such that the sum of the elements in each column is k. Prove that the sum of the elements in each column of A^{-1} is k^{-1}.

16. Prove that the matrix $\frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{pmatrix}$ is orthogonal.

Utilise this to solve the equations

$$\begin{aligned} x - 2y + 2z &= 2 \\ 2x - y - 2z &= 1 \\ 2x + 2y + z &= 7. \end{aligned}$$

17. If (a_{ij}) be an orthogonal matrix of order 3 and

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= 1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= 0 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= 0, \end{aligned}$$

show that $x_1 = a_{11}$, $x_2 = a_{12}$, $x_3 = a_{13}$.

18. If a_i, b_i, $c_i (i = 1, 2, 3)$ be all real and $a_1^2 + b_1^2 + c_1^2 = a_2^2 + b_2^2 + c_2^2 = a_3^2 + b_3^2 + c_3^2 = 1$; $a_1a_2 + b_1b_2 + c_1c_2 = a_2a_3 + b_2b_3 + c_2c_3 = a_3a_1 + b_3b_1 + c_3c_1 = 0$, prove that

$a_1^2 + a_2^2 + a_3^2 = b_1^2 + b_2^2 + b_3^2 = c_1^2 + c_2^2 + c_3^2 = 1$ and
$a_1b_1 + a_2b_2 + a_3b_3 = b_1c_1 + b_2c_2 + b_3c_3 = c_1a_1 + c_2a_2 + c_3a_3 = 0$.

19. A and B are real orthogonal matrices of the same order and det $A +$ det $B = 0$. Show that $A + B$ is a singular matrix.

[Hint. Consider the product $A^t(A + B)B^t$.]

20. P is an $n \times n$ real orthogonal matrix with det $P = -1$. Prove that $P + I_n$ is a singular matrix.

21. A is a real skew symmetric matrix and $I + A$ is non-singular, I being the identity matrix. Prove that the matrix $(I + A)^{-1}(I - A)$ is orthogonal.

22. A is a real square matrix and the matrix $(I + A)^{-1}(I - A)$ is orthogonal, I being the identity matrix. Prove that A is skew symmetric.

23. A is a real orthogonal matrix and $I + A$ is non-singular, I being the identity matrix. Prove that the matrix $(I + A)^{-1}(I - A)$ is skew symmetric.

24. A is a real square matrix and the matrix $(I + A)^{-1}(I - A)$ is skew symmetric, I being the identity matrix. Prove that A is orthogonal.

25. Let A be a square matrix such that $I + A$ is non-singular, I being the identity matrix. Let $\bar{A} = (I + A)^{-1}(I - A)$. Prove that

(i) $I + \bar{A}$ is non-singular, (ii) $\bar{\bar{A}} = A$.

26. A and B are non-singular matrices such that $AA^t = BB^t$. Prove that there exist orthogonal matrices P, Q such that $A = BP$ and $B = AQ$.

27. A is an $n \times n$ matrix and there exists a unique matrix B such that $AB = I_n$. Prove that $BA = I_n$ too and $B = A^{-1}$.

[Hint. Consider the product $A(BA - I_n + B)$.]

28. Find a real orthogonal matrix of order 3 having the elements

(i) $\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$ as the elements of a row;

(ii) $\frac{1}{3}, \frac{-2}{3}, \frac{2}{3}$ as the elements of a column.

29. Find a real orthogonal matrix of order 3, other than $\pm I_3$, having all integer elements.

30. Express the matrix $A = \begin{pmatrix} 1 & 2+i & 1-i \\ 2-i & 1+2i & 3 \\ 2+i & 2 & 1+i \end{pmatrix}$ as the sum of a Hermitian matrix and a skew Hermitian matrix.

31. (a) If H be a Hermitian matrix, prove that det H is a real number.

(b) If S be a skew Hermitian matrix of order n, prove that

(i) if n be even, then det S is a real number;

(ii) if n be odd, then det S is a purely imaginary number or zero.

32. A is a Hermitian matrix and $A^2 = O$. Prove that $A = O$.

33. A is a skew Hermitian matrix and $I+A$ is non-singular, I being the identity matrix. Prove that the matrix $(I+A)^{-1}(I-A)$ is unitary.

34. A is a unitary matrix and $I+A$ is non-singular, I being the identity matrix. Prove that the matrix $(I+A)^{-1}(I-A)$ is skew Hermitian.

35. Show that every complex square matrix A can be expressed as $P+iQ$, where P and Q are both Hermitian matrices.

36. If $a, b, c, d \in \mathbb{R}$ and $a^2+b^2+c^2+d^2=1$, prove that the matrix $\begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$ is unitary.

3.5. Rank of a matrix.

Let A be a non-zero matrix of order $m \times n$. The *rank* of A is defined to be the greatest positive integer r such that A has at least one non-zero minor of order r. The rank of a zero matrix A is defined to be 0.

The rank of A is also called the *determinant rank* of A.

If the rank of A be r, every minor of order $r+1$ is zero. Therefore every minor of order $r+2$ is also zero because each such can be expressed in terms of minors of order $r+1$. By similar arguments it can be proved that every minor of order greater than r is zero.

For a non-zero $m \times n$ matrix A, $0 < \text{rank of } A \leq \min\{m, n\}$.

For a square matrix A of order n, rank of $A < n$, or $= n$ according as A is singular or non-singular.

Rank of A = rank of A^t, because A and A^t have identical minors.

Examples.

1. Let $A = \begin{pmatrix} 1 & 0 & 3 \\ 4 & -1 & 5 \\ 2 & 0 & 6 \end{pmatrix}$.

Since det $A = 0$, rank of $A < 3$ and since there is a second order minor $\begin{vmatrix} 1 & 0 \\ 4 & -1 \end{vmatrix} \neq 0$, rank of $A = 2$.

2. Let $A = \begin{pmatrix} 2 & 3 & -1 & 1 \\ 3 & 0 & 4 & 2 \\ 6 & 9 & -3 & 3 \end{pmatrix}$.

Every minor of order 3 is zero, because the third row is a multiple of the first and therefore rank of $A < 3$. There is a second order minor $\begin{vmatrix} 2 & 3 \\ 3 & 0 \end{vmatrix} \neq 0$ and so rank of $A = 2$.

The rank of a matrix depends on the nature of its minors. The process of determining the rank of a matrix starts from the evaluation of the minor or minors of highest possible order and continues until a non-zero minor is obtained. For instance, let us consider the matrix

$$\begin{pmatrix} 2 & 3 & 0 & 4 & -1 \\ 1 & 2 & 5 & -7 & 3 \\ 4 & 6 & 0 & 8 & -2 \\ 3 & 5 & 5 & -3 & 2 \end{pmatrix}.$$

There are five minors of order 4 and each such is zero. Next to examine are minors of order 3. There are thirty such and each of them is zero. Next to examine are minors of order 2. There exists a non-zero minor $\begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix}$ and therefore the rank of the matrix is 2.

However this method of determining the rank of a matrix is much laborious and therefore we need some easier method for this purpose.

With this end in view, we now introduce some transformations on the matrix, called the *elementary operations* on the matrix.

3.6. Elementary operations.

An *elementary operation* on a matrix A over a field F is an operation of the following three types.

1. Interchange of two rows (or columns) of A.
2. Multiplication of a row (or column) by a non-zero scalar c in F.
3. Addition of a scalar multiple of one row (or column) to another row (or column).

When applied to rows, the elementary operations are said to be *elementary row operations*. And when applied to columns they are said to be *elementary column operations*.

The interchange of the ith and jth row is denoted by R_{ij}.

Multiplication of the ith row by a non-zero scalar c is denoted by cR_i [or $R_i(c)$].

Addition of c times the jth row to the ith row is denoted by $R_i + cR_j$ [or $R_{ij}(c)$].

In a similar manner, the *elementary column operations* C_{ij}, cC_i [or $C_i(c)$], $C_i + cC_j$ [or $C_{ij}(c)$] are defined.

If T be an elementary operation on the matrix A the transformed matrix is denoted as $T(A)$. If $B = T(A)$ the operation is expressed as

$$A \xrightarrow{T} B.$$

For example,

$$\begin{pmatrix} 2 & 4 & 0 \\ 4 & 9 & 5 \\ 1 & 3 & 7 \end{pmatrix} \xrightarrow{R_{23}} \begin{pmatrix} 2 & 4 & 0 \\ 1 & 3 & 7 \\ 4 & 9 & 5 \end{pmatrix}, \begin{pmatrix} 2 & 4 & 0 \\ 4 & 9 & 5 \\ 1 & 3 & 7 \end{pmatrix} \xrightarrow{C_{23}} \begin{pmatrix} 2 & 0 & 4 \\ 4 & 5 & 9 \\ 1 & 7 & 3 \end{pmatrix};$$

$$\begin{pmatrix} 2 & 1 & 3 \\ 4 & 5 & 0 \\ 3 & 2 & 1 \end{pmatrix} \xrightarrow{2R_3} \begin{pmatrix} 2 & 1 & 3 \\ 4 & 5 & 0 \\ 6 & 4 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 3 \\ 4 & 5 & 0 \\ 3 & 2 & 1 \end{pmatrix} \xrightarrow{2C_3} \begin{pmatrix} 2 & 1 & 6 \\ 4 & 5 & 0 \\ 3 & 2 & 2 \end{pmatrix};$$

$$\begin{pmatrix} 2 & 4 \\ 4 & 9 \\ 1 & 3 \end{pmatrix} \xrightarrow{R_2-2R_1} \begin{pmatrix} 2 & 4 \\ 0 & 1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 4 \\ 4 & 9 \\ 1 & 3 \end{pmatrix} \xrightarrow{C_2-2C_1} \begin{pmatrix} 2 & 0 \\ 4 & 1 \\ 1 & 1 \end{pmatrix}.$$

If T be an elementary operation on A such that $T(A) = B$ and T_1 be an elementary operation on B such that $T_1(B) = C$ then $C = T_1\{T(A)\}$. C is obtained from A by applying two elementary operations T and T_1 successively.

If T be an elementary row (column) operation on a matrix A then T^{-1}, the inverse of T, is defined to be an elementary row (column) operation such that $T^{-1}(TA) = A$.

For example, if $T = R_{ij}$ then $T^{-1} = R_{ij}$;

if $T = R_i(c)$ then $T^{-1} = R_i(c^{-1})$;

if $T = R_{ij}(c)$ then $T^{-1} = R_{ij}(-c)$.

Clearly, the inverse of an elementary row (column) operation is an elementary row (column) operation of the same type.

Row equivalence. Column equivalence.

Let us consider the set S of all $m \times n$ matrices over a field F. A matrix B in S is said to be *row equivalent* (*column equivalent*) to a matrix A in S if B can be obtained by successive application of a finite number of elementary row operations (column operations) on A.

The relation of row equivalence (column equivalence) on the set S is an equivalence relation. Consequently, the set S is partitioned into classes of row equivalent (column equivalent) matrices.

We shall now discuss some properties of *row equivalent* matrices. Analogous properties hold in case of *column equivalent* matrices.

Row-reduced matrix. Row echelon matrix.

Definition. An $m \times n$ matrix A is said to be *row-reduced* if

(a) the first non-zero element in each non-zero row is 1 (called the leading 1); and

(b) in each column containing the leading 1 of some row, the leading 1 is the only non-zero element.

Examples of a row-reduced matrix are

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Definition. An $m \times n$ matrix A is said to be a *row-reduced echelon* matrix (or a *row echelon* matrix) if

(a) A is row-reduced;

(b) there is an integer $r(0 \le r \le m)$ such that the first r rows of A are non-zero rows and the remaining rows (if there be any) are all zero rows; and

(c) if the leading element of the ith non-zero row occurs in the k_ith column of A, then $k_1 < k_2 < \cdots < k_r$.

Examples of a row echelon matrix are

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Theorem 3.6.1. A matrix A can be made row equivalent to a row-reduced matrix B by elementary row operations of the types 2 and 3.

The proof is given by describing the actual method of reduction.

Let $A = (a_{ij})_{mn}$. Let the t_1th row be the first non-zero row with the first non-zero element $a_{t_1k_1}$ in its k_1th column. Multiply the row by $(a_{t_1k_1})^{-1}$. The leading element of the first non-zero row becomes 1. Subtract a_{rk_1} times the new t_1th row from the rth row for every $r \ne t_1$. This reduces every other element in the k_1th column to zero. Thus the two conditions for a row-reduced matrix are satisfied for the first t_1 rows.

Then consider the next non-zero row. Let the t_2th row be the next non-zero row with the first non-zero element $a_{t_2k_2}$ in its k_2th column. Multiply the row by $(a_{t_2k_2})^{-1}$. The leading element of the second non-zero row becomes 1. Subtract a_{rk_2} times the new t_2th row from the rth row for every $r \neq t_2$. This reduces every other element in the k_2th column to zero. This operation will affect neither the elements of the previous rows in columns $1, 2, ..., k_1$, nor the elements of column k_1. Thus the two conditions for a row-reduced matrix are satisfied for the first t_2 rows.

Working with the succeeding non-zero rows one at a time, after a finite number of steps the resulting matrix B is obtained in the row-reduced form.

Since B is obtained from A by a finite number of elementary row operations, B is row equivalent to A.

Worked Example.

1. Find a row-reduced matrix which is row equivalent to

$$\begin{pmatrix} 0 & 0 & 2 & 2 & 0 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 6 & 2 & 6 & 2 \end{pmatrix}.$$

Let us apply elementary row operations on the matrix.

$$\begin{pmatrix} 0 & 0 & 2 & 2 & 0 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 6 & 2 & 6 & 2 \end{pmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 6 & 2 & 6 & 2 \end{pmatrix}$$

$$\underset{R_3-2R_1}{\overset{R_2-2R_1}{\longrightarrow}} \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 3 & 0 & 2 & 1 \\ 2 & 6 & 0 & 4 & 2 \end{pmatrix} \xrightarrow{R_3-2R_2} \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 3 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Working procedure. The first row is a non-zero row and the first non-zero element is 2 (in the third column).

Step 1. Multiply the first row by $\frac{1}{2}$. The leading element in the row becomes 1.

Step 2. To reduce the other elements in the third column to zero, perform the operations $R_2 - 2R_1, R_3 - 2R_1$.

Observe that the first non-zero element in the 2nd row is 1 (in the first column).

Step 3. To reduce the other elements in the first column to zero, perform the operation $R_3 - 2R_2$.

Observe that the 3rd row becomes a zero row.

All the rows are exhausted and the process terminates.

Theorem 3.6.2. A matrix A can be made row equivalent to a row echelon matrix B by elementary row operations.

Proof. Let $A = (a_{ij})_{mn}$. A can be transformed to a row-reduced matrix M by a finite succession of elementary row operations of the type 2 and 3.

Now apply a finite number of elementary row operations of the type 1 (i.e., interchange of rows) to bring all zero rows below every non-zero row and then rearrange the non-zero rows in such a manner that the leading 1's in the non-zero rows $1, 2, \ldots, r$ occur in columns $k_1, k_2, \ldots, k_r$ respectively such that $k_1 < k_2 < \cdots < k_r$.

The resulting matrix B becomes a row echelon matrix. Since B is obtained from A by a finite number of elementary row operations, B is row equivalent to A.

This completes the proof.

Worked Examples (continued).

2. Apply elementay row operations to reduce the following matrix to a row echelon matrix

$$\begin{pmatrix} 2 & 0 & 4 & 2 \\ 3 & 2 & 6 & 5 \\ 5 & 2 & 10 & 7 \\ 0 & 3 & 2 & 5 \end{pmatrix}.$$

$$\begin{pmatrix} 2 & 0 & 4 & 2 \\ 3 & 2 & 6 & 5 \\ 5 & 2 & 10 & 7 \\ 0 & 3 & 2 & 5 \end{pmatrix} \xrightarrow{\frac{1}{2}R_1} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 3 & 2 & 6 & 5 \\ 5 & 2 & 10 & 7 \\ 0 & 3 & 2 & 5 \end{pmatrix}$$

$$\underset{R_3-5R_1}{\overset{R_2-3R_1}{\longrightarrow}} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 2 & 0 & 2 \\ 0 & 3 & 2 & 5 \end{pmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 3 & 2 & 5 \end{pmatrix}$$

$$\underset{R_4-3R_2}{\overset{R_3-2R_2}{\longrightarrow}} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix} \xrightarrow{R_{34}} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{2}R_3} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1-2R_3} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = R, \text{ say.}$$

R is a row echelon matrix.

Working procedure. The element in the $(1,1)$ position is 2.

Step 1. Multiply the 1st row by $\frac{1}{2}$. The leading element in the first row becomes 1 in the $(1,1)$ position.

Step 2. To reduce all other elements in the first column to zero, perform the operations $R_2 - 3R_1, R_3 - 5R_1$.

Consider the submatrix obtained by deleting the first row and the first column. Observe that the element in the $(1, 1)$ position of the submatrix, i.e., the element in the $(2, 2)$ position of the original matrix is 2.

Step 3. Multiply the second row by $\frac{1}{2}$. The leading element in the second row becomes 1 in the $(2, 2)$ position .

Step 4. To reduce all other elements in the second column to zero, perform the operations $R_3 - 2R_2, R_4 - 3R_2$.

Observe that the third row becomes a zero row.

Step 5. Perform R_{34} to bring the zero row to the last.

Consider the submatrix obtained by deleting first two rows and first two columns. Observe that the element in the $(1, 1)$ position of the submatrix is 2.

Step 6. Multiply the third row by $\frac{1}{2}$. The leading element in the third row becomes 1 in the $(3, 3)$ position.

Step 7. To reduce all other elements in the third column to zero, perform the operation $R_1 - 2R_3$.

The rows are exhausted and the process terminates.

3. Find a row echelon matrix which is row equivalent to

$$\begin{pmatrix} 0 & 0 & 2 & 2 & 0 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 6 & 2 & 6 & 2 \\ 3 & 9 & 1 & 10 & 6 \end{pmatrix}.$$

Let us apply elementary row operations on the matrix.

$$\begin{pmatrix} 0 & 0 & 2 & 2 & 0 \\ 1 & 3 & 2 & 4 & 1 \\ 2 & 6 & 2 & 6 & 2 \\ 3 & 9 & 1 & 10 & 6 \end{pmatrix} \xrightarrow{R_{12}} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 2 & 6 & 2 & 6 & 2 \\ 3 & 9 & 1 & 10 & 6 \end{pmatrix}$$

$$\xrightarrow[R_4-3R_1]{R_3-2R_1} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & -2 & -2 & 0 \\ 0 & 0 & -5 & -2 & 3 \end{pmatrix} \xrightarrow{\frac{1}{2}R_2} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -2 & -2 & 0 \\ 0 & 0 & -5 & -2 & 3 \end{pmatrix}$$

$$\xrightarrow[R_3+2R_2,\ R_4+5R_2]{R_1-2R_2} \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 3 \end{pmatrix} \xrightarrow{R_{34}} \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{3}R_3} \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[R_2-R_3]{R_1-2R_3} \begin{pmatrix} 1 & 3 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Working procedure. The first column is a non-zero column. The element in the $(1,1)$ position is zero.

Step 1. Perform R_{12} to bring a non-zero element to $(1, 1)$ position. The leading 1 in the first row occurs in the first column.

Step 2. To reduce the other elements in the first column to zero, perform the operations $R_3 - 2R_1, R_4 - 3R_1$.

Consider the submatrix obtained by deleting the first row and the first column. Observe that the first column of the submatrix is a zero column, the second column is the next non-zero column. The element in the $(2, 3)$ position is a non-zero element 2.

Step 3. Multiply the second row by $\frac{1}{2}$. The leading element in the row becomes 1.

Step 4. Reduce all other elememts in the 3rd column to 0 by performing the operations $R_1 - 2R_2, R_3 + 2R_2, R_4 + 5R_2$.

Observe that the third row becomes a zero row.

Step 5. Perform R_{34} to bring the zero row to the last.

Step 6. Multiply the 3rd row by $\frac{1}{3}$. The leading element in the row becomes 1 in the 4th column.

Step 7. Reduce all other elements in the 4th column to zero by performing the operations $R_1 - 2R_3, R_2 - R_3$.

The rows are exhausted and the process terminates.

Theorem 3.6.3. If a row echelon matrix R has r non-zero rows, then rank of $R = r$.

Proof. Let the leading 1's of the first, second, ..., rth non-zero rows of R occur in columns $k_1, k_2, \ldots, k_r$ respectively. Then $k_1 < k_2 < \cdots < k_r$. Clearly, the submatrix formed from the rows $1, 2, \ldots, r$ and the columns $k_1, k_2, \ldots, k_r$ is the identity matrix I_r and it has a non-zero determinant. So R has a non-zero minor of order r.

If the number of rows of R be r then clearly the rank of R is r.

If the number of rows of R be greater than r, then every square submatrix of order $r+1$ contains a zero row and therefore each minor of order $r+1$ is zero. Therefore rank of R is r.
This completes the proof.

Theorem 3.6.4. The rank of a matrix remains invariant under an elementary row operation.

Proof. Let A be a matrix of rank r.

Let us apply an elementary row operation of the type 1.

Let $R_{ij}(A) = B$. Let M be a square submatrix of B of order $r+1$. The rows of M are also rows of A with possibly some interchanges. Since A is of rank r, det $M = 0$. So every minor of B of order $r+1$ is zero.

Since rank of A is r, there is a non-zero minor $|N|$ of A of order r. The submatrix N may contain

(i) none of the ith and the jth row of A, or

(ii) both of the ith and the jth row, or

(iii) one of them, say the ith.

In case (i), N remains as a submatrix of B.

In case (ii), N is transformed to a submatrix of B in which the ith and the jth row of A are interchanged and so the associated minor of B has the same non-zero magnitude of $|N|$ with a change of sign.

In case (iii), there is a submatrix of B which contains all the rows and columns of N but possibly with some interchanges of rows and so the associated minor of B has the same non-zero magnitude of $|N|$ with a possible change of sign.

Thus in B every minor of order $r+1$ is zero and there is a non-zero minor of order r. Therefore rank of $B = r$.

Let us apply an elementary row operation of the type 2.

Let $R_i(c)A = B$. Let M be a square submatrix of B of order $r+1$. If M contains the ith row of B, det M is c times a minor of A of order $r+1$ and so det $M = 0$. If M does not contain the ith row of B, det M is equal to a minor of A of order $r+1$ and so det $M = 0$. So every minor of B of order $r+1$ is zero.

There is a non-zero minor $|N|$ of A of order r. If the submatrix N does not contain the ith row of A, it remains as a submatrix of B. If N contains the ith row of A, it is transformed to a submatrix of B, the determinant of which is equal to $c|N| \neq 0$.

Thus in B every minor of order $r+1$ is zero and there is a non-zero minor of order r. Therefore rank of $B = r$.

Let us apply an elementary row operation of the type 3.

Let $R_{ij}(c)A = B$. Let M be a square submatrix of B of order $r+1$. If M contains the ith row of B, det M is the sum of two minors of order $r+1$ of A and so det $M = 0$. If M does not contain the ith row of B, det M is a minor of order $r+1$ of A and so det $M = 0$.

There is a non-zero minor $|N|$ of A of order r. The submatrix N

(i) may not contain the ith row, or

(ii) may contain the ith and the jth row, or

(iii) may contain the ith row and not the jth.

In case (i), N remains as a submatrix of B.

In case (ii), the determinant of the transformed submatrix in B can be expressed as the sum of two determinants $\Delta + c\Delta'$, where $\Delta = |N|$ and $\Delta' = 0$ (containing two identical rows).

In case (iii), the determinant of the transformed submatrix in B can be expressed as the sum of two determinants $\Delta + c\Delta'$, where $\Delta = |N|$ and Δ' differs from Δ in containing the jth row in place of ith row of A.

If $\Delta + c\Delta' \neq 0$, the minor of B of order r is non-zero. If $\Delta + c\Delta' = 0$, then $c\Delta' = -\Delta \neq 0$.

Therefore there exists a non-zero minor of A of order r which does not contain the ith row of A and by case (i) the corresponding minor in B is non-zero.

Thus in B every minor of order $r+1$ is zero and there exists a non-zero minor of order r. Therefore rank of $B = r$.
The proof is complete.

Note 1. The theorem can be re-stated in the following form–

Two row equivalent matrices have the same rank.

Note 2. The rank of a matrix remains invariant under elementary column operations.

Theorem 3.6.5. If a matrix A be row equivalent to a row echelon matrix having r non-zero rows, then rank of $A = r$.

This follows from the combination of the Theorems 3.6.3 and 3.6.4.

Theorem 3.6.6. An $n \times n$ matrix A is non-singular if and only if A is row equivalent to the identity matrix I_n.

Proof. Let A be a non-singular matrix. Then rank of $A = n$. Let A be row equivalent to a row echelon matrix R. Then R must have n non-zero rows. Let the leading 1's occur in columns $k_1, k_2, \ldots, k_n$.

Then $k_1 < k_2 < \cdots < k_n$. But R contains n columns only. Therefore $k_1 = 1, k_2 = 2, \ldots, k_n = n$. So each column of R contains a leading 1. Clearly, $R = I_n$. So A is row equivalent to I_n.

Conversely, let A be row equivalent to I_n. Then rank of A = rank of I_n $= n$. Hence det $A \neq 0$ and so A is non-singular.

This completes the proof.

Worked Example (continued).

4. Determine the rank of the matrix $A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 8 & 6 \\ 3 & 6 & 6 & 3 \end{pmatrix}$.

Let us apply elementary row operations on A to reduce it to a row echelon matrix.

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 8 & 6 \\ 3 & 6 & 6 & 3 \end{pmatrix} \xrightarrow[R_3-3R_1]{R_2-2R_1} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 6 & 6 \\ 0 & 0 & 3 & 3 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{6}R_2} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{pmatrix} \xrightarrow[R_3-3R_2]{R_1-R_2} \begin{pmatrix} 1 & 2 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = R, \text{ say.}$$

R is a row echelon matrix and R has 2 non-zero rows. Therefore rank of $R = 2$. Since A is row equivalent to R, rank of $A = 2$.

Fully reduced normal form.

An $m \times n$ matrix B is said to be *equivalent* to an $m \times n$ matrix A over the same field F, if B can be obtained from A by a finite number of elementary row and column operations. Thus row equivalence and column equivalence are particular cases of equivalence of matrices.

For a given $m \times n$ matrix A, a row-reduced echelon matrix B can be obtained by applying on A a finite number of elementary row operations. Now by applying suitable column operations on B, a column-reduced echelon matrix C can be obtained. C has the following properties–

(i) No zero row is followed by a non-zero row;

(ii) no zero column is followed by a non-zero column;

(iii) the leading 1 in each non-zero row is the only non-zero element;

(iv) the leading 1 in each non-zero column is the only non-zero element;

(v) the leading 1 in the kth row is the leading 1 in the kth column.

Thus C takes the form $\begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix}$

where I_r is the identity matrix of order r and O_{pq} is a zero matrix of order $p \times q$.

C is said to be the *fully reduced normal form* of the matrix A.

Note. In particular, if A be an $n \times n$ matrix of rank r, then C is a diagonal matrix of order n whose first r diagonal elements are 1 and the remaining diagonal elements (if $r < n$) are all 0.

Example 5. Find the fully reduced normal form of the matrix

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 1 \\ 1 & 3 & 1 & 0 & 3 \\ 2 & 6 & 4 & 2 & 8 \\ 3 & 9 & 4 & 2 & 10 \end{pmatrix}.$$

Let us apply elementary operations on the matrix.

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 1 \\ 1 & 3 & 1 & 0 & 3 \\ 2 & 6 & 4 & 2 & 8 \\ 3 & 9 & 4 & 2 & 10 \end{pmatrix} \xrightarrow{R_{12}} \begin{pmatrix} 1 & 3 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 & 1 \\ 2 & 6 & 4 & 2 & 8 \\ 3 & 9 & 4 & 2 & 10 \end{pmatrix}$$

$$\xrightarrow[R_4-3R_1]{R_3-2R_1} \begin{pmatrix} 1 & 3 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 & 1 \end{pmatrix} \xrightarrow[R_3-2R_2,\ R_4-R_2]{R_1-R_2} \begin{pmatrix} 1 & 3 & 0 & -2 & 2 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{-\frac{1}{2}R_3} \begin{pmatrix} 1 & 3 & 0 & -2 & 2 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[R_2-2R_3]{R_1+2R_3} \begin{pmatrix} 1 & 3 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[C_5-2C_1]{C_2-3C_1} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{C_5-C_3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{C_{23}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{C_{34}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = R, \text{ say.}$$

R is the fully reduced normal form.

3.7. Elementary matrices.

An $n \times n$ matrix obtained by applying a single elementary row operation on I_n is said to be an *elementary matrix* of order n. There are three types of elementary matrices.

1. The elementary matrix obtained by applying R_{ij} on I_n is denoted by E_{ij}. $R_{ij}(I_n) = E_{ij}$.

2. The elementary matrix obtained by applying cR_i on I_n is denoted by $E_i(c)$. $cR_i(I_n) = E_i(c)$.

3. The elementary matrix obtained by applying $R_i + cR_j$ (or $R_{ij}(c)$) on I_n is denoted by $E_{ij}(c)$. $R_{ij}(c)(I_n) = E_{ij}(c)$.

Three types of elementary matrices are illustrated below.

$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, E_2(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}, E_{13}(c) = \begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Note. An elementary matrix of any of the three types can also be obtained by applying an elementary column operation on I.

Theorem 3.7.1. The effect of the elementary row operations R_{ij}, $cR_i (c \neq 0)$ and $R_i + cR_j$ on an $m \times n$ matrix A is obtained by premultiplying A by $E_{ij}, E_i(c)$ and $E_{ij}(c)$ respectively, where each elementary matrix is of order m.

Proof. Let P be an $r \times m$ matrix. Then PA is defined.

Let the rows (row matrices) of P be $P_1, P_2, ..., P_r$ and the columns (column matrices) of A be $A^{(1)}, A^{(2)}, ..., A^{(n)}$. The product PA can be expressed as

$$PA = \begin{pmatrix} P_1A^{(1)} & P_1A^{(2)} & \ldots & P_1A^{(n)} \\ P_2A^{(1)} & P_2A^{(2)} & \ldots & P_2A^{(n)} \\ \ldots & \ldots & \ldots & \\ P_rA^{(1)} & P_rA^{(2)} & \ldots & P_rA^{(n)} \end{pmatrix}.$$

If two rows of P, say P_i and P_j, be interchanged, then the ith row and jth row of the product PA are also interchanged.

Hence $\{R_{ij}(P)\}.A = R_{ij}(PA)$.

If a row of P, say P_i, be multiplied by a non-zero scalar c, then the ith row of PA is also multiplied by c.

Hence $\{R_i(c)(P)\}.A = R_i(c)(PA)$.

If c times the jth row of P be added to the ith row of P, then the ith row of PA is added by c times the jth row of it.

Hence $\{R_{ij}(c)(P)\}.A = R_{ij}(c)(PA)$.

A can be expressed as $A = I_mA$, where I_m is the $m \times m$ identity matrix.

So $R_{ij}(A) = R_{ij}(I_mA) = \{R_{ij}(I_m)\}.A = E_{ij}.A$,

$R_i(c)(A) = R_i(c)(I_mA) = \{R_i(c)(I_m)\}.A = E_i(c).A$;

$R_{ij}(c)(A) = R_{ij}(c)(I_mA) = \{R_{ij}(c)(I_m)\}.A = E_{ij}(c).A$.

Theorem 3.7.2. The effect of the elementary column operations C_{ij}, $cC_i(c \neq 0)$ and $C_i + cC_j$ on an $m \times n$ matrix A is obtained by post multiplying A by $\{E_{ij}\}^t, \{E_i(c)\}^t$ and $\{E_{ij}(c)\}^t$ respectively, where each elementary matrix is of order n.

Proof left to the reader.

Theorem 3.7.3. Each elementary matrix is non-singular. The inverse of an elementary matrix is an elementary matrix of the same type.

Proof. An elementary matrix of order n is obtained by an elementary row operation on the identity matrix I_n. Since the rank of a matrix remains same under an elementary row operation and rank of I_n is n, the rank of an elementary matrix of order n is n. So an elementary matrix is non-singular.

We have $E_{ij}.E_{ij} = R_{ij}[E_{ij}] = R_{ij}[R_{ij}(I_n)] = I_n$ and this implies $E_{ij}^{-1} = E_{ij}$.

We have $E_i(c^{-1}).E_i(c) = c^{-1}R_i[E_i(c)] = c^{-1}R_i[cR_i(I_n)] = I_n$ and similarly, $E_i(c).E_i(c^{-1}) = I_n$. This implies $\{E_i(c)\}^{-1} = E_i(c^{-1})$.

Likewise, $\{E_{ij}(c)\}^{-1} = E_{ij}(-c)$.

Thus the inverse of an elementary matrix is an elementary matrix of the same type.

Theorem 3.7.4. A matrix is non-singular if and only if it can be expressed as the product of a finite number of elementary matrices.

Proof. Let A be a non-singular matrix of order n. Then A is row equivalent to the identity matrix I_n. So I_n can be obtained from A by a finite number of elementary row operations.

Since the effect of an elementary row operation on A is expressed by the product EA where E is an $n \times n$ elementary matrix, I_n can be expressed as $I_n = E_kE_{k-1}...E_2E_1A$, where $E_1, E_2, ..., E_k$ are elementary matrices of order n.

Since each elementary matrix is non-singular, E_i^{-1} exists and therefore $E_1^{-1}E_2^{-1}...E_k^{-1}I_n = (E_1^{-1}E_2^{-1}...E_k^{-1})(E_kE_{k-1}...E_2E_1)A = A$

or, $A = P_1P_2...P_k$, where $P_i(= E_i^{-1})$ is an elementary matrix.

Conversely, let $A = E_1E_2...E_k$ where E_i is an elementary matrix. Since each E_i is a non-singular matrix, A is non-singular.
This completes the proof.

Theorem 3.7.5. An $m \times n$ matrix B is row equivalent to an $m \times n$ matrix A if and only if $B = PA$, where P is a non-singular matrix of order m.

Proof. Let B be row equivalent to A. Then B can be obtained by a finite number of elementary row operations on A.

So B can be expressed as $E_kE_{k-1}...E_1A$ where $E_1, E_2, ..., E_k$ are elementary matrices of order m.

Therefore $B = PA$ where $P = E_kE_{k-1}...E_1$ and P is non-singular, since each E_i is non-singular.

Conversely, let $B = PA$, where P is a non-singular matrix of order m.

Since P is non-singular, $P = E_1E_2...E_r$, where each E_i is an elementary matrix. Then $B = E_1E_2...E_rA$.

Since pre multiplication of A by an elementary matrix is equivalent to an elementary row operation, B is obtained from A by applying r elementary row operations on A. So B is row equivalent to A.
This completes the proof.

Theorem 3.7.6. An $m \times n$ matrix B is column equivalent to an $m \times n$ matrix A if and only if $B = AQ$, where Q is a non-singular matrix of order n.

Proof left to the reader.

Theorem 3.7.7. An $m \times n$ matrix B is equivalent to an $m \times n$ matrix A if and only if $B = PAQ$ where P, Q are non-singular matrices.

Proof left to the reader.

Theorem 3.7.8. Equivalence of matrices is an equivalence relation.

Proof left to the reader.

Theorem 3.7.9. Two $m \times n$ matrices are equivalent if and only if they have the same rank.

Proof. Let the $m \times n$ matrices A and B be equivalent. Then $B = PAQ$ for some non-singular matrices P, Q. Therefore rank of A= rank of B.

Conversely, let rank of A= rank of $B = r$. Then each is equivalent to the $m \times n$ matrix $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ and therefore they are equivalent.
This completes the proof.

Computation of the inverse of a matrix by elementary row operations.

Let A be a non-singular matrix of order n. Then A is row equivalent to the identity matrix I_n. Hence for suitable elementary matrices E_i, $E_rE_{r-1}...E_2E_1A = I_n$.

Multiplying by A^{-1} from the right, we have $E_rE_{r-1}...E_1I_n = A^{-1}$.

Thus if $E_1, E_2, ..., E_r$ be a finite number of elementary matrices which reduce A on successive premultiplication (in the order $E_1, E_2, ..., E_r$) to I_n, then the same sequence of elementary matrices will reduce I_n (on successive premultiplication) to A^{-1}.

Therefore if a sequence of elementary row operations applied successively on A reduces A to I_n, the same sequence of operations applied successively on I_n will reduce I_n to A^{-1}.

This gives us a technique for finding A^{-1} described below.

Form the $n \times 2n$ matrix $(A|I_n)$. Apply elementary row operations $R_1, R_2, ..., R_k$ successively on the matrix $(A|I_n)$ which will reduce A to I_n. Then I_n will be reduced by the operations to A^{-1}. So the matrix $(A|I_n)$ will be reduced to $(I_n|A^{-1})$.

Worked Examples.

1. Find the inverse of A where $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 4 & 4 \\ 3 & 3 & 7 \end{pmatrix}$.

Let us form the 3×6 matrix $(A|I_3)$ and apply elementary row operations to reduce A to I_3.

$$(A|I_3) = \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 4 & 4 & 0 & 1 & 0 \\ 3 & 3 & 7 & 0 & 0 & 1 \end{bmatrix} \xrightarrow[R_3-3R_1]{R_2-2R_1} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}R_2} \begin{bmatrix} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{bmatrix} \xrightarrow{R_1-R_2} \begin{bmatrix} 1 & 0 & 2 & 2 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1-2R_3} \begin{bmatrix} 1 & 0 & 0 & 8 & -\frac{1}{2} & -2 \\ 0 & 1 & 0 & -1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & -3 & 0 & 1 \end{bmatrix} = (I_3|A^{-1}).$$

Therefore $A^{-1} = \begin{bmatrix} 8 & -\frac{1}{2} & -2 \\ -1 & \frac{1}{2} & 0 \\ -3 & 0 & 1 \end{bmatrix}$.

2. Show that the matrix $\begin{pmatrix} 2 & 0 & 1 \\ 3 & 3 & 0 \\ 6 & 2 & 3 \end{pmatrix}$ is non-singular and express it as a product of elementary matrices.

Let A be the given matrix and let us apply elementary row operations on A to reduce it to a row reduced echelon matrix.

$$A \xrightarrow{\frac{1}{2}R_1} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 3 & 3 & 0 \\ 6 & 2 & 3 \end{pmatrix} \xrightarrow[R_3-6R_1]{R_2-3R_1} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 3 & -\frac{3}{2} \\ 0 & 2 & 0 \end{pmatrix} \xrightarrow{\frac{1}{3}R_2} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 2 & 0 \end{pmatrix}$$

$$\xrightarrow{R_3-2R_2} \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow[R_2+\frac{1}{2}R_3]{R_1-\frac{1}{2}R_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since A is row equivalent to I_3, A is non-singular.

$[(R_2+\frac{1}{2}R_3)(R_1-\frac{1}{2}R_3)(R_3-2R_2)(\frac{1}{3}R_2)(R_3-6R_1)(R_2-3R_1)\frac{1}{2}R_1]A = I_3$,

or, $E_{23}(\frac{1}{2})E_{13}(-\frac{1}{2})E_{32}(-2)E_2(\frac{1}{3})E_{31}(-6)E_{21}(-3)E_1(\frac{1}{2})A = I_3$.

$A = \{E_1(\frac{1}{2})\}^{-1}\{E_{21}(-3)\}^{-1}\{E_{31}(-6)\}^{-1}\{E_2(\frac{1}{3})\}^{-1}\{E_{32}(-2)\}^{-1}$
$\{E_{13}(-\frac{1}{2})\}^{-1}\{E_{23}(\frac{1}{2})\}^{-1}$

$= E_1(2)E_{21}(3)E_{31}(6)E_2(3)E_{32}(2)E_{13}(\frac{1}{2})E_{23}(-\frac{1}{2})$.

Note. This factorisation is not unique since different sequences of elementary row operations can be used to reduce the matrix A to a row-reduced echelon matrix.

3. $A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}$. Find non-singular matrices P and Q such that PAQ is the fully reduced normal form.

Let us apply elementary operations on A.

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix} \xrightarrow[R_3-R_1]{R_2-2R_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{R_3-R_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{C_3-C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{C_3-C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = R, \text{ say.}$$

R is the fully reduced normal form of A.

$E_{32}(-1)E_{31}(-1)E_{21}(-2)A\{E_{31}(-1)\}^t\{E_{32}(-1)\}^t = R$.

Let $P = E_{32}(-1)E_{31}(-1)E_{21}(-2), Q = E_{13}(-1)E_{23}(-1)$. Then

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

$$Q = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

$PAQ = R$. P and Q are non-singular, since each is the product of elementary matrices.

Exercises 4

1. A is a non-singular matrix of order 4. Determine the rank of the matrix
(i) A^{-1}, (ii) A^3, (iii) $adj\ A$, (iv) $2A$.

2. If the rank of a real symmetric matrix be 1 show that the diagonal elements of the matrix cannot be all zero.

3. Prove that the rank of a real skew-symmetric matrix cannot be 1.

4. A is an $n \times n$ matrix of rank $n-1$. Prove that the rank of $adj\ A$ is 1.

[Hint. Since det $A = 0$, rank$(adj\ A) \leq 1$, by Theorem 2.4.3. Since rank of A is $n-1$, there is at least one non-zero minor of A of order $n-1$ and therefore rank$(adj\ A) \geq 1$.]

5. Find all real x for which the rank of the matrix is less than 4.

(i) $\begin{pmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{pmatrix}$, (ii) $\begin{pmatrix} 1+x & 2 & 3 & 4 \\ 1 & 2+x & 3 & 4 \\ 1 & 2 & 3+x & 4 \\ 1 & 2 & 3 & 4+x \end{pmatrix}$,

(iii) $\begin{pmatrix} 1+x & 2 & 3 & 4 \\ 2 & 1+x & 3 & 4 \\ 2 & 3 & 1+x & 4 \\ 2 & 3 & 4 & 1+x \end{pmatrix}$.

6. Find all real λ for which the rank of the matrix A is 2.

(i) $A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 2 & 5 & 3 & \lambda \\ 1 & 1 & 6 & \lambda+1 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & \lambda \\ 5 & 7 & 1 & \lambda^2 \end{pmatrix}$.

7. If $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ find the rank of the matrix

(i) $A + A^2$, (ii) $A + A^2 + A^3$.

8. Reduce the matrix A to the fully reduced normal form and find non-singular matrices P and Q such that PAQ is the fully reduced normal form.

(i) $A = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & 1 & 4 & 6 \\ 3 & 0 & 7 & 9 \end{pmatrix}$.

9. Use elementary row operations on A to obtain A^{-1} where A is

(i) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, (ii) $\begin{pmatrix} 2 & 0 & 0 \\ 4 & 3 & 0 \\ 6 & 4 & 1 \end{pmatrix}$, (iii) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$.

10. Express A as the product of elementary matrices.

(i) $A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & 5 \\ 0 & 0 & 2 \end{pmatrix}$, (iii) $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

11. A and B are equivalent square matrices. Prove that

(i) the matrices A^t and B^t are equivalent;

(ii) the matrices A^2 and B^2 are not necessarily equivalent;

(iii) the matrices AB and BA are equivalent if one of A and B be non-singular;

(iv) the matrices AB and BA are not necessarily equivalent if none of A and B be non-singular.

3.8. Congruence operations and congruence of matrices.

Let A be an $n \times n$ matrix over a field F. Let us consider three types of elementary matrices $E_{ij}, E_i(c)$ and $E_{ij}(c)$ of order n.

Premultiplication of A by E_{ij} gives the effect of the row operation R_{ij} on A. Post multiplication of A by $E_{ij}{}^t$ gives the effect of the column operation C_{ij} on A.

Premultiplication of A by $E_i(c)$ gives the effect of the row operation cR_i on A. Post multiplication of A by $\{E_i(c)\}^t$ gives the effect of the column operation cC_i on A.

Premultiplication of A by $E_{ij}(c)$ gives the effect of the row operation $R_i + cR_j$ on A. Post multiplication of A by $\{E_{ij}(c)\}^t$ gives the effect of the column operation $C_i + cC_j$ on A.

Therefore whatever the elementary matrix E may be, the product EAE^t gives an elementary row operation together with the corresponding elementary column operation on A. Such an operation given by the product EAE^t is called a *congruence operation* on A.

Definition.

Let S be the set of all $n \times n$ matrices over a field F. A matrix $B \in S$ is said to be *congruent* to a matrix $A \in S$ if there exists a non-singular matrix $P \in S$ such that $B = P^tAP$.

Since P is non-singular, it can be expressed as the product $E_1E_2 \ldots E_k$ where E's are elementary matrices of order n.

Therefore $B = E_k{}^tE_{k-1}{}^t \ldots E_2{}^tE_1{}^tAE_1E_2 \ldots E_{k-1}E_k$.

Since $E_i{}^tAE_i$ gives a congruence operation on A, B is obtained by applying k successive congruence operations on A.

Since the rank of a matrix remains unaltered by an elementary operation, the rank remains invariant under a congruence operation. In other words, two congruent matrices have the same rank.

Theorem 3.8.1. The relation of congruence on the set of all $n \times n$ matrices over a field F is an equivalence relation.

Proof. Let S be the set of all $n \times n$ matrices over a field F.

Let ρ be the relation of congruence on S.

(i) $A \rho A$ holds for all $A \in S$ because I_n is a non-singular matrix in S and $A = I_n{}^tAI_n$. Therefore ρ is reflexive.

(ii) Let $A, B \in S$ and $A \rho B$ hold. Then there exists a non-singular matrix P in S such that $B = P^tAP$.

$$\begin{aligned} \text{Therefore } A &= (P^t)^{-1}BP^{-1} \\ &= (P^{-1})^tBP^{-1} \\ &= Q^tBQ, \text{ where } Q \in S \text{ and is non-singular.} \end{aligned}$$

This implies $B \rho A$. Thus $A \rho B \Rightarrow B \rho A$ and therefore ρ is symmetric.

(iii) Let $A, B, C \in S$ and $A \rho B$ and $B \rho C$ hold.

Then there exist non-singular matrices P, Q in S such that $B = P^tAP$ and $C = Q^tBQ$.

$$\begin{aligned} \text{Therefore } C &= Q^t(P^tAP)Q \\ &= (PQ)^tA(PQ) \\ &= M^tAM, \text{ where } M \in S \text{ and is non-singular.} \end{aligned}$$

This implies $A \rho C$.

Thus $A \rho B$ and $B \rho C \Rightarrow A \rho C$ and therefore ρ is transitive.

Hence ρ is an equivalence relation on S.

In view of the theorem, the set S is partitioned into equivalence classes of congruent matrices.

If however, A is symmetric then P^tAP is also symmetric, since $(P^tAP)^t = P^tAP$. Thus each matrix in a congruence class of a symmetric matrix is symmetric.

We shall be mainly interested in the congruence class of a real symmetric matrix.

Theorem 3.8.2. An $n\times n$ real symmetric matrix A of rank r is congruent to an $n \times n$ real diagonal matrix D with non-zero elements in the first r diagonal positions and zero elsewhere.

Proof. When $A = O$ then clearly, $D = O$ and the theorem holds.

Let $A = (a_{ij}) \neq O$. We consider the following cases.

Case 1. $a_{11} \neq 0$.

Let us apply the row operations $R_2 - \frac{a_{12}}{a_{11}}R_1$, $R_3 - \frac{a_{21}}{a_{11}}R_1, \ldots, R_n - \frac{a_{n1}}{a_{11}}R_1$ and the corresponding column operations $C_2 - \frac{a_{12}}{a_{11}}C_1$, $C_3 - \frac{a_{13}}{a_{11}}C_1, \ldots, C_n - \frac{a_{1n}}{a_{11}}C_1$ on A. Then A is transformed to the congruent symmetric matrix $B = (b_{ij})$ in which $b_{11} = a_{11} \neq 0$, and $b_{1j} = b_{j1} = 0, j = 2, 3, \ldots, n$.

Case 2. $a_{11} = 0$ but at least one of the diagonal elements of A, say $a_{rr} \neq 0$.

Let us apply the row operation R_{1r} and the corresponding column operation C_{1r} on A. Then A is transformed to the congruent symmetric matrix $B = (b_{ij})$ in which $b_{11} = a_{rr} \neq 0$.

Next we proceed as in Case 1 and obtain a congruent symmetric matrix $C = (c_{ij})$ in which $c_{11} = b_{11} \neq 0$ and $c_{1j} = c_{j1} = 0, j = 2, 3, \ldots, n$.

Case 3. All diagonal elements of A are zero, $a_{rs} \neq 0$.

Let us apply the row operation $R_r + R_s$ and the corresponding column operation $C_r + C_s$ on A. Then A is transformed to a congruent matrix $B = (b_{ij})$ in which $b_{rr} = 2a_{rs} \neq 0$.

Then proceeding as in case 2 we obtain a congruent symmetric matrix $C = (c_{ij})$ in which $c_{11} \neq 0$ and $c_{1j} = c_{j1} = 0, j = 2, 3, \ldots, n$.

Thus in all cases A is transformed to a congruent symmetric matrix $C = (c_{ij})$ in which $c_{11} \neq 0$ and $c_{1j} = c_{j1} = 0, j = 2, 3, \ldots, n$.

Therefore $C = \begin{pmatrix} c_{11} & O \\ O & M \end{pmatrix}$, where M is a symmetric matrix of order $(n-1)$.

If $M = O$, C is a diagonal matrix and the process terminates.

If $M \neq O$, the same procedure as described in the previous cases can be applied to reduce M to a congruent symmetric matrix with a non-zero element in the $(1,1)$ position and zeros elsewhere in the first row and first column of M. All these operations however are to be applied on the matrix C but these will not affect the first row and the first column of C.

Thus A is transformed to a congruent symmetric matrix $G = \begin{pmatrix} g_{11} & & \\ & g_{22} & \\ & & H \end{pmatrix}$, where $g_{11} \neq 0, g_{22} \neq 0, g_{1j} = g_{j1} = 0, j = 2, 3, \ldots, n, g_{2j} = g_{j2} = 0, j = 1, 3, \ldots, n$ and H is a symmetric matrix of order $(n-2)$. Repeating the same procedure, if necessary, we obtain in a finite number of steps a congruent diagonal matrix $D = (d_{ij})$.

Since the rank remains invariant under congruence operations, rank of $D = r$ and therefore

$$\begin{aligned} d_{ii} &\neq 0, i = 1, 2, \ldots, r \\ d_{ii} &= 0, i = r+1, r+2, \ldots, n. \end{aligned}$$

Theorem 3.8.3. An $n\times n$ real symmetric matrix A of rank r is congruent to the diagonal matrix G whose first m diagonal elements are 1, the next $r-m$ diagonal elements are -1 and the remaining diagonal elements, if there be any, are all zero.

Proof. By the previous theorem, A is congruent to the diagonal matrix $D = (d_{ij})$ where $d_{ii} \neq 0, i = 1, 2, \ldots, r$ and $d_{ii} = 0, i = r+1, r+2, \ldots, n$.

If the non-zero elements in D be not all positive, without loss of generality, we can assume that first m diagonal elements are positive and the next $r-m$ diagonal elements are negative. Because, if a positive diagonal element appears after a negative element, they can be interchanged by suitable congruence operations.

Next let us apply the row operations $\frac{1}{\sqrt{d_{11}}}R_1, \frac{1}{\sqrt{d_{22}}}R_2, \cdots, \frac{1}{\sqrt{d_{mm}}}R_m, \frac{1}{\sqrt{-d_{m+1m+1}}}R_{m+1}, \cdots, \frac{1}{\sqrt{-d_{rr}}}R_r$ and the corresponding column operations on D.

Then D is transformed to a congruent diagonal matrix $G = (g_{ij})$, where $g_{11} = g_{22} = \cdots = g_{mm} = 1, g_{m+1,m+1} = \cdots = g_{rr} = -1, g_{ii} = 0$, for $i = r+1, r+2, \ldots, n$.

Thus G takes the form $\begin{pmatrix} I_m & & \\ & -I_{r-m} & \\ & & O \end{pmatrix}$.

G is said to be the *normal form of A under congruence*.

Definition. The integer m which is the number of positive 1's in the normal form of a real symmetric matrix A under congruence is invariant. This m is said to be the *index* of the symmetric matrix A.

Since the rank r is invariant and the index m is invariant under congruence, the integer $m-(r-m)=2m-r$ is also invariant under congruence. This integer $m-(r-m)$ is called the *signature* of the symmetric matrix A.

Theorem 3.8.4. Two real symmetric matrices of the same order are congruent if and only if they have the same rank and same signature.

Proof. Let A and B be two $n\times n$ real symmetric matrices and let A and B be congruent. Since the rank of a matrix remains invariant under congruence, rank of A = rank of B. Since the signature of a symmetric matrix remains invariant under congruence, the signature of A = the signature of B.

Conversely, let A and B have the same rank and signature.

Then both the matrices A and B have the same normal form under congruence. Let G be the normal form of both A and B. Then A and B being both congruent to G, are congruent to each other.

Worked Examples.

1. Obtain the normal form under congruence and find the rank and signature of the symmetric matrices

(i) $A=\begin{bmatrix} 2 & 4 & 3\\ 4 & 6 & 3\\ 3 & 3 & 1\end{bmatrix}$, (ii) $B=\begin{pmatrix} 0 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 1\end{pmatrix}$, (iii) $M=\begin{bmatrix} 0 & 2 & 1\\ 2 & 0 & 1\\ 1 & 1 & 0\end{bmatrix}$

Let us apply congruence operations on the matrices.

$$\text{(i) } A \xrightarrow[R_3-\frac{3}{2}R_1]{R_2-2R_1} \begin{bmatrix} 2 & 4 & 3\\ 0 & -2 & -3\\ 0 & -3 & -\frac{7}{2}\end{bmatrix} \xrightarrow[C_3-\frac{3}{2}C_1]{C_2-2C_1} \begin{bmatrix} 2 & 0 & 0\\ 0 & -2 & -3\\ 0 & -3 & -\frac{7}{2}\end{bmatrix}$$

$$\xrightarrow{R_3-\frac{3}{2}R_2} \begin{bmatrix} 2 & 0 & 0\\ 0 & -2 & -3\\ 0 & 0 & 1\end{bmatrix} \xrightarrow{C_3-\frac{3}{2}C_2} \begin{bmatrix} 2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$\xrightarrow{R_{23}} \begin{bmatrix} 2 & 0 & 0\\ 0 & 0 & 1\\ 0 & -2 & 0\end{bmatrix} \xrightarrow{C_{23}} \begin{bmatrix} 2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -2\end{bmatrix}$$

$$\xrightarrow[\frac{1}{\sqrt{2}}R_3]{\frac{1}{\sqrt{2}}R_1} \begin{bmatrix} \sqrt{2} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -\sqrt{2}\end{bmatrix} \xrightarrow[\frac{1}{\sqrt{2}}C_3]{\frac{1}{\sqrt{2}}C_1} \begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{bmatrix}.$$

The rank of $A=3$ and the signature of $A=1$.

$$\text{(ii) } B \xrightarrow{R_{12}} \begin{pmatrix} 2 & 4 & 2\\ 0 & 2 & 1\\ 1 & 2 & 1\end{pmatrix} \xrightarrow{C_{12}} \begin{pmatrix} 4 & 2 & 2\\ 2 & 0 & 1\\ 2 & 1 & 1\end{pmatrix}$$

$$\xrightarrow[R_3-\frac{1}{2}R_1]{R_2-\frac{1}{2}R_1} \begin{pmatrix} 4 & 2 & 2\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix} \xrightarrow[C_3-\frac{1}{2}C_1]{C_2-\frac{1}{2}C_1} \begin{pmatrix} 4 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}$$

$$\xrightarrow{\frac{1}{2}R_1} \begin{pmatrix} 2 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix} \xrightarrow{\frac{1}{2}C_1} \begin{pmatrix} 1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

The rank of $B=2$ and the signature of $B=0$.

$$\text{(iii) } M \xrightarrow{R_1+R_2} \begin{bmatrix} 2 & 2 & 2\\ 2 & 0 & 1\\ 1 & 1 & 0\end{bmatrix} \xrightarrow{C_1+C_2} \begin{bmatrix} 4 & 2 & 2\\ 2 & 0 & 1\\ 2 & 1 & 0\end{bmatrix}$$

$$\xrightarrow[R_3-\frac{1}{2}R_1]{R_2-\frac{1}{2}R_1} \begin{bmatrix} 4 & 2 & 2\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix} \xrightarrow[C_3-\frac{1}{2}C_1]{C_2-\frac{1}{2}C_1} \begin{bmatrix} 4 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}$$

$$\xrightarrow{\frac{1}{2}R_1} \begin{bmatrix} 2 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix} \xrightarrow{\frac{1}{2}C_1} \begin{bmatrix} 1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}.$$

The rank of $M=3$ and the signature of $M=-1$.

2. Find a non-singular matrix P such that P^tAP is a diagonal matrix, where A is the matrix of the previous example.

From the previous example, we have

$$E_{32}(-\tfrac{3}{2})E_{31}(-\tfrac{3}{2})E_{21}(-2)A\{E_{21}(-2)\}^t\{E_{31}(-\tfrac{3}{2})\}^t\{E_{32}(-\tfrac{3}{2})\}^t=D,$$

where $D=\begin{bmatrix} 2 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 1\end{bmatrix}$.

Let $P=\{E_{21}(-2)\}^t\{E_{31}(-\frac{3}{2})\}^t\{E_{32}(-\frac{3}{2})\}^t$.

Then $P^tAP=D$ and P is non-singular, since P is the product of elementary matrices.

$P=E_{12}(-2)E_{13}(-\frac{3}{2})E_{23}(-\frac{3}{2})=$

$$\begin{bmatrix} 1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix} 1 & 0 & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & -\frac{3}{2}\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix} 1 & -2 & \frac{3}{2}\\ 0 & 1 & -\frac{3}{2}\\ 0 & 0 & 1\end{bmatrix}.$$

Exercises 5

1. Determine the rank and the signature of the following symmetric matrices.

(i) $\begin{bmatrix} 2 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 10 \end{bmatrix}$, (ii) $\begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & 2 \\ 0 & 2 & 4 \end{bmatrix}$, (iii) $\begin{bmatrix} 5 & 4 & 5 \\ 4 & 5 & 7 \\ 5 & 7 & 10 \end{bmatrix}$,

(iv) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 2 \\ 2 & 2 & 0 \end{bmatrix}$, (v) $\begin{bmatrix} 0 & 2 & 1 \\ 2 & 3 & 4 \\ 1 & 4 & 2 \end{bmatrix}$, (vi) $\begin{bmatrix} 0 & 2 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 0 \end{bmatrix}$.

2. Examine if the matrices A and B are congruent.

(i) $A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 4 & 1 \\ 4 & 9 & 4 \\ 1 & 4 & 2 \end{bmatrix}$.

(ii) $A = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 & 1 \\ 4 & 9 & 4 \\ 1 & 4 & 3 \end{bmatrix}$.

3. Find a non-singular matrix P such that $P^t AP$ is the normal form of A under congruence.

(i) $A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 3 \end{pmatrix}$.

4. Find a non-singular matrix P such that $P^t AP$ is a diagonal matrix, where

(i) $A = \begin{bmatrix} 2 & 2 & 3 \\ 2 & 5 & 6 \\ 3 & 6 & 10 \end{bmatrix}$, (ii) $A = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 3 & 2 \\ 0 & 2 & 4 \end{bmatrix}$.

5. Prove that a non-singular symmetric matrix A is congruent to A^{-1}.

[**Hint.** Let G be the normal form of A under congruence. Then G is a non-singular diagonal matrix. The matrices G and G^{-1} have the same rank and the same index. Again A is congruent to G implies A^{-1} is congruent to G^{-1}.]

4. Vector Spaces

4.1. External composition.

Let U and S be two non-empty sets. A mapping $f : U \times S \to S$ is said to be an *external composition* of U with S. Each ordered pair (u, s) in $U \times S$ has a definite image $f(u, s)$ in S.

Examples.

1. Let S be the set of all real matrices of order $m \times n$. The mapping $* : \mathbb{R} \times S \to S$ defined by $c * A = cA$, $c \in \mathbb{R}$, $A \in S$ is an external composition of $\mathbb{R}$ with S. For each real number c and each A in S, cA is an $m \times n$ matrix in S. $*$ is called 'multiplication of $m \times n$ matrices by real numbers.'

2. Let G be a group under a multiplicative composition. Let a mapping $* : \mathbb{Z} \times G \to G$ be defined by $n * a = a^n$, $n \in \mathbb{Z}, a \in G$. For each integer n and each a in G, a^n is an element in G. $*$ is an external composition of $\mathbb{Z}$ with G.

4.2. Vector space over a Field.

Let V be a non-empty set and $\oplus$ be a binary composition on V. Let $(F, +, .)$ be a field and let $\odot$ be an external composition of F with V.

[$\oplus$ is a mapping from $V \times V$ to V. $\odot$ is a mapping from $F \times V$ to V and it maps the ordered pair (c, α) of $F \times V$ to a definite element $c \odot \alpha$ in V.]

V is said to be a *vector space* (or a *linear space*) *over the field F* if the following conditions are satisfied.

V1. $\alpha \oplus \beta \in V$ for all $\alpha, \beta \in V$;
V2. $\alpha \oplus \beta = \beta \oplus \alpha$ for all $\alpha, \beta \in V$;
V3. $\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma$ for all $\alpha, \beta, \gamma \in V$;
V4. there exists an element θ in V such that $\alpha \oplus \theta = \alpha$ for all $\alpha \in V$;
V5. for each α in V there exists an element $-\alpha$ in V such that $\alpha \oplus (-\alpha) = \theta$;
V6. $c \odot \alpha \in V$ for all $c \in F$, all $\alpha \in V$;
V7. $c \odot (d \odot \alpha) = (c.d) \odot \alpha$ for all $c, d \in F$, all $\alpha \in V$;
V8. $c \odot (\alpha \oplus \beta) = (c \odot \alpha) \oplus (c \odot \beta)$ for all $c \in F$, all $\alpha, \beta \in V$;

V9. $(c+d)\odot\alpha = (c\odot\alpha)\oplus(d\odot\alpha)$ for all $c,d\in F$, all $\alpha\in V$;
V10. $1\odot\alpha=\alpha$ for all $\alpha\in V$, 1 being the identity element in F.

The vector space is denoted by $(V,F,+,.,\oplus,\odot)$. The elements of V are called *vectors* and the elements of F are called *scalars*. F is called the *ground field* (or the *field of scalars*) of the vector space.

Four symbols $+,.,\oplus,\odot$ denote four different compositions $+ : F\times F\to F$; $. : F\times F\to F$; $\oplus : V\times V\to V$; $\odot : F\times V\to V$. We shall dispense with $\oplus$ and use only $+$ to denote both types of addition. We shall dispense with $\odot$ and . both and denote $c.d$ in F by cd and denote $c\odot\alpha$ in V by $c\alpha$.

Therefore a non-empty set V is said to form a vector space over a field F if

(i) there is a binary composition $+$ on V, called 'addition', satisfying the conditions –

V1. $\alpha+\beta\in V$ for all $\alpha,\beta\in V$,
V2. $\alpha+\beta=\beta+\alpha$ for all $\alpha,\beta\in V$,
V3. $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$ for all $\alpha,\beta,\gamma\in V$,
V4. there exists an element θ in V such that $\alpha+\theta=\alpha$ for all α in V,
V5. for each α in V there exists an element $-\alpha$ in V such that $\alpha+(-\alpha)=\theta$;

and (ii) there is an external composition of F with V, called 'multiplication by the elements of F' satisfying the conditions –

V6. $c\alpha\in V$ for all $c\in F$, all $\alpha\in V$,
V7. $c(d\alpha)=(cd)\alpha$ for all $c,d\in F$, all $\alpha\in V$,
V8. $c(\alpha+\beta)=c\alpha+c\beta$ for all $c\in F$, all $\alpha,\beta\in V$,
V9. $(c+d)\alpha=c\alpha+d\alpha$ for all $c,d\in F$, all $\alpha\in V$,
V10. $1\alpha=\alpha$ for all $\alpha\in V$, 1 being the identity element in F.

The elements of V are called *vectors* and the elements of F are called *scalars*. The external composition of F with V is also called, 'multiplication by scalars'.

In particular, V is said to be a *real vector space* (or a *complex vector space*) if the field F be $\mathbb{R}$ (or $\mathbb{C}$).

Since F is a field, it has the zero element 0 and the identity element 1. Since V is a commutative group with respect to addition, the identity element with respect to addition is the zero element in V. This is said to be the *zero vector* or the *null vector* and is denoted by θ. In order to avoid confusion, two zero elements (the scalar zero in F and the vector zero in V) will appear with different symbols. But when no such confusion occurs the symbol 0 will be used for both the zero elements.

Real vector space.

A non-empty set V is said to form a *real vector space* (or a vector space over the field $\mathbb{R}$) if

(i) there is a binary composition $(+)$ on V, called 'addition', satisfying the conditions –

V1. $\alpha+\beta\in V$ for all $\alpha,\beta\in V$;
V2. $\alpha+\beta=\beta+\alpha$ for all $\alpha,\beta\in V$;
V3. $\alpha+(\beta+\gamma)=(\alpha+\beta)+\gamma$ for all $\alpha,\beta,\gamma\in V$;
V4. there exists an element θ in V such that $\alpha+\theta=\alpha$ for all $\alpha\in V$;
V5. for each α in V there exists an element $-\alpha$ in V such that $\alpha+(-\alpha)=\theta$;

and (ii) there is an external composition of $\mathbb{R}$ with V, called 'multiplication by real numbers' satisfying the conditions –

V6. $c\alpha\in V$ for all $c\in\mathbb{R}$, all $\alpha\in V$;
V7. $c(d\alpha)=(cd)\alpha$ for all $c,d\in\mathbb{R}$, all $\alpha\in V$;
V8. $c(\alpha+\beta)=c\alpha+c\beta$ for all $c\in\mathbb{R}$, all $\alpha,\beta\in V$;
V9. $(c+d)\alpha=c\alpha+d\alpha$ for all $c,d\in\mathbb{R}$, all $\alpha\in V$;
V10. $1\alpha=\alpha$ for all $\alpha\in V$, 1 being the identity element in $\mathbb{R}$.

The elements of V are called *vectors* and the elements of $\mathbb{R}$ are called *scalars*. $\mathbb{R}$ is said to be the *ground field* (or the *field of scalars*) of the vector space V.

Examples.

1. **Real vector space $\mathbb{R}^n$.**

Let V be the set of all ordered n-tuples $\{(a_1,a_2,\ldots,a_n) : a_i\in\mathbb{R}\}$.

Let $+$ be a composition on V, called 'addition', defined by

$(a_1,a_2,\ldots,a_n)+(b_1,b_2,\ldots,b_n)=(a_1+b_1,a_2+b_2,\cdots,a_n+b_n)$

and an external composition of $\mathbb{R}$ with V, called 'multiplication by real numbers' be defined by

$c(a_1,a_2,\ldots,a_n)=(ca_1,ca_2,\ldots,ca_n), c\in\mathbb{R}$.

Then the conditions **V1-V10** are satisfied. Therefore V is a real vector space and it is denoted by $\mathbb{R}^n$.

$(0,0,\ldots,0)$ is the null vector of $\mathbb{R}^n$ and it is denoted by θ.

In a similar manner the vector spaces $\mathbb{R}^2,\mathbb{R}^3,\mathbb{R}^4,\ldots$ are defined. The set $\mathbb{R}$ itself forms a real vector space.

2. **Real vector space $\mathbb{C}$.**

$\mathbb{C}$ is the set of all complex numbers $\{a+ib : a\in\mathbb{R}, b\in\mathbb{R}, i=\sqrt{(-1)}\}$.

Let $+$ be a composition on $\mathbb{C}$, called 'addition', defined by

$(a+ib)+(c+id)=(a+c)+i(b+d)$;

and an external composition of $\mathbb{R}$ with $\mathbf{C}$, called 'multiplication by real numbers' be defined by $c(a+ib) = (ca) + i(cb), c \in \mathbb{R}$.

Then the conditions **V1-V10** are satisfied. Therefore $\mathbf{C}$ is a real vector space.

3. Real vector space P_n.

A real polynomial of degree r is $a_0 + a_1x + a_2x^2 + \cdots + a_rx^r$, where $a_0, a_1, \ldots, a_r$ are real numbers with $a_r \neq 0$.

Let V be the set of all real polynomials of degree $< n$ together with the zero polynomial (a polynomial which is identically zero and this has no degree at all).

Let $+$ be a composition on V, called 'addition', defined by

$(a_0 + a_1x + \cdots + a_mx^m) + (b_0 + b_1x + \cdots + b_rx^r)$

$= (a_0 + b_0) + \cdots + (a_m + b_m)x^m + b_{m+1}x^{m+1} + \cdots + b_rx^r$, if $m < r$

$= (a_0 + b_0) + \cdots + (a_r + b_r)x^r + a_{r+1}x^{r+1} + \cdots + a_mx^m$, if $r < m$

$= (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_r + b_r)x^r$, if $r = m$;

and an external composition of $\mathbb{R}$ with V, called 'multiplication of polynomials by real numbers' be defined by

$$c(a_0 + a_1x + \cdots + a_rx^r) = (ca_0) + (ca_1)x + \cdots + (ca_r)x^r, c \in \mathbb{R}$$

Then the conditions **V1-V10** are satisfied. Therefore V is a real vector space and it is denoted by P_n.

4. Real vector space $\mathbb{R}_{m\times n}$.

Let V be the set of all $m \times n$ matrices over $\mathbb{R}$.

Let $+$ be a composition on V, called 'addition', defined by

$$(a_{ij})_{m,n} + (b_{ij})_{m,n} = (a_{ij} + b_{ij})_{m,n};$$

and an external composition of $\mathbb{R}$ with V, called 'multiplication of matrices by real numbers' be defined by

$$c(a_{ij})_{m,n} = (ca_{ij})_{m,n}, \ c \in \mathbb{R}.$$

Then the conditions **V1-V10** are satisfied. Therefore V is a real vector space and it is denoted by $\mathbb{R}_{m\times n}$.

5. Let F be a field and V_n be the set of all ordered n-tuples $\{(a_1, a_2, \ldots, a_n) : a_i \in F\}$.

Let $+$ be a composition on V_n, called 'addition', defined by

$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \cdots, a_n + b_n)$;

and an external composition of F with V_n, called 'multiplication by scalars in F' be defined by

$$c(a_1, a_2, \ldots, a_n) = (ca_1, ca_2, \ldots, ca_n), c \in F.$$

Then the conditions **V1-V10** are satisfied. Therefore V_n is a vector space over the field F and it is denoted by $V_n(F)$, or F^n.

Theorem 4.2.1. In a vector space V over a field F,

(i) $0\alpha = \theta$ for all $\alpha \in V$;

(ii) $c\theta = \theta$ for all $c \in F$;

(iii) $-1\alpha = -\alpha$ for all $\alpha \in V$, 1 being the identity element in F;

(iv) $c\alpha = \theta$ implies either $c = 0$ or $\alpha = \theta$.

Proof. (i) 0 is the zero element in F.

$0 + 0 = 0$ in F

$\Rightarrow (0+0)\alpha = 0\alpha$ in V

$\Rightarrow 0\alpha + 0\alpha = 0\alpha$, by V_9.

$-0\alpha \in V$, since $0\alpha \in V$.

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

or, $(-0\alpha + 0\alpha) + 0\alpha = \theta$, by V_3 and V_5

or, $\theta + 0\alpha = \theta$, by V_5

or, $0\alpha = \theta$, by V_4.

(ii) θ is the zero element in V.

$\theta + \theta = \theta$ in V

$\Rightarrow c(\theta + \theta) = c\theta$

$\Rightarrow c\theta + c\theta = c\theta$, by V_8.

$-c\theta \in V$, since $c\theta \in V$.

Therefore $-c\theta + (c\theta + c\theta) = -c\theta + c\theta$

or, $(-c\theta + c\theta) + c\theta = \theta$, by V_3 and V_5

or, $\theta + c\theta = \theta$, by V_5

or, $c\theta = \theta$, by V_4.

(iii) We have
$$\begin{aligned} \theta &= 0\alpha, \text{ by (i)} \\ &= [1 + (-1)]\alpha \\ &= 1\alpha + (-1)\alpha \\ &= \alpha + (-1)\alpha, \text{ by } V_{10}. \end{aligned}$$

Therefore
$$\begin{aligned} -\alpha + \theta &= -\alpha + \{\alpha + (-1)\alpha\} \\ &= (-\alpha + \alpha) + (-1)\alpha \\ &= \theta + (-1)\alpha, \text{ by } V_5 \\ &= (-1)\alpha \\ \text{or, } -\alpha &= (-1)\alpha. \end{aligned}$$

(iv) Let $c\alpha = \theta$ and let $c \neq 0$. Then c^{-1} exists in F.

Now
$$\begin{aligned} c\alpha = \theta &\Rightarrow c^{-1}(c\alpha) = c^{-1}\theta \\ &\Rightarrow (c^{-1}c)\alpha = c^{-1}\theta, \text{ by } V_7 \\ &\Rightarrow 1\alpha = \theta, \text{ by (ii)} \\ &\Rightarrow \alpha = \theta, \text{ by } V_{10}. \end{aligned}$$

Therefore $c\alpha = \theta$ and $c \neq 0 \Rightarrow \alpha = \theta$.

Contrapositively, $c\alpha = \theta$ and $\alpha \neq \theta \Rightarrow c = 0$.

Hence $c\alpha = \theta$ implies either $c = 0$ or $\alpha = \theta$.

4.3. Sub-spaces.

Let V be a vector space over a field F with respect to addition (+) and multiplication by elements of F.

Let W be a non-empty subset of V. If W be stable under + and . then the restriction of + to $W \times W$ is a mapping from $W \times W$ to W and the restriction of . to $F \times W$ is a mapping from $F \times W$ to W. The restriction of +, say $\oplus$, is a composition on W and is defined by $\alpha \oplus \beta = \alpha + \beta$ for all $\alpha, \beta \in W$. The restriction of ., say $\odot$, is an external composition of F with W and is defined by $c \odot \alpha = c.\alpha$ for all $c \in F$ and all $\alpha \in W$.

If W forms a vector space over F with respect to $\oplus$ and $\odot$, then W is said to be a *sub-vector space* or a *linear subspace* or a *subspace* of V.

Theorem 4.3.1. A non-empty subset W of a vector space V over a field F is a subspace of V if and only if

(i) $\alpha \in W,\ \beta \in W \Rightarrow \alpha + \beta \in W$; and (ii) $\alpha \in W,\ c \in F \Rightarrow c\alpha \in W$.

Proof. Let the conditions hold in W.

Let $\alpha, \beta \in W$. Since F is a field, $-1 \in F$ where 1 is the identity element in F. By (ii) we have $-1\beta \in W$, i.e., $-\beta \in W$.

Then by (i) we have $\alpha + (-\beta) \in W$, i.e., $\alpha - \beta \in W$.

Thus $\alpha, \beta \in W \Rightarrow \alpha - \beta \in W$.

This proves that W is a subgroup of the additive group V. Since V is a commutative group, W is also a commutative subgroup of V.

Therefore the conditions **V1-V5** for a vector space are satisfied in W. **V6** is satisfied in W by(ii). The conditions **V7-V10** are satisfied in W, since they are hereditary properties. Thus W is by itself a vector space over F and so W is a subspace of V.

The necessity of the conditions (i) and (ii) follows from the definition of a vector space.

Note. The two conditions (i) and (ii) can also be expressed as the single condition— $a\alpha + b\beta \in W$ for all $\alpha, \beta \in W$ and all $a, b \in F$.

Examples.

1. Let V be a vector space over a field F. Then V itself is a subspace of V. This subspace is called the *improper subspace* of V.

The set consisting only of the null vector θ of V forms a subspace of V. This subspace is called the *trivial subspace* of V.

2. Let S be the subset of $\mathbb{R}^3$ defined by $S = \{(x, y, z) \in \mathbb{R}^3 : y = z = 0\}$. Then S is a non-empty subset of $\mathbb{R}^3$, since $(0,0,0) \in S$.

Let $\alpha = (x_1, 0, 0)$, $\beta = (x_2, 0, 0) \in S$; Then x_1, x_2 are real.

Let $c \in \mathbb{R}, d \in \mathbb{R}$. Then $c\alpha + d\beta = c(x_1, 0, 0) + d(x_2, 0, 0) = (cx_1 + dx_2, 0, 0) \in S$, since $cx_1 + dx_2 \in \mathbb{R}$.

This proves that S is a subspace of $\mathbb{R}^3$.

3. Let T be the subset defined by $T = \{(x, y, z) \in \mathbb{R}^3 : x = z = 0\}$. Then T is a subspace of $\mathbb{R}^3$.

4. Let U be the subset defined by $U = \{(x, y, z) \in \mathbb{R}^3 : x = y = 0\}$. Then U is a subspace of $\mathbb{R}^3$.

5. Let S be the subset defined by $S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = z^2\}$. Then S is a non-empty subset of $\mathbb{R}^3$, since $(0, 0, 0) \in S$.

Let $\alpha = (x_1, y_1, z_1) \in S$, $\beta = (x_2, y_2, z_2) \in S$. Then $x_i, y_i, z_i \in \mathbb{R}$ and $x_1^2 + y_1^2 = z_1^2$, $x_2^2 + y_2^2 = z_2^2$.

$\alpha + \beta = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$. $\alpha + \beta$ may not belong to S because $(x_1 + x_2)^2 + (y_1 + y_2)^2$ may not be equal to $(z_1 + z_2)^2$.

For example, let $\alpha = (3, -4, 5)$; $\beta = (-3, 4, 5)$. Then $\alpha \in S$, $\beta \in S$ but $\alpha + \beta = (0, 0, 10) \notin S$.

So S is not a subspace of $\mathbb{R}^3$.

6. Let S be the subset defined by $S = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}$.

S is a non-empty subset of $\mathbb{R}^3$ since $(0, 0, 1) \in S$. But S does not contain the null vector $(0, 0, 0)$. S is not a subspace of $\mathbb{R}^3$, since every subspace W of a vector space V must contain the null vector θ of V.

Theorem 4.3.2. The intersection of two subspaces of a vector space V over a field F is a subspace of V.

Proof. Let W_1 and W_2 be two subspaces of V. $W_1 \cap W_2$ is not empty, because $\theta \in W_1 \cap W_2$.

Case 1. Let $W_1 \cap W_2 = \{\theta\}$. Then $W_1 \cap W_2$ is a subspace of V.

Case 2. Let $W_1 \cap W_2 \neq \{\theta\}$ and let $\alpha_1 \in W_1 \cap W_2$, $\alpha_2 \in W_1 \cap W_2$.

Then $\alpha_1, \alpha_2 \in W_1$ and $\alpha_1, \alpha_2 \in W_2$.

Since W_1 is a subspace of V, (i) $\alpha_1 + \alpha_2 \in W_1$ and (ii) $c\alpha_1 \in W_1$, c being a scalar in F.

Since W_2 is a subspace of V, (i) $\alpha_1 + \alpha_2 \in W_2$ and (ii) $c\alpha_1 \in W_2$, c being a scalar in F.

Therefore $\alpha_1 + \alpha_2 \in W_1 \cap W_2$ and $c\alpha_1 \in W_1 \cap W_2$ for all $c \in F$. This proves that $W_1 \cap W_2$ is a subspace of V.

Note 1. Since $W_1 \cap W_2$ is the largest subset contained in both of W_1 and W_2, $W_1 \cap W_2$ is the largest subspace contained in W_1 and W_2.

Note 2. The intersection of a family of subspaces of V is a subspace of V.

Note 3. The union of two subspaces of V is not, in general, a subspace of V.

For example, let us consider two subspaces S and T of the vector space $\mathbb{R}^3$, where $S = \{(x, y, z) \in \mathbb{R}^3 : y = 0, z = 0\}$, $T = \{(x, y, z) \in \mathbb{R}^3 : x = 0, z = 0\}$.

Let $\alpha = (1, 0, 0) \in S$, $\beta = (0, 1, 0) \in T$. Then $\alpha + \beta = (1, 1, 0)$.

Here $\alpha \in S \cup T$, $\beta \in S \cup T$ but $\alpha + \beta \notin S$, $\alpha + \beta \notin T$ and therefore $\alpha + \beta \notin S \cup T$.

Hence $S \cup T$ is not a subspace of $\mathbb{R}^3$.

Theorem 4.3.3. If U and W be subspaces of a vector space V over a field F, then the union $U \cup W$ is a subspace of V if and only if either $U \subset W$ or $W \subset U$.

Proof. Let $U \cup W$ be a subspace of V. We prove that either $U \subset W$ or $W \subset U$, i.e., either $U - W = \phi$ or $W - U = \phi$.

Let us assume that both $U - W \neq \phi$ and $W - U \neq \phi$. Then there exists a vector α such that $\alpha \in U$, but $\alpha \notin W$ and a vector β such that $\beta \in W$, but $\beta \notin U$.

$\alpha \in U \Rightarrow \alpha \in U \cup W$ and $\beta \in W \Rightarrow \beta \in U \cup W$.

Since $U \cup W$ is a subspace of V, $\alpha + \beta \in U \cup W$. This implies $\alpha + \beta \in U$ or $\alpha + \beta \in W$.

$\alpha + \beta \in U$ and $\alpha \in U \Rightarrow (\alpha + \beta) - \alpha \in U$, since U is a subspace
$\Rightarrow \beta \in U$, a contradiction.

$\alpha + \beta \in W$ and $\beta \in W \Rightarrow (\alpha + \beta) - \beta \in W$, since W is a subspace
$\Rightarrow \alpha \in W$, a contradiction.

Therefore $\alpha + \beta \notin U$, $\alpha + \beta \notin W$ and therefore $\alpha + \beta \notin U \cup W$.

So our assumption that both $U - W \neq \phi$ and $W - U \neq \phi$ is not tenable and therefore either $U - W = \phi$ or $W - U = \phi$, i.e., either $U \subset W$ or $W \subset U$.

Conversely, let U and W be subspaces of V such that either $U \subset W$ or $W \subset U$.

If $U \subset W$ then $U \cup W = W$ and it is a subspace of V.

If $W \subset U$ then $U \cup W = U$ and it is a subspace of V.

Therefore in any case, $U \cup W$ is a subspace of V.
This proves the theorem.

Note. A vector space V cannot be the union of two proper subspaces

Linear sum of two subspaces.

Let U and W be subspaces of a vector space V over a field F. Then the subset $\{u + w : u \in U, w \in W\}$ is said to be the *linear sum* of the subspaces U and W.

Let $\alpha \in U$. Then $\alpha = \alpha + \theta$, where $\alpha \in U, \theta \in W$. This shows that $\alpha \in U + W$. Therefore $U \subset U + W$.

Let $\alpha \in W$. Then $\alpha = \theta + \alpha$ where $\theta \in U, \alpha \in W$. This shows that $\alpha \in U + W$. Therefore $W \subset U + W$.

Theorem 4.3.4. Let U and W be subspaces of a vector space V over a field F. Then the linear sum $U + W$ is a subspace of V.

Proof. Let $S = U + W = \{u + w : u \in U, w \in W\}$.

$\theta \in U, \theta \in W \Rightarrow \theta \in S$ and therefore S is non-empty.

Let $\alpha_1, \alpha_2 \in S$. Then $\alpha_1 = u_1 + w_1$ for some $u_1 \in U, w_1 \in W$;
$\alpha_2 = u_2 + w_2$ for some $u_2 \in U, w_2 \in W$.

$\alpha_1 + \alpha_2 = (u_1 + u_2) + (w_1 + w_2) \in S$, since $u_1 + u_2 \in U$, $w_1 + w_2 \in W$.

Let c be a scalar in F.

Then $c\alpha_1 = c(u_1 + w_1) = cu_1 + cw_1 \in S$, since $cu_1 \in U$, $cw_1 \in W$.

Therefore $\alpha_1 \in S$, $\alpha_2 \in S \Rightarrow \alpha_1 + \alpha_2 \in S$; and $c \in F$, $\alpha_1 \in S \Rightarrow c\alpha_1 \in S$.

This proves that S is a subspace of V, i.e., $U + W$ is a subspace of V.

Theorem 4.3.5. The subspace $U + W$ is the smallest subspace of V containing the subspaces U and W.

Proof. Let P be any subspace of V containing the subspaces U and W.

Let α be an element of $U + W$.

Then $\alpha = u_1 + w_1$ for some $u_1 \in U$, $w_1 \in W$.

Since $U \subset P$, $u_1 \in P$ and since $W \subset P$, $w_1 \in P$.

Since P is a subspace of V and $u_1, w_1 \in P$, $u_1 + w_1 \in P$, i.e., $\alpha \in P$.

Thus $\alpha \in U + W \Rightarrow \alpha \in P$ and therefore $U + W \subset P$.

This proves that $U + W$ is the smallest subspace containing U and V.

Some important subspaces of a vector space.

1. Let V be a vector space over a field F and let $\alpha \in V$. Then the set $W = \{c\alpha : c \in F\}$ forms a subspace of V.

Case 1. Let $\alpha = \theta$. Then $W = \{\theta\}$ and W is a subspace of V.

Case 2. Let $\alpha \neq \theta$.

W is a non-empty subset of V, since $\alpha \in W$.

Let $\alpha_1, \alpha_2 \in W$. Then $\alpha_1 = c_1\alpha, \alpha_2 = c_2\alpha$ for some scalars c_1, c_2 in F.

$\alpha_1 + \alpha_2 = c_1\alpha + c_2\alpha = (c_1 + c_2)\alpha \in W$, since $c_1 + c_2 \in F$.

Therefore $\alpha_1 \in W, \alpha_2 \in W \Rightarrow \alpha_1 + \alpha_2 \in W \quad \ldots \quad$ (i)

Let p be a scalar in F.

Then $p\alpha_1 = p(c_1\alpha) = (pc_1)\alpha \in W$, since $pc_1 \in F$.

Therefore $p \in F, \alpha_1 \in W \Rightarrow p\alpha_1 \in W \quad \ldots \quad$ (ii)

From (i) and (ii) it follows that W is subspace of V.

Note. This subspace is said to be *generated by* the vector α and α is said to be a *generator* of the subspace W.

2. Let V be a vector space over a field F and let $\alpha, \beta \in V$. Then the set $W = \{c\alpha + d\beta : c \in F, d \in F\}$ forms a subspace of V.

W is a non-empty subset of V, since $\theta(= 0\alpha + 0\beta) \in W$.

Let $\alpha_1 = c_1\alpha + d_1\beta \in W, \alpha_2 = c_2\alpha + d_2\beta \in W$, where $c_1, d_1, c_2, d_2 \in F$.

$\alpha_1 + \alpha_2 = (c_1\alpha + d_1\beta) + (c_2\alpha + d_2\beta) = (c_1 + c_2)\alpha + (d_1 + d_2)\beta \in W$, since $c_1 + c_2 \in F, d_1 + d_2 \in F$.

Therefore $\alpha_1 \in W, \alpha_2 \in W \Rightarrow \alpha_1 + \alpha_2 \in W \quad \ldots \quad$ (i)

Let p be a scalar in F.

Then $p\alpha_1 = p(c_1\alpha + d_1\beta)$

$= (pc_1)\alpha + (pd_1)\beta \in W$, since $pc_1 \in F, pd_1 \in F$.

Therefore $p \in F, \alpha_1 \in W \Rightarrow p\alpha_1 \in W \quad \ldots \quad$ (ii)

From (i) and (ii) it follows that W is a subspace of V.

Note. This subspace is said to be generated by the vectors α, β. The set $\{\alpha, \beta\}$ is said to be a *generating set* of the subspace W.

3. Let V be a vector space over a field F and let $\alpha_1, \alpha_2, \ldots, \alpha_r \in V$. Then the set $W = \{c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r : c_i \in F\}$ forms a subspace of V.

The set of vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is said to be a *generating set* of the subspace W.

Definition. Let V be a vector space over a field F. Let $\alpha_1, \alpha_2, \ldots, \alpha_r \in V$. A vector β in V is said to be a *linear combination* of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ if β can be expressed as

$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$ for some scalars $c_1, c_2, \ldots, c_r$ in F.

Example.

Let V be a real vector space and $\alpha, \beta, \gamma \in V$. Then $\alpha + \beta + \gamma$, $\alpha + 2\beta + 3\gamma, 0\alpha + \beta + 0\gamma, 0\alpha + 0\beta + \gamma, 0\alpha + 0\beta + 0\gamma$ are linear combinations of α, β, γ.

Theorem 4.3.6. Let V be a vector space over a field F and let S be a non-empty finite subset of V. Then the set W of all linear combinations of the vectors in S forms a subspace of V and this is the smallest subspace containing the set S.

Proof. Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. Then $W = \{c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n : c_1, c_2, \ldots, c_n \in F\}$.

W is a non-empty subset of V, since $\alpha_1 \in W$.

Let $\alpha = r_1\alpha_1 + r_2\alpha_2 + \cdots + r_n\alpha_n \in W$, $\beta = s_1\alpha_1 + s_2\alpha_2 + \cdots + s_n\alpha_n \in W$. Then $r_1, r_2, \ldots, r_n \in F$, $s_1, s_2, \ldots, s_n \in F$.

$\alpha + \beta = (r_1\alpha_1 + r_2\alpha_2 + \cdots + r_n\alpha_n) + (s_1\alpha_1 + s_2\alpha_2 + \cdots + s_n\alpha_n)$

$= (r_1 + s_1)\alpha_1 + (r_2 + s_2)\alpha_2 + \cdots + (r_n + s_n)\alpha_n \in W$, since $r_i + s_i \in F$.

Therefore $\alpha \in W, \beta \in W \Rightarrow \alpha + \beta \in W \quad \ldots \quad$ (i)

Let $p \in F$.

Then $p\alpha = p(r_1\alpha_1 + r_2\alpha_2 + \cdots + r_n\alpha_n)$

$= (pr_1)\alpha_1 + (pr_2)\alpha_2 + \cdots + (pr_n)\alpha_n \in W$, since $pr_i \in F$.

Therefore $p \in F, \alpha \in W \Rightarrow p\alpha \in W \quad \ldots \quad$ (ii)

From (i) and (ii) it follows that W is a subspace of V.

Let P be any subspace of V containing the set S.

Let $\xi \in W$. Then $\xi = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$ for some scalars $x_i \in F$.

Since P is a subspace of V containing α_i and $x_i \in F$, it follows that $x_i\alpha_i \in P$ for $i = 1, 2, \ldots, n$.

Since P is a subspace and $x_1\alpha_1, x_2\alpha_2, \ldots, x_n\alpha_n \in P$, it follows that $x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \in P$, i.e., $\xi \in P$.

Thus $\xi \in W \Rightarrow \xi \in P$ and therefore $W \subset P$.

This proves that W is the smallest subspace containing S.

Note. The smallest subspace containing S is the intersection of all subspaces containing S.

Definition. The smallest subspace containing a finite set S of vectors of a vector space V is said to be the *linear span* of S and is denoted by $L(S)$. $L(S)$ is said to be generated (or spanned) by the set S and S is said to be the *generating set* (or *spanning set*) of $L(S)$.

Note. If S be a non-empty finite set $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$, then $L(S)$ is the set of all linear combinations of the vectors $\alpha_1, \alpha_2, \dots, \alpha_n$.

If $S = \phi$, then $L(S) = \{\theta\}$, since $\{\theta\}$ is the intersection of all subspaces containing ϕ.

Theorem 4.3.7. If S and T be two non-empty finite subsets of a vector space V over a field F and $S \subset T$, then $L(S) \subset L(T)$.

Proof. Let $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and let $\xi \in L(S)$.

Then $\xi = r_1\alpha_1 + r_2\alpha_2 + \dots + r_n\alpha_n$ for some scalars $r_i \in F$.

$\alpha_1 \in S \Rightarrow \alpha_1 \in T$ and therefore $\alpha_1 \in L(T)$.

Since $L(T)$ is a subspace of $V, r_1 \in F$, $\alpha_1 \in L(T) \Rightarrow r_1\alpha_1 \in L(T)$.

Similarly, $r_2\alpha_2 \in L(T), \dots, r_n\alpha_n \in L(T)$.

Since $L(T)$ is a subspace, it also follows that

$r_1\alpha_1 + r_2\alpha_2 + \dots + r_n\alpha_n \in L(T)$, i.e., $\xi \in L(T)$.

Thus $\xi \in L(S) \Rightarrow \xi \in L(T)$ and therefore $L(S) \subset L(T)$.

Theorem 4.3.8. If S and T be two non-empty finite subsets of a vector space V over a field F and each element of T is a linear combination of the vectors of S, then $L(T) \subset L(S)$.

Proof. Let $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $T = \{\beta_1, \beta_2, \dots, \beta_m\}$ and let

$\beta_i = c_{i1}\alpha_1 + c_{i2}\alpha_2 + \dots + c_{in}\alpha_n$ for some $c_{ij} \in F$, $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

Let ξ be an element of $L(T)$.

Then $\xi = p_1\beta_1 + p_2\beta_2 + \dots + p_m\beta_m$ for some scalars $p_i \in F$.

$$\begin{aligned}\xi &= p_1(c_{11}\alpha_1 + c_{12}\alpha_2 + \dots + c_{1n}\alpha_n) + p_2(c_{21}\alpha_1 + c_{22}\alpha_2 + \dots + c_{2n}\alpha_n) \\ &\quad + \dots + p_m(c_{m1}\alpha_1 + c_{m2}\alpha_2 + \dots + c_{mn}\alpha_n) \\ &= d_1\alpha_1 + d_2\alpha_2 + \dots + d_n\alpha_n \in L(S), \text{ since } d_i = p_1c_{1i} + p_2c_{2i} + \dots \\ &\qquad + p_mc_{mi} \in F; i = 1, 2, \dots, n.\end{aligned}$$

Thus $\xi \in L(T) \Rightarrow \xi \in L(S)$ and therefore $L(T) \subset L(S)$.

Note. The theorem says that if $T \subset L(S)$ then $L(T) \subset L(S)$.

Worked Examples.

1. In $\mathbb{R}^3$, $\alpha = (4,3,5)$, $\beta = (0,1,3)$, $\gamma = (2,1,1)$, $\delta = (4,2,2)$.

Examine if (i) α is a linear combination of β and γ,

(ii) β is a linear combination of γ and δ.

(i) Let $\alpha = c\beta + d\gamma$, where $c, d \in \mathbb{R}$.

Then $(4,3,5) = c(0,1,3) + d(2,1,1) = (0 + 2d, c + d, 3c + d)$.

Therefore $2d = 4, c + d = 3, 3c + d = 5$ giving $c = 1, d = 2$.

Hence $\alpha = \beta + 2\gamma$ and α is a linear combination of β and γ.

(ii) Let $\beta = c\gamma + d\delta$ where $c, d \in \mathbb{R}$.

Then $(0,1,3) = c(2,1,1) + d(4,2,2) = (2c + 4d, c + 2d, c + 2d)$.

Therefore $2c + 4d = 0, c + 2d = 1, c + 2d = 3$.

The equations are inconsistent. Therefore β cannot be expressed as $c\gamma + d\delta$ for real c, d. Hence β is not a linear combination of γ and δ.

2. Determine the subspace of $\mathbb{R}^3$ spanned by the vectors $\alpha = (1,2,3)$ and $\beta = (3,1,0)$. Examine if

(i) $\gamma = (2,1,3)$ is in the subspace,

(ii) $\delta = (-1,3,6)$ is in the subspace.

$L\{\alpha, \beta\}$ is the set of vectors $\{c\alpha + d\beta\} : c \in \mathbb{R}, d \in \mathbb{R}\}$.

$c\alpha + d\beta = c(1,2,3) + d(3,1,0) = (c + 3d, 2c + d, 3c)$.

(i) If $\gamma \in L\{\alpha, \beta\}$ then there must be real numbers c, d such that $(2,1,3) = (c + 3d, 2c + d, 3c)$.

Therefore $c + 3d = 2, 2c + d = 1$ and $3c = 3$.

These equations are inconsistent and so γ is not in $L\{\alpha, \beta\}$.

(ii) If $\delta \in L\{\alpha, \beta\}$, then there must be real numbers c, d such that $(-1,3,6) = (c + 3d, 2c + d, 3c)$

Therefore $c + 3d = -1, 2c + d = 3$ and $3c = 6$, giving $c = 2, d = -1$.

Therefore $\delta = 2(1,2,3) - 1(3,1,0)$, showing that $\delta \in L\{\alpha, \beta\}$.

3. Let $S = \{\alpha, \beta, \gamma\}$, $T = \{\alpha, \beta, \alpha + \beta, \beta + \gamma\}$ be subsets of a real vector space V. Show that $L(S) = L(T)$.

S and T are finite subsets of V and each element of T is a linear combination of the vectors of S and therefore $L(T) \subset L(S)$.

Again $\alpha = \alpha + 0\beta + 0(\alpha + \beta) + 0(\beta + \gamma)$,

$\beta = 0\alpha + \beta + 0(\alpha + \beta) + 0(\beta + \gamma)$

and $\gamma = 0\alpha - \beta + 0(\alpha + \beta) + (\beta + \gamma)$.

This shows that each element of S is a linear combination of the vectors of T and therefore $L(S) \subset L(T)$.

It follows that $L(S) = L(T)$.

Definition. Linear span of a set.

Let S be a non-empty subset of a vector space V over a field F. The set of all *finite* linear combinations of the elements of S is said to be the *linear span* of S and it is denoted by $L(S)$.

Theorem 4.3.9. Let S be a non-empty subset of a vector space V over a field F. Then $L(S)$ is the smallest subspace of V containing the set S.

Proof. Let $\alpha \in S$. Then $\alpha(= 1\alpha)$ being a linear combination of one (a finite number) element of S, belongs to $L(S)$.

$\alpha \in S \Rightarrow \alpha \in L(S)$ and therefore $S \subset L(S)$...(i)

Let $\xi \in L(S)$, $\eta \in L(S)$. Then
$\xi = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$, $\alpha_i \in S$, $c_i \in F$;
$\eta = d_1\beta_1 + d_2\beta_2 + \cdots + d_k\beta_k$, $\beta_i \in S$, $d_i \in F$.

Let $c \in F$, $d \in F$. Then
$$c\xi + d\eta = c(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r) + d(d_1\beta_1 + d_2\beta_2 + \cdots + d_k\beta_k)$$
$$= p_1\alpha_1 + p_2\alpha_2 + \cdots + p_r\alpha_r + p_{r+1}\beta_1 + p_{r+2}\beta_2 + \cdots + p_{r+k}\beta_k,$$
where $p_i = cc_i(i = 1, 2, .., r) \in F$, $p_i = dd_{i-r}(i = r+1, ..., r+k) \in F$.

$c\xi + d\eta$ is a finite linear combination of the elements of S and therefore belongs to $L(S)$.

$\xi \in L(S), \eta \in L(S) \Rightarrow c\xi + d\eta \in L(S)$ for all $c, d \in F$...(ii)

From (i) and (ii) it follows that $L(S)$ is a subspace of V containing the set S.

Let P be any subspace of V containing the set S.

Let $\xi \in L(S)$. Then $\xi = r_1\alpha_1 + r_2\alpha_2 + \cdots + r_m\alpha_m$, some finite linear combination of the elements of S.

$\alpha_i \in S \Rightarrow \alpha_i \in P$, $i = 1, 2, \ldots, m$.

Since P is a subspace, $r_i \in F, \alpha_i \in P \Rightarrow r_i\alpha_i \in P$ for $i = 1, 2, ..., m$. Therefore $\xi \in P$, since P is a subspace.

Thus $\xi \in L(S) \Rightarrow \xi \in P$ and therefore $L(S) \subset P$. This proves that $L(S)$ is the smallest subspace of V containing the set S.

Note. Since $L(S)$ is the smallest subspace containing the set S, it is the intersection of all subspaces of V containing the set S.

Theorem 4.3.10. If S be a non-empty subset of a vector space V over a field F then $L(L(S)) = L(S)$.

Proof. Since $L(L(S))$ is the smallest subspace containing $L(S)$ and $L(S)$ is itself a subspace, $L(L(S)) = L(S)$.

Theorem 4.3.11. Let S and T be non-empty subsets of a vector space V over a field F. Then $L(S \cup T) = L(S) + L(T)$.

Proof. $L(S \cup T)$ is the set of all finite linear combinations of the elements of $S \cup T$.

Let $\xi \in L(S) + L(T)$. Then $\xi = \sigma + \tau$ for some $\sigma \in L(S)$, some $\tau \in L(T)$.

$\sigma \in L(S) \Rightarrow \sigma = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$, some finite linear combination of the vectors of S.

$\tau \in L(T) \Rightarrow \tau = d_1\beta_1 + d_2\beta_2 + \cdots + d_k\beta_k$, some finite linear combination of the vectors of T.

Therefore ξ is a finite linear combination of the vectors of $S \cup T$ and so $\xi \in L(S \cup T)$.

Thus $\xi \in L(S) + L(T) \Rightarrow \xi \in L(S \cup T)$ and therefore $L(S) + L(T) \subset L(S \cup T)$ ... ... (i)

$S \subset L(S) \Rightarrow S \subset L(S) + L(T)$ and $T \subset L(T) \Rightarrow T \subset L(S) + L(T)$. Therefore $S \cup T \subset L(S) + L(T)$.

$L(S \cup T)$ being the smallest subspace of V containing $S \cup T$, $L(S \cup T) \subset L(S) + L(T)$ ... ... (ii)

From (i) and (ii) it follows that $L(S \cup T) = L(S) + L(T)$.

Note. In particular, if S and T be subspaces of V then $L(S) = S$ and $L(T) = T$. Therefore if S and T be subspaces of V then $L(S \cup T) = S + T$, by the theorem.

Exercises 6

1. Prove that in a real vector space V
   (i) $c(\alpha - \beta) = c\alpha - c\beta$, where $c \in \mathbb{R}$; $\alpha, \beta \in V$;
   (ii) $(c - d)\alpha = c\alpha - d\alpha$, where $c, d \in \mathbb{R}$; $\alpha \in V$;
   (iii) $(-c)\alpha = -(c\alpha) = c(-\alpha)$, where $c, d \in \mathbb{R}$; $\alpha, \beta \in V$.
2. Examine if the set S is a subspace of $\mathbb{R}^3$.
   (i) $S = \{(x, y, z) \in \mathbb{R}^3 : x = 0\}$;
   (ii) $S = \{(x, y, z) \in \mathbb{R}^3 : x = 1\}$;
   (iii) $S = \{(x, y, z) \in \mathbb{R}^3 : xy = z\}$;
   (iv) $S = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$;
   (v) $S = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 1\}$;
   (vi) $S = \{(x, y, z) \in \mathbb{R}^3 : x + 2y - z = 0, 2x - y + z = 0\}$;
   (vii) $S = \{(x, y, z) \in \mathbb{R}^3 : x + 2y - z = 1, 2x - y + z = 2\}$.
3. If $\alpha = (1, 1, 2), \beta = (0, 2, 1), \gamma = (2, 2, 4)$, determine whether
   (i) α is a linear combination of β and γ,

(ii) β is a linear combination of γ and α,

(iii) γ is a linear combination of α and β.

4. In $\mathbb{R}^3$, $\alpha = (1,3,0), \beta = (2,1,-2)$. Determine $L\{\alpha,\beta\}$. Examine if $\gamma = (-1,3,2), \delta = (4,7,-2)$ are in $L\{\alpha,\beta\}$.

5. Let $\alpha_1,\alpha_2,\alpha_3$ are vectors in a real vector space V such that $\alpha_1+\alpha_2+\alpha_3=\theta$. Prove that $L\{\alpha_1,\alpha_2\} = L\{\alpha_2,\alpha_3\} = L\{\alpha_3,\alpha_1\}$.

6. Let $S = \{\alpha,\beta,\gamma\}$, $T = \{\alpha,\alpha+\beta,\alpha+\beta+\gamma\}$, $P = \{\alpha+\beta,\beta+\gamma,\gamma+\alpha\}$ be subsets in a real vector space V. Prove that $L(S) = L(T) = L(P)$.

7. If $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ are vectors in a real vector space V such that $c_1\alpha_1+c_2\alpha_2+c_3\alpha_3+c_4\alpha_4=\theta$, where c_1,c_2,c_3,c_4 are real numbers with $c_1c_4 \neq 0$. Prove that $L\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\} = L\{\alpha_2,\alpha_3,\alpha_4\} = L\{\alpha_1,\alpha_2,\alpha_3\}$.

8. Prove that the set $C[a,b]$ of all real valued continuous functions defined on the closed interval $[a,b]$ forms a real vector space if

(i) addition is defined by $(f+g)(x) = f(x)+g(x), f,g \in C[a,b]$,

(ii) multiplication by a real number r is defined by $(r.f)(x) = rf(x), f \in C[a,b]$.

9. Prove that the subset $D[a,b]$ of all real valued differentiable functions defined on $[a,b]$ is a subspace of $C[a,b]$.

10. Examine if the set S is a subspace of the vector space $\mathbb{R}_{2\times 2}$, where

(i) $S = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}_{2\times 2} : a+b=0\right\}$;

(ii) $S = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}_{2\times 2} : a+b+c+d=0\right\}$;

(iii) $S = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}_{2\times 2} : \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 0\right\}$;

(iv) S is the set of all 2×2 real diagonal matrices;

(v) S is the set of all 2×2 real symmetric matrices;

(vi) S is the set of all 2×2 real skew symmetric matrices;

(vii) S is the set of all 2×2 real upper triangular matrices;

(viii) S is the set of all 2×2 real lower triangular matrices.

11. Show that the set S is a subspace of the vector space $C[0,1]$, where

(i) $S = \{f \in C[0,1] : f(0) = 0\}$; (ii) $S = \{f \in C[0,1] : f(0) = 0, f(1) = 0\}$.

12. A real vector space V is the union of three subspaces P,Q,R of V. Prove that one of P,Q,R is V.

4.4. Linear dependence and Linear independence.

A *finite set* of vectors $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$ of a vector space V over a field F is said to be *linearly dependent* in V if there exist scalars $c_1,c_2,\ldots,c_n$, not all zero, in F such that

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \theta \quad \cdots \quad \text{(i)}$$

The set is said to be *linearly independent* in V if the equality (i) is satisfied only when $c_1 = c_2 = \cdots = c_n = 0$.

An *arbitrary set* S of vectors of a vector space V over a field F is said to be linearly dependent in V if there exists a *finite* subset of S which is linearly dependent in V.

A set of vectors which is not linearly dependent is said to be linearly independent.

Although the definition of linear dependence or linear independence refers to a set of vectors, we shall also state that the individual vectors $\alpha_1,\alpha_2,\ldots,\alpha_n$ are linearly dependent or independent.

Theorem 4.4.1. A superset of a linearly dependent set of vectors in a vector space V over a field F is linearly dependent.

Proof. **Case (i).** Let S be a linearly dependent set of vectors containing a finite number of elements $\alpha_1,\alpha_2,\ldots,\alpha_n$. Let T be a superset of S. S being a finite subset of T and being linearly dependent, T is linearly dependent, by definition.

Case (ii). Let S be a linearly dependent set of vectors containing an infinite number of elements and T be a superset of S. Since S is linearly dependent, there exists a finite subset P of S such that P is linearly dependent. P being a linearly dependent finite subset of T, T is linearly dependent, by definition.

This completes the proof.

Theorem 4.4.2. A subset of a linearly independent set of vectors in a vector space V over a field F is linearly independent.

Proof. Let S be a linearly independent set of vectors and P be a subset of S. If P be linearly dependent then S being a superset of the linearly dependent set P, must be linearly dependent, by Theorem 4.4.1. But it is not so. This proves that P is linearly independent.

Note. The set ϕ is linearly independent.

Theorem 4.4.3. A set of vectors containing the null vector θ in a vector space V over a field F is linearly dependent.

Proof. Let $S = \{\theta\}$. The set S is linearly dependent since the relation $c\theta = \theta$ holds for a non-zero scalar c.

Let T be an arbitrary set of vectors containing the null vector θ. Then T being a superset of S, is linearly dependent by Theorem 4.4.1.

Theorem 4.4.4. The set consisting of a single non-zero vector α in a vector space V over a field F is linearly independent.

The proof follows from the Theorem 4.2.1.(iv).

Worked Examples.

1. Examine if the set of vectors $\{(2,1,1)(1,2,2),(1,1,1)\}$ is linearly dependent in $\mathbb{R}^3$.

Let $\alpha = (2,1,1)$, $\beta = (1,2,2)$, $\gamma = (1,1,1)$.

Let us consider the relation $c_1\alpha + c_2\beta + c_2\gamma = \theta$, where c_1, c_2, c_3 are real numbers.

Then $c_1(2,1,1) + c_2(1,2,2) + c_3(1,1,1) = (0,0,0)$.

Therefore $2c_1 + c_2 + c_3 = 0,\quad c_1 + 2c_2 + c_3 = 0,\quad c_1 + 2c_2 + c_3 = 0$.

or equivalently, $2c_1 + c_2 + c_3 = 0,\quad c_1 + 2c_2 + c_3 = 0$.

The solution is $c_1 = -k,\ c_2 = -k,\ c_3 = 3k$, where k is a real number.

Since k is arbitrary, there exist c_1, c_2, c_3, not all zero, such that $c_1\alpha + c_2\beta + c_3\gamma = \theta$. For example, $c_1 = 1, c_2 = 1, c_3 = -3$.

Therefore the given set of vectors is linearly dependent in $\mathbb{R}^3$.

2. Prove that the set of vectors $\{(1,2,2),(2,1,2),(2,2,1)\}$ is linearly independent in $\mathbb{R}^3$.

Let $\alpha = (1,2,2)$, $\beta = (2,1,2)$, $\gamma = (2,2,1)$.

Let us consider the relation $c_1\alpha + c_2\beta + c_3\gamma = \theta$, where c_1, c_2, c_3 are real numbers.

Then $c_1(1,2,2) + c_2(2,1,2) + c_3(2,2,1) = (0,0,0)$.

Therefore $c_1 + 2c_2 + 2c_3 = 0,\quad 2c_1 + c_2 + 2c_3 = 0,\quad 2c_1 + 2c_2 + c_3 = 0$.

This is a homogeneous system of three equations in c_1, c_2, c_3.

The co-efficient determinant of the system is $\begin{vmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{vmatrix} = 5 \neq 0$.

By Cramer's rule, there exists a unique solution for c_1, c_2, c_3 and the solution is $c_1 = 0, c_2 = 0, c_3 = 0$.

This proves that the given set of vectors is linearly independent in $\mathbb{R}^3$.

Theorem 4.4.5. If the set of vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ in a vector space V over a field F be linearly dependent, then at least one of the vectors of the set can be expressed as a linear combination of the remaining others.

Conversely, if one of the vectors of the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linear combination of the remaining others, the set is linearly dependent.

Proof. Since the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is linearly dependent, there exist scalars $c_1, c_2, \ldots, c_n$ in F, not all zero, such that

$$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \theta.$$

Let $c_j \neq 0$. Then $c_j^{-1} \in F$ and $c_j^{-1}c_j = 1$, 1 being the identity element in F.

$$c_j\alpha_j = -c_1\alpha_1 - c_2\alpha_2 - \cdots - c_{j-1}\alpha_{j-1} - c_{j+1}\alpha_{j+1} - \cdots - c_n\alpha_n.$$

Therefore
$$\begin{aligned}\alpha_j &= c_j^{-1}[-c_1\alpha_1 - c_2\alpha_2 - \cdots - c_{j-1}\alpha_{j-1} - c_{j+1}\alpha_{j+1} - \cdots - c_n\alpha_n] \\ &= d_1\alpha_1 + d_2\alpha_2 + \cdots + d_{j-1}\alpha_{j-1} + d_{j+1}\alpha_{j+1} + \cdots + d_n\alpha_n,\end{aligned}$$
where $d_i = -c_j^{-1}c_i \in F$, $i = 1, 2, \ldots, j-1, j+1, \ldots, n$.

This shows that α_j is a linear combination of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_n$.

Conversely, let one of the vectors, say α_j, is a linear combination of the other vectors of the set.

Then $\alpha_j = r_1\alpha_1 + r_2\alpha_2 + \cdots + r_{j-1}\alpha_{j-1} + r_{j+1}\alpha_{j+1} + \cdots + r_n\alpha_n$, for some scalars $r_i \in F$, $i = 1, 2, \ldots, j-1, j+1, \ldots, n$.

Therefore $r_1\alpha_1 + r_2\alpha_2 + \cdots + r_{j-1}\alpha_{j-1} + (-1)\alpha_j + r_{j+1}\alpha_{j+1} + \cdots + r_n\alpha_n = \theta$.

Since the above equality holds for scalars $r_1, r_2, \ldots, r_{j-1}, -1, r_{j+1}, \ldots, r_n$ in F and since one of them at least is non-zero, the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is linearly dependent.

This completes the proof.

Corollary. Two vectors α, β in a vector space V over a field F are linearly dependent if at least one of them is a scalar multiple of the other.

Examples.

1. The set of vectors $S = \{\alpha, 2\alpha, \beta\}$ of a real vector space V is linearly dependent, since $2\alpha \in S$ and $2\alpha (= 2\alpha + 0\beta)$ is a linear combination of the remaining vectors of S.

2. The set of vectors $S = \{\alpha, \beta, \gamma, \beta + \gamma\}$ of a real vector space V is linearly dependent, since $\beta + \gamma \in S$ and $\beta + \gamma (= 0\alpha + \beta + \gamma)$ is a linear combination of the remaining vectors of S.

3. The set of vectors $S = \{\alpha, \beta, \gamma, \theta\}$ of a real vector space V is linearly dependent, since $\theta \in S$ and $\theta(= 0\alpha + 0\beta + 0\gamma)$ is a linear combination of the remaining vectors of S.

Note. The set S in Ex.3 is also linearly dependent by Theorem 4.4.3.

Alternative Criterion.

Theorem 4.4.6. A set of vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}(\alpha_i \neq \theta, i = 1, 2, \ldots, n)$ in a vector space V over a field F is linearly dependent if and only if there exists a vector in the set which is a linear combination of the preceding others.

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be linearly dependent. Then there exist scalars $c_1, c_2, \ldots, c_n$, not all zero, in F such that

$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \theta$.

Let i be the largest subscript such that $c_i \neq 0$. Then $i \neq 1$.

We have $c_i\alpha_i = -c_1\alpha_1 - c_2\alpha_2 - \cdots - c_{i-1}\alpha_{i-1}$.

Since $c_i \neq 0, c_i^{-1} \in F$ and therefore

$\alpha_i = -c_i^{-1}c_1\alpha_1 - c_i^{-1}c_2\alpha_2 - \cdots - c_i^{-1}c_{i-1}\alpha_{i-1}$

i.e., α_i is a linear combination of the preceding vectors of the set.

Conversely, let α_j be a linear combination of the preceding vectors of the set. Then $\alpha_j = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_{j-1}\alpha_{j-1}$ for some scalars d_i in F.

or, $d_1\alpha_1 + d_2\alpha_2 + \cdots + d_{j-1}\alpha_{j-1} - \alpha_j + 0\alpha_{j+1} + \cdots + 0\alpha_n = \theta$.

Since the above equality holds for scalars $d_1, d_2, \ldots, d_{j-1}, -1$ and 0 in F and one of them at least is non-zero, the set of vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is linearly dependent.

This completes the proof.

Theorem 4.4.7. Deletion theorem.

If a non-null vector space V over a field F be spanned by a linearly dependent set $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, then V can also be spanned by a suitable proper subset of S. [That is, some vector can be deleted from a linearly dependent spanning set of V].

Proof. Since S is a linearly dependent set, at least one of the vectors of the set, say α_j, can be expressed as a linear combination of the remaining others.

Let $\alpha_j = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_{j-1}\alpha_{j-1} + d_{j+1}\alpha_{j+1} + \cdots + d_n\alpha_n$, for some scalars $d_i \in F$, $i = 1, 2, \ldots, j-1, j+1, \ldots, n$. ... (i)

Here $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ and $L(S) = V$.

Let $T = \{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \alpha_{j+1}, \ldots, \alpha_n\}$.

As T is a subset of S, $L(T) \subset L(S)$, by Theorem 4.3.7.

Using (i) we see that each element of S is a linear combination of the vectors of T.

Therefore $L(S) \subset L(T)$, by Theorem 4.3.8.

Hence $L(T) = L(S) = V$. Thus V is spanned by a proper subset T of S. This completes the proof.

Note. If V be the null space then $S = \{\theta\}$ is a generating set of V. V can also be considered as $L\{\phi\}$ and the set ϕ is an improper subset of S.

Theorem 4.4.8. If $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linearly independent set of generators of a vector space V, then no proper subset of S can be a spanning set of V.

Proof. If possible, let a proper subset T of S be a spanning set of V. Let some α_j belong to $S - T$.

Since $L(T) = V$ and $\alpha_j \in V$, α_j can be expressed as a linear combination of the vectors of T. Therefore the set $T \cup \{\alpha_j\}$ is linearly dependent, by Theorem 4.4.5.

The set S being a superset of $T \cup \{\alpha_j\}$ becomes a linearly dependent set, and this is a contradiction which proves the theorem.

Worked Examples (continued).

3. Let $\alpha_1 = (1, 2, 0), \alpha_2 = (3, -1, 1), \alpha_3 = (4, 1, 1)$. Show that the set $S = \{\alpha_1, \alpha_2, \alpha_3\}$ is linearly dependent. Apply Deletion Theorem to find a proper subset of S that can generate $L(S)$.

Let $c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = \theta$, where c_1, c_2, c_3 are real numbers.

Then $c_1(1, 2, 0) + c_2(3, -1, 1) + c_3(4, 1, 1) = (0, 0, 0)$.

Therefore $c_1 + 3c_2 + 4c_3 = 0$, $2c_1 - c_2 + c_3 = 0$, $c_2 + c_3 = 0$.

This gives $c_1 = c_2 = -c_3$. Taking $c_1 = 1$, we have $c_2 = 1, c_3 = -1$ and then $\alpha_1 + \alpha_2 - \alpha_3 = \theta$.

This shows that the set S is linearly dependent.

As $\alpha_3 = \alpha_1 + \alpha_2$, α_3 can be deleted from the generating set of $L(S)$, by Deletion theorem. That is, $L\{\alpha_1, \alpha_2\} = L\{\alpha_1, \alpha_2, \alpha_3\}$.

Note. Since $\alpha_1 = -\alpha_2 + \alpha_3$, α_1 can also be deleted by the theorem and $L\{\alpha_2, \alpha_3\} = L\{\alpha_1, \alpha_2, \alpha_3\}$. Similarly, $L\{\alpha_1, \alpha_3\} = L\{\alpha_1, \alpha_2, \alpha_3\}$.

4.5. Basis and dimension of a vector space.

Let V be a vector space over a field F. V is said to be *finitely generated*, or *finite dimensional* if there exists a finite set of vectors in V generating V. Otherwise, V is said to be *infinite dimensional*.

The null space $\{\theta\}$ is finite dimensional, since it is generated by the empty set ϕ.

We shall be mainly concerned with finite dimensional vector spaces in this treatise.

Basis of a vector space.

Definition. Let V be a vector space over a field F. A set S of vectors in V is said to be a *basis* of V if

(i) S is linearly independent in V, and

(ii) S generates V.

Theorem 4.5.1. There exists a basis for every finitely generated vector space.

Proof. **Case 1.** Let V be a finitely generated vector space other than the null space.

Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a finite set of generators of V. If S is linearly independent in V, then S itself is a basis of V and the theorem is done.

If S is linearly dependent, then we can delete, by Deletion theorem, at least one vector from S and obtain a proper subset S_1 (of S) spanning the same space V. If S_1 is linearly independent in V, then S_1 is a basis of V and the theorem is done.

If S_1 be not linearly independent then we repeat the process of deletion and finally obtain, after $k(< n)$ steps of deletion, a subset S_k which is linearly independent in V and which also spans V.

This is possible, because S is a finite set of n elements and in the extreme unfavourable case we can come, after $n - 1$ steps of deletion, to a subset S_{n-1} containing only one non-zero vector that generates V and that is linearly independent.

Therefore our assertion that S_k is a linearly independent set for some $k(< n)$ is true and hence it is a basis of V.

Case 2. Let $V = \{\theta\}$. Since the set ϕ is linearly independent and $L(\phi) = \{\theta\}$, ϕ is a basis of V.
This completes the proof.

Examples.

1. The set $E = \{\epsilon_1 = (1, 0), \epsilon_2 = (0, 1)\}$ is a basis of the vector space $\mathbb{R}^2$.

To prove linear independence of the set E, let us consider the relation $c_1\epsilon_1 + c_2\epsilon_2 = \theta$ for real c_1, c_2.

Then $c_1(1, 0) + c_2(0, 1) = (0, 0)$.
This gives $c_1 = 0, c_2 = 0$ and this proves that the set E is linearly independent.

Let $\xi = (a, b)$ be an arbitrary vector of $\mathbb{R}^2$. Then $a, b \in \mathbb{R}$ and ξ can be expressed as $a(1, 0) + b(0, 1) = a\epsilon_1 + b\epsilon_2$.
This shows that $\xi \in L(E)$ and therefore $\mathbb{R}^2 \subset L(E)$.

Again $E \subset \mathbb{R}^2$ and $L(E)$ being the smallest subspace of $\mathbb{R}^2$ containing E, $L(E) \subset \mathbb{R}^2$. It follows that $L(E) = \mathbb{R}^2$.
Thus the set E fulfils both the conditions for a basis of $\mathbb{R}^2$.

Note. This basis E is said to be the *standard basis* of $\mathbb{R}^2$.

2. The set $E = \{\epsilon_1 = (1, 0, 0), \epsilon_2 = (0, 1, 0), \epsilon_3 = (0, 0, 1)\}$ is a basis of the vector space $\mathbb{R}^3$. E is the standard basis of $\mathbb{R}^3$.

3. The set $E = \{\epsilon_1 = (1, 0, 0, \ldots, 0), \epsilon_2 = (0, 1, 0, \ldots, 0), \ldots, \epsilon_n = (0, 0, 0, \ldots, 1)\}$ is a basis of $\mathbb{R}^n$. E is the standard basis of $\mathbb{R}^n$.

Worked Example.

1. Prove that the set $S = \{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ is a basis of $\mathbb{R}^3$.

Let $\alpha_1 = (0, 1, 1), \alpha_2 = (1, 0, 1), \alpha_3 = (1, 1, 0)$.
To examine linear independence of the set, let us consider the relation $c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = \theta$, where c_1, c_2, c_3 are real numbers.

Then $c_1(0, 1, 1) + c_2(1, 0, 1) + c_3(1, 1, 0) = (0, 0, 0)$.
This gives $c_2 + c_3 = 0, c_1 + c_3 = 0, c_1 + c_2 = 0$.
Therefore $c_1 = c_2 = c_3 = 0$ and this proves that the set S is linearly independent.

Let $\xi = (a, b, c)$ be an arbitrary vector of $\mathbb{R}^3$. Let us examine if $\xi \in L(S)$.
If possible, let $\xi = r_1\alpha_1 + r_2\alpha_2 + r_3\alpha_3$ for real r_1, r_2, r_3.
Then $r_2 + r_3 = a, r_1 + r_3 = b, r_1 + r_2 = c$.
This is a non-homogeneous system of three equations in r_1, r_2, r_3.

The co-efficient determinant of the system is $\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix} = 2 \neq 0$.

By Cramer's rule, there exists a unique solution for r_1, r_2, r_3.
This proves that $\xi \in L(S)$ and therefore $\mathbb{R}^3 \subset L(S)$.

Again $S \subset \mathbb{R}^3$ and $L(S)$ being the smallest subspace of $\mathbb{R}^3$ containing S, $L(S) \subset \mathbb{R}^3$. Consequently, $L(S) = \mathbb{R}^3$.
Thus the set S fulfils both the conditions for a basis of $\mathbb{R}^3$.

Theorem 4.5.2. Replacement theorem.

If $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of a vector space V over a field F and a non-zero vector β of V is expressed as $\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n, c_i \in F$, then if $c_j \neq 0, \{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\}$ is a new basis of V. [That is, β can replace α_j in the basis.]

Proof. $\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_{j-1}\alpha_{j-1} + c_j\alpha_j + c_{j+1}\alpha_{j+1} + \cdots + c_n\alpha_n$

or, $c_j\alpha_j = -c_1\alpha_1 - c_2\alpha_2 - \cdots - c_{j-1}\alpha_{j-1} + \beta - c_{j+1}\alpha_{j+1} - \cdots - c_n\alpha_n$.

Since $c_j \neq 0, c_j^{-1}$ exists in F and $c_j^{-1}c_j = 1, 1$ being the identity element in F. Therefore

$$\begin{aligned} \alpha_j &= -c_j^{-1}c_1\alpha_1 - c_j^{-1}c_2\alpha_2 - \cdots - c_j^{-1}c_{j-1}\alpha_{j-1} + c_j^{-1}\beta \\ &\quad - c_j^{-1}c_{j+1}\alpha_{j+1} - \cdots - c_j^{-1}c_n\alpha_n \\ &= p_1\alpha_1 + \cdots + p_{j-1}\alpha_{j-1} + p_j\beta + p_{j+1}\alpha_{j+1} + \cdots + p_n\alpha_n, \end{aligned}$$

where $p_i = -c_j^{-1}c_i$, if $i = 1, 2, \ldots, j-1, j+1, \ldots, n$; and $p_j = c_j^{-1}$.

So $p_i \in F$ for $i = 1, 2, \ldots, n$ and therefore α_j is a linear combination of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n$.

To prove that the set $\{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\}$ is linearly independent, let us consider the relation $d_1\alpha_1 + d_2\alpha_2 + \cdots + d_{j-1}\alpha_{j-1} + d_j\beta + d_{j+1}\alpha_{j+1} + \cdots + d_n\alpha_n = \theta$, where $d_i \in F,\ i = 1, 2, \ldots, n$.

Then $d_1\alpha_1 + d_2\alpha_2 + \cdots + d_{j-1}\alpha_{j-1} + d_j(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_{j-1}\alpha_{j-1} + c_j\alpha_j + c_{j+1}\alpha_{j+1} + \cdots + c_n\alpha_n) + d_{j+1}\alpha_{j+1} + \cdots + d_n\alpha_n = \theta$

or, $(d_1 + d_jc_1)\alpha_1 + (d_2 + d_jc_2)\alpha_2 + \cdots + (d_{j-1} + d_jc_{j-1})\alpha_{j-1} + d_jc_j\alpha_j + (d_{j+1} + d_jc_{j+1})\alpha_{j+1} + \cdots + (d_n + d_jc_n)\alpha_n = \theta$.

Since the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is linearly independent, we have $d_1 + d_jc_1 = 0, d_2 + d_jc_2 = 0, \cdots, d_{j-1} + d_jc_{j-1} = 0, d_jc_j = 0, d_{j+1} + d_jc_{j+1} = 0, \cdots, d_n + d_jc_n = 0$.

$d_jc_j = 0 \Rightarrow d_j = 0$, since $c_j \neq 0$. And $d_j = 0$ gives $d_i = 0$ for $i = 1, 2, \ldots, n$. This proves that the set $\{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\}$ is linearly independent.

We now prove that $L\{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\} = V$.

Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \alpha_j, \alpha_{j+1}, \ldots, \alpha_n\}$ and
$T = \{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\}$.

Since β is a linear combination of the vectors of S, each element of T is a linear combination of the vectors of S. Therefore $L(T) \subset L(S)$.

Since α_j is a linear combination of the vectors of T, each element of S is a linear combination of the vectors of T. Therefore $L(S) \subset L(T)$.

Consequently, $L(T) = L(S) = V$.

Hence $T = \{\alpha_1, \alpha_2, \ldots, \alpha_{j-1}, \beta, \alpha_{j+1}, \ldots, \alpha_n\}$ fulfils both the conditions for a basis of V. This completes the proof.

***Theorem* 4.5.3.** If $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of a finite dimensional vector space V over a field F, then any linearly independent set of vectors in V contains at most n vectors.

Proof. Let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be a linearly independent set of vectors in V. None of β_i is a zero vector. Since $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis of V and β_1 is a non-zero vector in V, $\beta_1 = c_1\alpha_1 + \cdots + c_n\alpha_n$, where $c_1, c_2, \ldots, c_n$ are scalars in F, not all of which are zero. Let $c_i \neq 0$

By Replacement theorem, $\{\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \beta_1, \alpha_{i+1}, \ldots, \alpha_n\}$ is a basis of V. Since $\beta_2 \neq 0$ and $\{\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \beta_1, \alpha_{i+1}, \ldots, \alpha_n\}$ is a basis of V, $\beta_2 = d_1\alpha_2 + d_2\alpha_2 + \cdots + d_{i-1}\alpha_{i-1} + d_i\beta_1 + d_{i+1}\alpha_{i+1} + \cdots + d_n\alpha_n$, where d_i's are scalars, not all zero.

We assert that at least one of $d_1, d_2, \ldots, d_{i-1}, d_{i+1}, \ldots, d_n$ is non-zero. Because, if all of them be zero, then $\beta_2 = d_i\beta_1$ and this will imply linear dependence of β_1, β_2 which is a contradiction.

Let $d_j \neq 0, j \neq i$. By Replacement theorem, $\{\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \beta_1, \alpha_{i+1}, \ldots, \alpha_{j-1}, \beta_2, \alpha_{j+1}, \ldots, \alpha_n\}$ is a new basis of V. Since $\beta_3 \neq 0$, $\beta_3 = t_1\alpha_1 + t_2\alpha_2 + \cdots + t_{i-1}\alpha_{i-1} + t_i\beta_1 + t_{i+1}\alpha_{i+1} + \cdots + t_{j-1}\alpha_{j-1} + t_j\beta_2 + t_{j+1}\alpha_{j+1} + \cdots + t_n\alpha_n$, where t_i's are scalars, not all zero.

We assert that at least one of $t_1, t_2, \ldots, t_{i-1}, t_{i+1}, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n$ is non-zero. Because, otherwise, $\beta_3 = t_i\beta_1 + t_j\beta_2$ and this will imply linear dependence of $\beta_1, \beta_2, \beta_3$, which is a contradiction.

Proceeding in this way we observe that at each step one α is replaced by one β and the resulting set remains a basis of V. The following cases may arise.

(i) $\beta_1, \beta_2, \ldots, \beta_r$ all come to the new basis containing some α's. In this case $r < n$.

(ii) $\beta_1, \beta_2, \ldots, \beta_r$ exhaust all α's and form the new basis. In this case $r = n$.

It can not happen that $r > n$. Because, then by Replacement theorem, n vectors $\beta_1, \beta_2, \ldots, \beta_n$ will come to the basis replacing all α's one after another and $\{\beta_1, \beta_2, \ldots, \beta_n\}$ becomes a new basis of V. Therefore the remaining vectors $\beta_{n+1}, \beta_{n+2}, \ldots, \beta_r$ of V will be each a linear combination of $\beta_1, \ldots, \beta_n$ showing that the set $\{\beta_1, \beta_2, \ldots, \beta_n, \beta_{n+1}, \ldots, \beta_r\}$ is linearly dependent, a contradiction.

Therefore $r \leq n$. This completes the proof.

Theorem 4.5.4. Any two bases of a finite dimensional vector space V have the same number of vectors.

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}, \{\beta_1, \beta_2, \ldots, \beta_m\}$ be two bases of a finite dimensional vector space V.

Since $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ is a basis of V and $\{\beta_1, \beta_2, \dots, \beta_m\}$ is a linearly independent set of vectors in V, $m \leq n$.

Since $\{\beta_1, \beta_2, \dots, \beta_m\}$ is a basis of V and $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ is a linearly independent set of vectors in V, $n \leq m$.

$m \leq n$ and $n \leq m \Rightarrow m = n$ and the theorem is done.

Dimension of a vector space.

Definition. The number of vectors in a basis of a vector space V is said to be the *dimension* (or *rank*) of V and is denoted by $\dim V$. The null space $\{\theta\}$ is said to be of dimension 0.

Examples (continued).

4. The demension of the vector space $\mathbb{R}^2$ is 2, since $E = \{(1,0),(0,1)\}$ is a basis.

The dimension of the vector space $\mathbb{R}^3$ is 3, since $E = \{(1,0,0),(0,1,0),(0,0,1)\}$ is a basis.

The dimension of the vector space $\mathbb{R}^n$ is n, since $E = \{(1,0,0,\dots,0),(0,1,0,\dots,0),\dots,(0,0,0,\dots,1)\}$ is a basis.

5. The dimension of the real vector space $\mathbb{R}_{m\times n}$ of all $m \times n$ real matrices is mn, since the set $\{E_{11}, E_{12}, \dots, E_{mn}\}$, where E_{ij} is an $m \times n$ matrix having 1 as the ijth element and 0 elsewhere, is a basis.

6. The dimension of the real vector space P_n is n. The set of polynomials $\{1, x, x^2, \dots, x^{n-1}\}$ is a basis.

7. The vector space P of all real polynomials is infinite dimensional.

Let S be a finite subset of P. Since there is only a finite number of polynomials in S, there is a polynomial f in S having a maximum degree, say m. Therefore every polynomial in S is of degree $\leq m$.

Let us consider an arbitrary polynomial p in $L(S)$. The degree of p cannot exceed m. This proves that $L(S)$ is a proper subspace of P because in P there are polynomials of degree $\geq m+1$. So S cannot be a basis of P.

Since S is an arbitrary finite set, P cannot have a finite basis. Therefore P is infinite demensional.

Theorem 4.5.5. Let V be a vector space of *dimension* n over a field F. Then any linearly independent set of n vectors of V is a basis of V.

Proof. Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a linearly independent set of vectors in V. Let β be an arbitrary vector of V and $\beta \neq \alpha_i$. Since $\dim V = n$, any basis of V contains n vectors and the set $\{\alpha_1, \alpha_2, \dots, \alpha_n, \beta\}$ which contains $n+1$ vectors is linearly dependent.

Therefore there exist scalars $c_1, c_2, \dots, c_n, c$, not all zero, in F such that $c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n + c\beta = \theta$ ... (i)

We assert that $c \neq 0$.

Because $c = 0$ implies $c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n = \theta$ where $c_1, c_2, \dots, c_n$ are not all zero, and this implies linear dependence of the set $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$, a contradiction.

Since $c \neq 0, c^{-1} \in F$ and $c^{-1}c = 1$, 1 being the unity in F.

From (i) $\quad c\beta = -c_1\alpha_1 - c_2\alpha_2 - \dots - c_n\alpha_n$.

Then $\beta = c^{-1}(-c_1\alpha_1 - c_2\alpha_2 - \dots - c_n\alpha_n)$
$\quad = d_1\alpha_1 + d_2\alpha_2 + \dots + d_n\alpha_n$, where $d_i = -c^{-1}c_i \in F$.

This shows that β is a linear combination of $\alpha_1, \alpha_2, \dots, \alpha_n$.

If however, $\beta = \alpha_i$ for some i, then also β is a linear combination of $\alpha_1, \alpha_2, \dots, \alpha_n$.

Thus $\beta \in L\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and so $V \subset L\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ ... (ii)

Again $\{\alpha_1, \alpha_2, \dots, \alpha_n\} \subset V$ and $L\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ being the smallest subspace containing $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $L\{\alpha_1, \alpha_2, \dots, \alpha_n\} \subset V$ ... (iii)

From (ii) and (iii), $V = L\{\alpha_1, \alpha_2, \dots, \alpha_n\}$. This proves that the set $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ is a basis of V.

This completes the proof.

Theorem 4.5.6. Let V be a vector space of *dimension* n over a field F. Then any subset of n vectors of V that generates V is a basis of V.

Proof. Since $\dim V = n$, any basis of V contains n vectors.

Let $S = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ be a generating set of V. We prove that it is linearly independent in V.

If it be linearly dependent, then by Deletion theorem there exists a proper subset S_1 of S such that S_1 is linearly independent in V and S_1 generates V.

Therefore S_1 is a basis of V containing less than n vectors of V and this contradicts that $\dim V = n$.

Therefore S is a linearly independent set in V. Thus the set S is a linearly independent generating set of V and therefore it is a basis of V.

This completes the proof.

Worked Examples (continued).

2. Find a basis for the vector space $\mathbb{R}^3$ that contains the vectors (1,2,0) and (1,3,1).

$\mathbb{R}^3$ is a vector space of dimension 3. The standard basis for $\mathbb{R}^3$ is $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ where $\epsilon_1 = (1,0,0), \epsilon_2 = (0,1,0), \epsilon_3 = (0,0,1)$.

Let $\alpha = (1,2,0), \beta = (1,3,1)$. Then $\alpha = 1\epsilon_1 + 2\epsilon_2 + 0\epsilon_3$.

Since the coefficient of ϵ_1 in the representation of α is non-zero, by Replacement theorem α can replace ϵ_1 in the basis $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ and $\{\alpha, \epsilon_2, \epsilon_3\}$ can be a new basis for $\mathbb{R}^3$.

Let $\beta = c_1\alpha + c_2\epsilon_2 + c_3\epsilon_3$.

Then $(1,3,1) = c_1(1,2,0) + c_2(0,1,0) + c_3(0,0,1)$.

Therefore $c_1 = 1, 2c_1 + c_2 = 3, c_3 = 1$.

We have $c_1 = 1, c_2 = 1, c_3 = 1$ and $\beta = \alpha + \epsilon_2 + \epsilon_3$.

Since the coefficient of ϵ_2 in the representation of β is non-zero, by Replacement theorem β can replace ϵ_2 in the basis $\{\alpha, \epsilon_2, \epsilon_3\}$ and $\{\alpha, \beta, \epsilon_3\}$ can be a new basis for $\mathbb{R}^3$.

Note. The replacement can be done in more than one ways and thus different bases for $\mathbb{R}^3$ can be obtained.

3. Prove that the set $S = \{(2,1,1), (1,2,1), (1,1,2)\}$ is a basis of $\mathbb{R}^3$.

Let $\alpha_1 = (2,1,1), \alpha_2 = (1,2,1), \alpha_3 = (1,1,2)$.

Let us consider the relation $c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = \theta$, where $c_1, c_2, c_3 \in \mathbb{R}$.

Then $c_1(2,1,1) + c_2(1,2,1) + c_3(1,1,2) = (0,0,0)$.

This gives $2c_1 + c_2 + c_3 = 0, \quad c_1 + 2c_2 + c_3 = 0, \quad c_1 + c_2 + 2c_3 = 0$.

This is a homogeneous system of equations in c_1, c_2, c_3. Here the coefficient determinant $= -4 \neq 0$.

By Cramer's rule, there exists a unique solution and the solution is $c_1 = 0, c_2 = 0, c_3 = 0$. This proves that the set S is linearly independent.

Since $\mathbb{R}^3$ is a vector space of dimension 3 and S is a linearly independent set containing 3 vectors of $\mathbb{R}^3$, S is a basis of $\mathbb{R}^3$.

4. Let V be a real vector space with $\{\alpha, \beta, \gamma\}$ as a basis. Prove that the set $\{\alpha + \beta + \gamma, \beta + \gamma, \gamma\}$ is also a basis of V.

Let $\alpha + \beta + \gamma = \alpha_1, \beta + \gamma = \alpha_2, \gamma = \alpha_3$.

Let us consider the relation $c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 = \theta$, where $c_1, c_2, c_3 \in \mathbb{R}$.

Then $c_1(\alpha + \beta + \gamma) + c_2(\beta + \gamma) + c_2\gamma = \theta$

or, $c_1\alpha + (c_1 + c_2)\beta + (c_1 + c_2 + c_3)\gamma = \theta$.

This implies $c_1 = 0, c_1 + c_2 = 0, c_1 + c_2 + c_3 = 0$, since the set $\{\alpha, \beta, \gamma\}$ is linearly independent.

The solution is $c_1 = 0, c_2 = 0, c_3 = 0$ and this proves that the set $\{\alpha_1, \alpha_2, \alpha_3\}$ is linearly independent.

V is a vector space of dimension 3 and $\{\alpha_1, \alpha_2, \alpha_3\}$ is a linearly independent set containing 3 vectors of V. Therefore $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis of V.

5. V is the vector space of all 2×2 real matrices. Prove that the set $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$ is a basis of V.

Let $\alpha_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \alpha_4 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$.

Let us consider the relation $c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 + c_4\alpha_4 = O$, where $c_1, c_2, c_3, c_4 \in \mathbb{R}$.

Then $c_1 + c_2 + c_3 + c_4 = 0, c_2 + c_3 + c_4 = 0, c_3 + c_4 = 0, c_4 = 0$.

This gives $c_1 = 0, \quad c_2 = 0, \quad c_3 = 0, \quad c_4 = 0$.

This proves that the set $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ is linearly independent.

V is a vector space of dimension 4 and $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ is a linearly independent set containing 4 vectors of V. Therefore the set $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ is a basis of V.

6. Find a basis and the dimension of the subspace W of $\mathbb{R}^3$, where $W = \{(x,y,z) \in \mathbb{R}^3 : x + y + z = 0\}$.

Let $\xi = (a,b,c) \in W$. Then $a, b, c \in \mathbb{R}$ and $a + b + c = 0$.

Therefore $\xi = (a, b, -a-b) = a(1,0,-1) + b(0,1,-1)$.

Let $\alpha = (1,0,-1), \beta = (0,1,-1)$. Then $\xi = a\alpha + b\beta \in L\{\alpha, \beta\}$.

Therefore $W \subset L\{\alpha, \beta\}$.

Again $\alpha \in W, \beta \in W$. This implies $L\{\alpha, \beta\} \subset W$, as W is a subspace.

Consequently, $W = L\{\alpha, \beta\}$.

α, β are linearly independent in W, since none of them is a scalar multiple of the other. [Corollary, Theorem 4.4.5.]

Hence the set $\{\alpha, \beta\}$ is a basis of W and $\dim W = 2$.

7. Find a basis and the dimension of the subspace W of $\mathbb{R}^3$, where $W = \{(x,y,z) \in \mathbb{R}^3 : x + 2y + z = 0, 2x + y + 3z = 0\}$.

Let $\xi = (a,b,c)$ be an arbitrary vector of W.

Then $a + 2b + c = 0$ and $2a + b + 3c = 0$, $a, b, c \in \mathbb{R}$.

Solving, we have $\frac{a}{5} = \frac{b}{-1} = \frac{c}{-3} = k$, say.

ξ takes the form $k(5,-1,-3)$, where k is an arbitrary real number.

Therefore $W = L\{\alpha\}$ where $\alpha = (5,-1,-3)$. Since $\{\alpha\}$ is a linearly independent set, $\{\alpha\}$ is a basis of W and $\dim W = 1$.

8. Find dim $S \cap T$, where S and T are subspaces of the vector space $\mathbb{R}^4$ given by

$S = \{(x, y, z, w) \in \mathbb{R}^4 : 2x + y + 3z + w = 0\}$,
$T = \{(x, y, z, w) \in \mathbb{R}^4 : x + 2y + z + 3w = 0\}$.

$S \cap T = \{(x, y, z, w)\} \in \mathbb{R}^4 : 2x + y + 3z + w = 0, x + 2y + z + 3w = 0\}$.

Let $\xi = (a, b, c, d) \in S \cap T$. Then $a, b, c, d \in \mathbb{R}$ and

$2a + b + 3c + d = 0, \quad a + 2b + c + 3d = 0.$

So we have $a = -5b - 8d$, $c = 3b + 5d$.

Hence $\xi = (-5b - 8d, b, 3b + 5d, d) = b(-5, 1, 3, 0) + d(-8, 0, 5, 1)$.

Let $\alpha = (-5, 1, 3, 0)$, $\beta = (-8, 0, 5, 1)$. Then $\xi \in L\{\alpha, \beta\}$ and therefore $S \cap T \subset L\{\alpha, \beta\}$ ... (i)

Again $\alpha \in S, \alpha \in T; \beta \in S, \beta \in T$. This implies $\alpha \in S \cap T, \beta \in S \cap T$ and therefore $L(\alpha, \beta) \subset S \cap T$ ... (ii)

From (i) and (ii) we have $S \cap T = L\{\alpha, \beta\}$.

α, β are linearly independent in $S \cap T$, since none of them is a scalar multiple of the other. [Corollary, Theorem 4.4.5.]

Hence dim $S \cap T = 2$.

Let V be a non-zero vector space of finite dimension. Then a basis of V contains a finite number of vectors. But can we construct a basis of V? Since V contains non-zero vectors, let α be one such. The singleton set $\{\alpha\}$ is linearly independent in V. If $L\{\alpha\} = V$ then $\{\alpha\}$ itself is a basis. If not, a basis can be constructed that contains $\{\alpha\}$.

The method of construction is discussed in the following theorem.

Theorem 4.5.7. Extension theorem.

A linearly independent set of vectors in a finite dimensional vector space V over a field F is either a basis of V, or it can be extended to a basis of V.

Proof. Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ be a linearly independent set in V. $L(S)$ being the smallest subspace containing S, $L(S) \subset V$.

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

If $L(S)$ be a proper subspace of V, then $V - L(S) \neq \phi$. Let $\beta \in V - L(S)$. We prove the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \beta\}$ is linearly independent.

Let us consider the relation $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r + b\beta = 0$ where $c_1, c_2, \ldots, c_r, b$ are scalars in F ... (i)

We assert that $b = 0$. Because if $b \neq 0$, then b^{-1} exists in F and β can be expressed as

$\beta = -b^{-1}(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r) = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_r\alpha_r$, where $d_i = -b^{-1}c_i \in F$, $i = 1, 2, \ldots, r$.

This implies $\beta \in L(S)$, a contradiction. Therefore $b = 0$.

The linear independence of the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ and $b = 0$ together imply $c_1 = c_2 = \cdots = c_r = b = 0$ in (i).

This proves linear independence of the set $S_1 = \{\alpha_1, \alpha_2, \ldots, \alpha_r, \beta\}$.

$L(S_1) \subset V$. If $L(S_1) = V$, then S_1 is a basis of V and as S_1 is an extension of S, the theorem is proved.

If however, $L(S_1)$ is a proper subspace of V, we can take a vector $\gamma \in V - L(S_1)$ and proceed as before.

Since V is finite dimensional, after a finite number of steps we come to a finite set of vectors as an extension of S and also as a basis of V. This completes the proof.

Worked Example (continued).

9. Extend the set $\{(1, 1, 1, 1), (1, -1, 1, -1)\}$ to a basis of $\mathbb{R}^4$.

Let $\alpha_1 = (1, 1, 1, 1), \alpha_2 = (1, -1, 1, -1)$.

Let us examine if $\epsilon_3 = (0, 0, 1, 0)$ belongs to $L\{\alpha_1, \alpha_2\}$.

Let $\epsilon_3 = c_1\alpha_1 + c_2\alpha_2$, where $c_1, c_2 \in \mathbb{R}$.

Then $c_1 + c_2 = 0, \quad c_1 - c_2 = 0, \quad c_1 + c_2 = 1$ and $c_1 - c_2 = 0$.

This is an inconsistent system of equations in c_1, c_2. Therefore $\epsilon_3 \notin L\{\alpha_1, \alpha_2\}$. Consequently, the set $\{\alpha_1, \alpha_2, \epsilon_3\}$ is a linearly independent set in $\mathbb{R}^4$.

Since $\mathbb{R}^4$ is a vector space of dimension 4, the set $\{\alpha_1, \alpha_2, \epsilon_3\}$ containing three vectors of $\mathbb{R}^4$ can not be a basis of $\mathbb{R}^4$.

Let us examine if $\epsilon_4 = (0, 0, 0, 1)$ belongs to $L\{\alpha_1, \alpha_2, \epsilon_3\}$.

Let $\epsilon_4 = d_1\alpha_1 + d_2\alpha_2 + d_3\epsilon_3$, where $d_1, d_2, d_3 \in \mathbb{R}$.

Then $d_1 + d_2 = 0, \quad d_1 - d_2 = 0, \quad d_1 + d_2 + d_3 = 0$ and $d_1 - d_2 = 1$.

This is an inconsistent system of equations in d_1, d_2, d_3. Therefore $\epsilon_4 \notin L\{\alpha_1, \alpha_2, \epsilon_3\}$. Consequently, the set $\{\alpha_1, \alpha_2, \epsilon_3, \epsilon_4\}$ is a linearly independent set in $\mathbb{R}^4$.

Since $\mathbb{R}^4$ is a vector space of dimension 4 and the set $\{\alpha_1, \alpha_2, \epsilon_3, \epsilon_4\}$ is a linearly independent set of 4 vectors in $\mathbb{R}^4$, $\{\alpha_1, \alpha_2, \epsilon_3, \epsilon_4\}$ is a basis of $\mathbb{R}^4$.

Clearly, this basis set is an extension of the given set of vectors $\{(1,1,1,1),(1,-1,1,-1)\}$.

Note. The basis $\{\alpha_1, \alpha_2, \epsilon_3, \epsilon_4\}$ obtained as an extension of the linearly independent set $\{\alpha_1, \alpha_2\}$ is not unique.

In the first step the vector ϵ_3 is selected at random. In fact we are to select a vector in $\mathbb{R}^4$ not contained in $L\{\alpha_1, \alpha_2\}$.

Theorem 4.5.8. Let V be a vector space over a field F. A subset $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of V is a basis of V if and only if every element of V has a *unique representation* as a linear combination of the vectors of B.

Proof. Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V.

Let $\alpha \in V$. Then $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$, for some scalars $c_i \in F$.

Let us assume that $\alpha = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_n\alpha_n$, for some other scalars $d_i \in F$.

Then $\theta = \alpha - \alpha = (c_1 - d_1)\alpha_1 + (c_2 - d_2)\alpha_2 + \cdots + (c_n - d_n)\alpha_n$.

Since the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is linearly independent, it follows that $c_i - d_i = 0$ for $i = 1, 2, \ldots, n$.

So $c_1 = d_1, c_2 = d_2, \ldots, c_n = d_n$ and therefore c_i's are unique.

Conversely, let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a subset of V such that every vector of V has a unique representation as a linear combination of the vectors of B.

Clearly, $V = L\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$. ... (i)

$\theta \in V$, and by the condition, θ has a unique representation as a linear combination of the vectors of B. Let $\theta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$.

This is obviously satisfied by $c_1 = 0, c_2 = 0, \cdots, c_n = 0$ and because of uniqueness in the condition, it follows that

$c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \theta \Rightarrow c_1 = 0, c_2 = 0, \cdots, c_n = 0.$

This proves that the set B is a linearly independent set. ... (ii)

From (i) and (ii) it follows that B is a basis of V.

Observation. Let V be a vector space over a field F. If we examine the following pairs of statements in each of the items from 1 to 4, we observe that a sort of *duality relations* exist between *a spanning set* of V and a *linearly independent set* of V.

1. (i) If dim $V = n$, no linearly independent set in V contains more than n vectors of V;

(ii) If dim $V = n$, no generating set of V contains less than n vectors of V.

2. (i) If dim $V = n$, every linearly independent set of n vectors in V is a basis of V; [Theorem 4.5.5.]

(ii) If dim $V = n$, every generating set of n vectors of V is a basis of V. [Theorem 4.5.6.]

3. (i) Every linearly independent set in V is a subset of a basis of V; [Extension theorem]

(ii) Every generating set of V is a superset of a basis of V. [Deletion theorem]

4. (i) A basis of V is a maximal linearly independent set in V;

(ii) A basis of V is a minimal generating set of V.

4.6. Co-ordinatisation of vectors.

Let $B = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an *ordered basis* of a vector space V over a field F. Then to each vector α in V there corresponds a well determined ordered set of n scalars $c_1, c_2, \ldots, c_n$ in F such that $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$.

The ordered n-tuple $(c_1, c_2, \ldots, c_n)$ is said to be the *co-ordinate vector* of α relative to the *ordered basis* B and is denoted by $(\alpha)_B$.

We insist that the set of vectors in B should be ordered, because a change in $(\alpha)_B$ occurs if the relative order of vectors in B be changed.

Thus the co-ordinate vectors in an abstract vector space V of dimension n over a field F relative to an ordered basis B are the elements of F^n. In particular, the co-ordinate vectors of the basis elements $\alpha_1, \alpha_2, \ldots, \alpha_n$ relative to B are $(1,0,0,\ldots,0)$, $(0,1,0,\ldots,0)$, ..., $(0,0,0,\ldots,1)$ respectively.

Worked Examples.

1. Find the co-ordinate vector of $\alpha = (1,3,1)$ relative to the ordered basis $B = (\alpha_1, \alpha_2, \alpha_3)$ of $\mathbb{R}^3$, where $\alpha_1 = (1,1,1), \alpha_2 = (1,1,0), \alpha_3 = (1,0,0)$.

Let us find scalars c_1, c_2, c_3 such that $(1,3,1) = c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3$.

Then $(1,3,1) = c_1(1,1,1) + c_2(1,1,0) + c_3(1,0,0)$.

This gives $c_1 + c_2 + c_3 = 1, c_1 + c_2 = 3, c_1 = 1$.

Solving, we have $c_1 = 1, c_2 = 2, c_3 = -2$. So $(\alpha)_B = (1,2,-2)$.

2. Find the vector α in $\mathbb{R}^3$ whose co-ordinate vector relative to the ordered basis B of Example 1 is $(3,2,1)$.

$\alpha = 3\alpha_1 + 2\alpha_2 + \alpha_3 = 3(1,1,1) + 2(1,1,0) + (1,0,0) = (6,5,3)$.

Exercises 7

1. Show that the set S of vectors is linearly dependent in $\mathbb{R}^3$.
   (i) $S = \{(2,3,1),(2,1,3),(1,1,1)\}$,
   (ii) $S = \{(0,-1,3),(3,4,3),(1,1,2)\}$.

2. Show that the set S of vectors is linearly dependent in $\mathbb{R}^4$.
   (i) $S = \{(1,1,1,0),(1,0,1,1),(1,2,1,2),(1,1,1,1)\}$,
   (ii) $S = \{(2,3,1,4),(3,2,4,1),(1,1,1,1)\}$.

3. Find $k \in \mathbb{R}$ so that the set S is linearly dependent in $\mathbb{R}^3$.
   (i) $S = \{(1,2,1),(k,3,1),(2,k,0)\}$,
   (ii) $S = \{(k,k,1),(k,1,k),(1,k,k)\}$.

4. Find the conditions on $a, b \in \mathbb{R}$ so that the set of vectors is linearly dependent in $\mathbb{R}^3$.
   (i) $\{(a,b,b),(b,a,b),(b,b,a)\}$; (ii) $\{(a,b,1),(b,1,a),(1,a,b)\}$.

5. Show that set of vectors $S = \{(1,2,0),(2,1,3),(1,1,1),(2,3,1)\}$ is linearly dependent in $\mathbb{R}^3$. Find a linearly independent subset S_1 of S such that $L(S_1) = L(S)$.

6. Show that the set of vectors $S = \{(1,2,3,0),(2,1,0,3),(1,1,1,1),(2,3,4,1)\}$ is linearly dependent in $\mathbb{R}^4$. Find a linearly independent subset S_1 of S such that $L(S_1) = L(S)$.

7. Show that the set S of vectors is linearly independent in $\mathbb{R}^3$.
   (i) $S = \{(1,2,3),(2,3,1),(3,1,2)\}$;
   (ii) $S = \{(0,1,1),(1,0,1),(1,1,0)\}$.

8. Show that the set S of vectors is linearly independent in $\mathbb{R}^4$.
   (i) $S = \{(1,2,3,0),(2,3,0,1),(3,0,1,2)\}$;
   (ii) $S = \{(1,1,1,0),(1,1,0,1),(1,0,1,1),(0,1,1,1)\}$.

9. For what real values of k does the set
   (i) $S = \{(k,1,1),(1,k,1),(1,1,k)\}$ form a basis of $\mathbb{R}^3$?
   (ii) $S = \{(k,1,1,1),(1,k,1,1),(1,1,k,1),(1,1,1,k)\}$ form a basis of $\mathbb{R}^4$?

10. Let $\{\alpha,\beta,\gamma\}$ be a basis of a real vector space V and c be a non-zero real number. Prove that
    (i) $\{c\alpha, c\beta, c\gamma\}$ is a basis of V,
    (ii) $\{\alpha + c\beta, \beta, \gamma\}$ is a basis of V,
    (iii) $\{\alpha + c\beta, \beta + c\gamma, \gamma + c\alpha\}$ may not be a basis of V.

11. If $\{\alpha_1,\alpha_2,\alpha_3\}$ be a basis of a real vector space V and $\beta_1 = \alpha_1 + \alpha_3$, $\beta_2 = 2\alpha_1 + 3\alpha_2 + 4\alpha_3$, $\beta_3 = \alpha_1 + 2\alpha_2 + 3\alpha_3$, prove that $\{\beta_1,\beta_2,\beta_3\}$ is also a basis of V.

12. Prove that the set $S = \{(1,1,0),(1,0,1),(0,1,1)\}$ is a basis of the vector space $\mathbb{R}^3$. Show that the vector $(1,1,1)$ may replace any one of the vectors of the set S to form a new basis for $\mathbb{R}^3$, but the same is not true for the vector $(3,1,2)$.

13. Find a basis for the vector space $\mathbb{R}^3$, that contains the vectors
    (i) (1,2,1) and (3,6,2); (ii) (1,0,1) and (1,1,1).

14. Find a basis for the vector space $\mathbb{R}^4$, that contains the vectors
    (i) (1,0,1,0) and (0,1,0,1); (ii) (1,1,0,0) and (1,1,1,0).

15. Extend the set S to obtain a basis of the vector space $\mathbb{R}^3$.
    (i) $S = \{(1,2,1),(2,1,1)\}$;
    (ii) $S = \{(1,1,0),(1,1,1)\}$.

16. Extend the set S to obtain a basis of the vector space $\mathbb{R}^4$.
    (i) $S = \{(1,0,1,0),(0,1,0,1)\}$;
    (ii) $S = \{(1,1,0,0),(1,1,1,0)\}$.

17. Let V be a real vector space with a basis $\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$. Examine if $\{\alpha_1+\alpha_2, \alpha_2+\alpha_3, \cdots, \alpha_n+\alpha_1\}$ is also a basis of V.

18. Find the dimension of the subspace S of $\mathbb{R}^3$ defined by
    (i) $S = \{(x,y,z) \in \mathbb{R}^3 : 2x + y - z = 0\}$;
    (ii) $S = \{(x,y,z) \in \mathbb{R}^3 : x + 2y = z, 2x + 3z = y\}$.

19. Find the dimension of the subspace S of $\mathbb{R}^4$ defined by
    (i) $S = \{(x,y,z,w) \in \mathbb{R}^4 : x + y + z + w = 0\}$;
    (ii) $S = \{(x,y,z,w) \in \mathbb{R}^4 : x + 2y - z = 0, 2x + y + w = 0\}$.

20. Find the dimension of the subspace $S \cap T$ of $\mathbb{R}^4$ where $S = \{(x,y,z,w) \in \mathbb{R}^4 : x + y + z + w = 0\}$, $T = \{(x,y,z,w) \in \mathbb{R}^4 : 2x + y - z + w = 0\}$.

21. Show that S is a subspace of $\mathbb{R}^3$. Find a basis for S.
    (i) $S = \left\{ X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : AX = O \right\}$, where $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 1 & 0 \end{pmatrix}$;
    (ii) $S = \left\{ X = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : AX = O \right\}$, where $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

22. Find a basis and determine the dimension of the subspace S of the vector space $\mathbb{R}_{2\times 2}$.

(i) $S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}_{2\times 2} : a + b = 0 \right\}$,

(ii) $S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}_{2\times 2} : a = d = 0 \right\}$,

(iii) $S =$ the set of all 2×2 real diagonal matrices,

(iv) $S =$ the set of all 2×2 real symmetric matrices,

(v) $S =$ the set of all 2×2 real skew symmetric matrices.

23. Show that the set S is a basis of the vector space $\mathbb{R}_{2\times 2}$.

(i) $S = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \right\}$;

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

24. Find the co-ordinate vector of α in $\mathbb{R}^3$ relative to the basis $(\alpha_1, \alpha_2, \alpha_3)$.

(i) $\alpha = (0,3,1)$, $\alpha_1 = (1,1,0)$, $\alpha_2 = (1,0,1)$, $\alpha_3 = (0,1,1)$;

(ii) $\alpha = (2,3,3)$, $\alpha_1 = (2,1,1)$, $\alpha_2 = (1,2,1)$, $\alpha_3 = (1,1,2)$.

25. Show that the set $\mathbb{C}$ (the set of all complex numbers) forms a vector space over the field $\mathbb{R}$. What is the dimension of $\mathbb{C}$ over $\mathbb{R}$?

26. Show that the set $\mathbb{R}$ (the set of all real numbers) forms a vector space over the field $\mathbb{Q}$. What is the dimension of $\mathbb{R}$ over $\mathbb{Q}$?

4.7. Complement of a subspace.

Theorem 4.7.1. If U be a subspace of a finite dimensional vector space V and $\dim V = n$, then U is finite dimensional and $\dim U \leq n$.

Proof. **Case 1.** $U = \{\theta\}$. Then $\dim U = 0 < n$.

Case 2. $U \neq \{\theta\}$. Then there exists a non-zero vector, say α, in U.

If $L\{\alpha\} = U$ then we have a finite basis $\{\alpha\}$ for U. If not, there is a non-zero vector β in U such that the set $\{\alpha, \beta\}$ is linearly independent.

If $L\{\alpha, \beta\} = U$ then we have a finite basis $\{\alpha, \beta\}$ for U. If not, the process can be continued and after a finite number of steps we get a linearly independent set S forming a basis of U. S being a linearly independent set in V also, S contains at most n vectors by Theorem 4.5.3. Therefore $\dim U \leq n$.

Combining the cases, U is finite dimensional and $\dim U \leq n$.

Theorem 4.7.2. Let U and W be two subspace of a finite dimensional vector space V over a field F. Then

$$\dim (U + W) = \dim U + \dim W - \dim (U \cap W).$$

Proof. Since U and W are subspaces of a finite dimensional vector space V, $\dim U$, $\dim W$, $\dim (U+W)$, $\dim U \cap W$ are each finite and

$\dim U \cap W \leq \dim U$, $\dim U \leq \dim (U + W) \leq \dim V$;

$\dim U \cap W \leq \dim W$, $\dim W \leq \dim (U + W) \leq \dim V$.

Let $\dim (U \cap W) = r$, $\dim U = r + s$ and $\dim W = r + t$.

Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ be a basis for $U \cap W$ and let S be supplemented by additional vectors $\beta_1, \beta_2, \ldots, \beta_s$ to make up a basis for U. Similarly, let S be supplemented by additional vectors $\gamma_1, \gamma_2, \ldots, \gamma_t$ to make up a basis for W.

We propose to prove that

$B = \{\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_s, \gamma_1, \gamma_2, \ldots, \gamma_t\}$ is a basis for $U + W$.

The set $B \subset U + W$ and therefore $L(B) \subset U + W$.

Any vector $\xi \in U + W$ is of the form $u + w$ where $u \in L\{\alpha_1, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_s\}$ and $w \in L\{\alpha_1, \ldots, \alpha_r, \gamma_1, \gamma_2, \ldots, \gamma_t\}$.

Therefore $U + W \subset L(B)$. Thus $L(B) = U + W$.

In order to establish linear independence of the set B, let

$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_r\alpha_r + b_1\beta_1 + b_2\beta_2 + \cdots + b_s\beta_s + c_1\gamma_1 + c_2\gamma_2 + \cdots + c_t\gamma_t = \theta$,

for some scalars $a_i, b_i, c_i \in F$. $\quad \ldots$ (i)

Then $a_1\alpha_1 + a_2\alpha_2 + \cdots + a_r\alpha_r + b_1\beta_1 + b_2\beta_2 + \cdots + b_s\beta_s$

$= -(c_1\gamma_1 + c_2\gamma_2 + \cdots + c_t\gamma_t). \quad \ldots \quad$ (ii)

We observe that the left hand vector belongs to U. Hence the right hand vector belongs to both U and W and so is in $U \cap W$.

Therefore $-(c_1\gamma_1 + c_2\gamma_2 + \cdots + c_t\gamma_t) = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_r\alpha_r$ for some $d_i \in F$.

or, $c_1\gamma_1 + c_2\gamma_2 + \cdots + c_t\gamma_t + d_1\alpha_1 + d_2\alpha_2 + d_r\alpha_r = \theta. \quad \ldots \quad$ (iii)

But since the vectors $\alpha_1, \alpha_2, \ldots, \alpha_r, \gamma_1, \gamma_2, \ldots, \gamma_t$ constitute a basis for W, the relation (iii) implies $c_1 = c_2 = \cdots = c_t = d_1 = d_2 = \cdots = d_r = 0$.

Therefore the relation (ii) becomes

$a_1\alpha_1 + a_2\alpha_2 + \cdots + a_r\alpha_r + b_1\beta_1 + b_2\beta_2 + \cdots + b_s\beta_s = \theta$.

Since the vectors $\alpha_1, \alpha_2, \ldots, \alpha_r, \beta_1, \beta_2, \ldots, \beta_s$ constitute a basis for U, we have $a_1 = a_2 = \cdots = a_r = b_1 = b_2 = \cdots = b_s = 0$.

Since all scalars a_i, b_i, c_i in relation (i) are shown to be zero, we conclude that the set B is linearly independent and we have already shown that B generates $U + W$. Thus B is a basis for $U + W$.

Hence $\dim (U + W) = r + s + t = \dim U + \dim W - \dim (U \cap W)$.

Theorem 4.7.3. If U and W be two subspaces of a vector space V over a field F such that $U \cap W = \{\theta\}$ and if $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ and $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be respectively the bases of U and W then $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$ is a basis of $U + W$.

Proof. Every vector of U is of the form $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m, c_i \in F$ and every vector of W is of the form $d_1\beta_1 + d_2\beta_2 + \cdots + d_n\beta_n, d_i \in F$.

Therefore every vector of $U+W$ is a linear combination of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n$.

So $U + W \subset L\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$.

Again every linear combination of $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$ can be expressed as the sum $u + w$, where $u \in U, w \in W$.

So $L\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\} \subset U + W$.

Thus $L\{\alpha_1, \ldots, \beta_n\} = U + W$.

Let us consider the relation $t_1\alpha_1 + t_2\alpha_2 + \cdots + t_m\alpha_m + t_{m+1}\beta_1 + t_{m+2}\beta_2 + \cdots + t_{m+n}\beta_n = \theta, t_i \in F$. This implies

$t_1\alpha_1 + t_2\alpha_2 + \cdots + t_m\alpha_m = -(t_{m+1}\beta_1 + t_{m+2}\beta_2 + \cdots + t_{m+n}\beta_n) \ldots$ (i)

But $t_1\alpha_1 + t_2\alpha_2 + \cdots + t_m\alpha_m$ is a vector in U and $-(t_{m+1}\beta_1 + \cdots + t_{m+n}\beta_n)$ is a vector in W. By hypothesis, the only vector common to U and W is θ. So the equality (i) implies

$t_1\alpha_1 + t_2\alpha_2 + \cdots + t_m\alpha_m = \theta = -(t_{m+1}\beta_1 + t_{m+2}\beta_2 + \cdots + t_{m+n}\beta_n)$.

But $t_1\alpha_1 + t_2\alpha_2 + \cdots + t_m\alpha_m = \theta \Rightarrow t_1 = t_2 = \cdots = t_m = 0$, since the set $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is linearly independent.

Also $t_{m+1}\beta_1 + t_{m+2}\beta_2 + \cdots + t_{m+n}\beta_n = \theta \Rightarrow t_{m+1} = \cdots = t_{m+n} = 0$, since the set $\{\beta_1, \beta_2, \ldots, \beta_n\}$ is linearly independent.

Therefore $t_1\alpha_1 + t_2\alpha_2 + \cdots + t_m\alpha_m + t_{m+1}\beta_1 + \cdots + t_{m+n}\beta_n = \theta$ $\Rightarrow t_1 = t_2 = \cdots = t_m = t_{m+1} = \cdots = t_{m+n} = 0$.

This proves that $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$ is a basis of $U+W$.

Definition. Direct sum.

Two subspaces U and W of a vector space V are said to be *complement* of each other if $U \cap W = \{\theta\}$ and $U + W = V$.

If two subspaces U and W are complements of each other in V, then V is said to be the *direct sum of* the subspaces U and W. This is expressed by $V = U \oplus W$.

Theorem 4.7.4. If V be a finite dimensional vector space and U and W are complements of each other in V, then $\dim V = \dim U + \dim W$.

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basis of U and $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be a basis of W. By the Theorem 4.7.3, $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$ is a basis of V. Therefore $\dim V = m + n = \dim U + \dim W$.

Theorem 4.7.5. If a finite dimensional vector space V be the direct sum of two subspaces U and W then every vector α in V has a unique representation of the form $\alpha = u + w$, where $u \in U, w \in W$.

Proof. Since V is the linear sum of two subspaces U and W every vector α in V can be expressed as $\alpha = u + w$, where $u \in U, w \in W$. We assert that this representation is unique if $V = U \oplus W$.

If possible, let there be two such repersentations for α given by $\alpha = u + w$ and $\alpha = u' + w'$ where $u, u' \in U; w, w' \in W$.

Then $\theta = \alpha - \alpha = (u - u') + (w - w')$ or, $u - u' = w' - w$.

Now $u - u' \in U, w' - w \in W$ and the equality shows that either of them is in $U \cap W$. Since $U \cap W = \theta$, $u = u'$ and $w = w'$ and our assertion is established.

Theorem 4.7.6. Every subspace of a finite dimensional vector space over a field F possesses a complement.

Proof. Let U be a subspace of a finite dimensional vector space V and let $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basis of U. By Extension theorem the set $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ can be extended to a basis of V.

Let $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ be a basis of V. Let W be the vector space spanned by $\{\alpha_{m+1}, \alpha_{m+2}, \ldots, \alpha_n\}$. We shall prove that W is a complement of U.

$U = L\{\alpha_1, \alpha_2, \ldots, \alpha_m\}, W = L\{\alpha_{m+1}, \alpha_{m+2}, \ldots, \alpha_n\}$

Therefore $U + W = L\{\alpha_1, \alpha_2, \ldots, \alpha_{m+1}, \ldots, \alpha_n\}$.

We assert that $U \cap W = \{\theta\}$. If not, let β be a non-null vector in $U \cap W$. Then $\beta \in L\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ and $\beta \in L\{\alpha_{m+1}, \alpha_{m+2}, \ldots, \alpha_n\}$.

So $\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m$ for some scalars $c_i \in F$, not all zero. Also $\beta = d_1\alpha_{m+1} + d_2\alpha_{m+2} + \cdots + d_{n-m}\alpha_n$ for some scalars $d_i \in F$, not all zero.

Therefore $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m - d_1\alpha_{m+1} - \cdots - d_{n-m}\alpha_n = \theta$ holds for some scalars $c_1, c_2, \ldots, c_m, d_1, d_2, \ldots, d_{n-m}$, not all zero.

This implies that $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \alpha_{m+1}, \ldots, \alpha_n\}$ is a linearly dependent set, which is a contradiction and our assertion is established.

Hence U and W are complements in V.

Note. The extension of the basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of U does not lead to a uniquely determined basis of V because the additional vectors for the completion of the basis can be chosen in different ways. Thus there may be different complements of U in V.

Worked Examples.

1. If $U = L\{(1,2,1),(2,1,3)\}, W = L\{(1,0,0),(0,0,1)\}$, show that U and W are subspaces of $\mathbb{R}^3$. Determine dim U, dim W, dim $(U \cap W)$. Deduce that dim $(U + W) = 3$.

Let $\alpha = (1,2,1), \beta = (2,1,3), \gamma = (1,0,0), \delta = (0,0,1)$.

$\{\alpha, \beta\}$ is linearly independent and therefore U is a subspace of $\mathbb{R}^3$ of dimension 2. $\{\gamma, \delta\}$ is linearly independent and therefore W is a subspace of $\mathbb{R}^3$ of dimension 2.

Let λ be a vector in $U \cap W$. Then $\lambda = a\alpha + b\beta$ for some real numbers a, b and also $\lambda = c\gamma + d\delta$ for some real numbers c, d.

Therefore $a(1,2,1) + b(2,1,3) = c(1,0,0) + d(0,0,1)$.

This gives $a + 2b = c, 2a + b = 0, a + 3b = d$.

So $a = -\frac{b}{2}, c = \frac{3}{2}b, d = \frac{5}{2}b$ and therefore $\lambda = b(\frac{3}{2}, 0, \frac{5}{2}), b \in \mathbb{R}$.

Hence $U \cap W$ is a subspace of dimension 1.

So dim $(U + W)$ =dim U+ dim W− dim $(U \cap W) = 2 + 2 - 1 = 3$.

2. $U = L\{(1,2,1),(2,1,3)\}$. Find two different complements of U in $\mathbb{R}^3$.

Let $\alpha = (1,2,1), \beta = (2,1,3)$. Then $\{\alpha, \beta\}$ is linearly independent. Therefore U is a subspace of $\mathbb{R}^3$ of dimension 2.

$\epsilon_1 = (1,0,0)$ is a vector in $\mathbb{R}^3$ but not in U, since $(1,0,0)$ cannot be expressed as $a(1,2,1) + b(2,1,3)$ for real a, b.

Therefore $\{\alpha, \beta, \epsilon_1\}$ is linearly independent and is a basis for $\mathbb{R}^3$.

Let W be the subspace of $\mathbb{R}^3$ spanned by ϵ_1. Then $U \cap W = \{\theta\}$ and $U + W = L\{\alpha, \beta, \epsilon_1\} = \mathbb{R}^3$. So W is a complement of U in $\mathbb{R}^3$.

Again $\epsilon_2 = (0,1,0)$ is a vector in $\mathbb{R}^3$ but not in U, because $(0,1,0)$ can not be expressed as $c(1,2,1) + d(2,1,3)$ for real c, d.

Let P be the subspace of $\mathbb{R}^3$ spanned by ϵ_2. Then $U \cap P = \{\theta\}$ and $U + P = L\{\alpha, \beta, \epsilon_2\} = \mathbb{R}^3$. So P is a complement of U in $\mathbb{R}^3$.

Thus P and W are both complements of U in $\mathbb{R}^3$ but $P \neq W$.

4.8. Quotient space.

Let V be a vector space over a field F. Let W be a subspace of V. Let $\alpha \in V$. Then the set $\{\alpha + w : w \in W\}$ is a subset of V. It is called a *coset* of W in V and is denoted by $\alpha + W$.

Theorem 4.8.1. Let W be a subspace of a vector space V over a field F. Let $\alpha, \beta \in V$. Then the cosets $\alpha + W = \beta + W$ if and only if $\alpha - \beta \in W$.

Proof. Let $\alpha + W = \beta + W$ and let $\gamma \in \alpha + W$. Then $\gamma = \alpha + w_1 = \beta + w_2$

for some $w_1, w_2 \in W$. Therefore $\alpha + w_1 = \beta + w_2$ and this implies $\alpha - \beta = w_2 - w_1 \in W$.

Conversely, let $\alpha - \beta \in W$. Then $\alpha = \beta + w_3$ for some $w_3 \in W$, and $\beta = \alpha + w_4$ for some $w_4 \in W$.

Let $\gamma \in \alpha + W$. Then $\gamma = \alpha + w_5$ for some $w_5 \in W$

$$= (\beta + w_3) + w_5 = \beta + w_6, \text{ where } w_6 = w_3 + w_5 \in W.$$

This shows that $\gamma \in \beta + W$ and therefore $\alpha + W \subset \beta + W$.

Let $\delta \in \beta + W$. Then $\delta = \beta + w_7$ for some $w_7 \in W$,

$$= (\alpha + w_4) + w_7 = \alpha + w_8, \text{ where } w_8 = w_4 + w_7 \in W.$$

This shows that $\delta \in \alpha + W$ and therefore $\beta + W \subset \alpha + W$.

It follows that $\alpha + W = \beta + W$. This completes the proof.

The set of all distinct cosets of W is denoted by V/W. Since $(V, +)$ is an additive group, $(W, +)$ is an additive subgroup of V. V/W is an additive group under the composition $+$ defined by $(\alpha + W) + (\beta + W) = (\alpha + \beta) + W$ for all $\alpha, \beta \in V$.

We like to equip the set with a structure of a vector space over F. For this purpose we define scalar multiplication in V/W by $c(\alpha + W) = c\alpha + W$ for all $\alpha \in V$, all $c \in F$.

First we prove that the scalar multiplication is well defined in the sense that if $\alpha + W = \alpha' + W$ then $c\alpha' + W = c\alpha + W$.

Now $\alpha + W = \alpha' + W \Rightarrow \alpha - \alpha' \in W$. Therefore $c\alpha' - c\alpha = c(\alpha' - \alpha) = cw_1 \in W$. Consequently, $c\alpha + W = c\alpha' + W$.

Hence the scalar multiplication is well defined on the set V/W.

It is a matter of simple verification that V/W is a vector space over the field F. The vector space V/W is called the *quotient space*.

Theorem 4.8.2. Let V be a vector space over a field F and W be a subspace of V. Then dim V/W = dim V − dim W.

Proof. Let dim $V = m + n$, dim $W = m$. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ be a basis of W. By extension theorem, let $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$ be a basis of V.

Let $\alpha + W \in V/W$. Then $\alpha \in V$. α can be expressed as $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m + d_1\beta_1 + d_2\beta_2 + \cdots + d_n\beta_n$, where $c_i, d_i \in F$.

$\alpha - (d_1\beta_1 + d_2\beta_2 + \cdots + d_n\beta_n) = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_m\alpha_m \in W$.

$$\begin{aligned} \text{Therefore } \alpha + W &= (d_1\beta_1 + d_2\beta_2 + \cdots + d_n\beta_n) + W \\ &= (d_1\beta_1 + W) + (d_2\beta_2 + W) + \cdots + (d_n\beta_n + W) \\ &= d_1(\beta_1 + W) + d_2(\beta_2 + W) + \cdots + d_n(\beta_n + W). \end{aligned}$$

This shows that $\alpha + W \in L\{\beta_1 + W, \beta_2 + W, \ldots, \beta_n + W\}$.

We prove that the set $\{\beta_1 + W, \beta_2 + W, \ldots, \beta_n + W\}$ is linearly independent. Let us consider the relation

$p_1(\beta_1 + W) + p_2(\beta_2 + W) + \cdots + p_n(\beta_n + W) = \theta + W$, where $p_i \in F$.

$p_1(\beta_1 + W) + p_2(\beta_2 + W) + \cdots + (p_n\beta_n + W) = W$

Then $(p_1\beta_1 + W) + (p_2\beta_2 + W) + \cdots + (p_n\beta_n + W) = W$

or, $(p_1\beta_1 + p_2\beta_2 + \cdots + p_n\beta_n) + W = W$.

Therefore $p_1\beta_1 + p_2\beta_2 + \cdots + p_n\beta_n \in W$.

Since $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ is a basis of W, $(p_1\beta_1 + p_2\beta_2 + \cdots + p_n\beta_n) = q_1\alpha_1 + q_2\alpha_2 + q_m\alpha_m$ for some $q_i \in F$.

This gives $q_1\alpha_1 + q_2\alpha_2 + q_m\alpha_m - p_1\beta_1 - p_2\beta_2 - \cdots - p_n\beta_n = \theta$.

Since $\{\alpha_1, \alpha_2, \ldots, \alpha_m, \beta_1, \beta_2, \ldots, \beta_n\}$ is a linearly independent set, we have $q_1 = q_2 = \cdots = q_m = 0, \quad p_1 = p_2 = \cdots = p_n = 0$.

This proves that the set $\{\beta_1 + W, \beta_2 + W, \ldots, \beta_n + W\}$ is linearly independent. So it is a basis of V/W.

The dimension of $V/W = n = (m+n) - m = \dim V - \dim W$.
This completes the proof.

Worked Example.

1. Let $V = \mathbb{R}^3$ and W be a subspace of V generated by the vectors $(1,0,0), (1,1,0)$. Find a basis of the quotient space V/W. Verify that $\dim V/W = \dim V - \dim W$.

Let $\alpha = (1,0,0), \beta = (1,1,0)$. Since the vectors α, β are linearly independent, $\{\alpha, \beta\}$ is a basis of W. The linearly independent set α, β in V can be extended to a basis of V.

If $\gamma = (0,0,1)$ then the set $\{\alpha, \beta, \gamma\}$ is linearly independent in V and therefore it is a basis of V. A basis of the quotient space V/W is $\gamma + W$ and therefore $\dim V/W = 1$.

Here dim V=3, dim W=2 and dim V/W= 1= dim $V-$ dim W.

Exercises 8

1. $U = L\{(2,0,1), (3,1,0)\}$, $W = L\{(1,0,0), (0,1,0)\}$. Find dim U, dim W, dim $(U \cap W)$ and dim $(U + W)$.

2. Two subspaces of $\mathbb{R}^3$ are $U = \{(x,y,z) : x + y + z = 0\}$ and $w = \{(x,y,z) : x + 2y - z = 0\}$. Find dim U, dim W, dim $U \cap W$, dim $(U + W)$.

3. $U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a + b = 0 \right\}$, $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : c + d = 0 \right\}$ are subspaces of $\mathbb{R}_{2\times 2}$. Find dim U, dim W, dim $(U \cap W)$ and dim $(U + W)$.

4. Let U be the subspace of $\mathbb{R}^3$ generated by the set $\{(1,0,1), (2,0,1)\}$. Find two different subspaces P, Q of $\mathbb{R}^3$ such that $U \oplus P = \mathbb{R}^3, U \oplus Q = \mathbb{R}^3$.

5. Let $U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a = b = 0 \right\}$ and $W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : c = d = 0 \right\}$ be subspaces of $\mathbb{R}_{2\times 2}$. Show that $U \oplus W = \mathbb{R}_{1\times 2}$

6. Let S be the vector space of all $n \times n$ real symmetric matrices and T be the vector space of all $n \times n$ real skew symmetric matrices. Prove that dim $S = \frac{n(n+1)}{2}$ and dim $T = \frac{n(n-1)}{2}$. Hence prove that the space $\mathbb{R}_{n\times n}$ of all $n \times n$ real matrices is the direct sum of S and T.

Deduce that an $n \times n$ real matrix can be expressed as the sum of an $n \times n$ real symmetric matrix and an $n \times n$ real skew symmetric matrix.

7. Let $V = \mathbb{R}^3$ and W be a subspace of V generated by the vectors $(1,0,0), (1,1,0)$. Find a basis of the quotient space V/W. Verify that dim $V/W =$ dim $V-$ dim W.

8. Let $V = \mathbb{R}^4$ and W be a subspace of V generated by the vectors $(1,0,0,0), (1,1,0,0)$. Find a basis of the quotient space V/W. Verify that dim $V/W =$ dim $V-$ dim W.

4.9. Row space and column space of a matrix.

Let A be an $m \times n$ matrix over a field F. Each row of A is an n-tuple vector in F^n. Let the row vectors of A be denoted by $\alpha_1, \alpha_2, \ldots, \alpha_m$. Each column of A is an m-tuple vector in F^m. Let the column vectors be denoted by $\bar{\alpha_1}, \bar{\alpha_2}, \ldots, \bar{\alpha_n}$.

The row vectors of A generate a vector space which is said to be the *row space* of A and is denoted by $R(A)$ and the column vectors generate a vector space which is said to be the *column space* of A and is denoted by $C(A)$. Clearly, $R(A)$ is a subspace of F^n and $C(A)$ is a subspace of F^m.

Row rank and column rank.

Definition. The dimension of $R(A)$ is said to be the *row rank* of A and the dimension of $C(A)$ is said to be the *column rank* of A.

Since $R(A) \subset F^n$, the row rank of $A \leq n$. Since $C(A) \subset F^m$, the column rank of $A \leq m$.

Example.

1. Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 0 & 1 \end{pmatrix}$. The row space of A is the linear span of the row vectors $\alpha_1 = (1,2,0), \alpha_2 = (0,1,3), \alpha_3 = (2,0,1)$.

Therefore the row space of $A = \{c_1\alpha_1 + c_2\alpha_2 + c_3\alpha_3 : c_i \in \mathbb{R}\}$.

The set $\{\alpha_1, \alpha_2, \alpha_3\}$ is linearly independent. Therefore the row rank of A is 3.

The column space of A is the linear span of the column vectors $\bar{\alpha_1} = (1, 0, 2), \bar{\alpha_2} = (2, 1, 0)\}, \bar{\alpha_3} = (0, 3, 1)$.

Therefore the column space of $A = \{c_1\bar{\alpha_1} + c_2\bar{\alpha_2} + c_3\bar{\alpha_3} : c_i \in \mathbb{R}\}$.

The set $\{\bar{\alpha_1}, \bar{\alpha_2}, \bar{\alpha_3}\}$ is linearly independent. Therefore the column rank of A is 3.

Theorem 4.9.1. Let A be an $m \times n$ matrix and P be an $m \times m$ matrix over the same field F. Then the row space of PA is a subspace of the row space of A.

In particular, if P be non-singular, then the matrices A and PA have the same row spaces.

Proof. Let $P = (p_{ij})_{m,m}$, $A = (a_{ij})_{m,n}$.

Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the row vectors of A and $\rho_1, \rho_2, \ldots, \rho_m$ be the row vectors of PA.

Then $\rho_i = (p_{i1}a_{11} + p_{i2}a_{21} + \cdots + p_{im}a_{m1}, p_{i1}a_{12} + p_{i2}a_{22} + \cdots + p_{im}a_{m2}, \ldots\ldots, p_{i1}a_{1n} + p_{i2}a_{2n} + \cdots + p_{im}a_{mn}) =$

$p_{i1}(a_{11}, a_{12}, \ldots, a_{1n}) + p_{i2}(a_{21}, a_{22}, \ldots, a_{2n}) + \cdots + p_{im}(a_{m1}, a_{m2}, \ldots, a_{mn})$
$= p_{i1}\alpha_1 + p_{i2}\alpha_2 + \cdots + p_{im}\alpha_m$.

Each ρ_i is a linear combination of the vectors $\alpha_1, \alpha_2, \ldots \alpha_m$.

Therefore $L\{\rho_1, \rho_2, \ldots, \rho_m\} \subset L\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, by Theorem 4.3.7.

That is, the row space of PA is a subspace of the row space of A...(i)

Second Part. Since P is non-singular, P^{-1} exists.

Considering the product $P^{-1}(PA)$, we have

the row space of $P^{-1}(PA)$ is a subspace of the row space of PA.

That is, the row space of A is a subspace of the row space of PA...(ii)

From (i) and (ii) it follows that the row space of $A =$ the row space of PA.

Corollary 1. If A be an $m \times n$ matrix and P be a non-singular $m \times m$ matrix, the row rank of $A =$ the row rank of PA.

2. If A be an $m \times n$ matrix and P be a non-singular $n \times n$ matrix, the column rank of $A =$ the column rank of AP.

Theorem 4.9.2. Row equivalent matrices have the same row spaces.

Proof. Let A and B be two row equivalent matrices. Then there exist elementary matrices $P_1, P_2, \ldots, P_r$ such that $B = P_1P_2\ldots P_rA$.

Since an elementary matrix is non-singular, the product $P_1P_2\ldots P_r$ is non-singular.

By Theorem 4.9.1, A and B have the same row spaces.

Theorem 4.9.3. Let R be a non-zero row reduced echelon matrix row equivalent to an $m \times n$ matrix A. Then the non-zero row vectors of R form a basis of the row space of A.

Proof. Let $R = (a_{ij})_{m,n}$ and $\alpha_1, \alpha_2, \ldots, \alpha_r$ be the non-zero row vectors of R. Then the row space of R is generated by $\alpha_1, \alpha_2, \ldots, \alpha_r, \theta, \ldots, \theta$ (θ being counted $m - r$ times). The set of generators $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \theta\}$ is linearly dependent as it contains the null vector θ. The null vector θ can be deleted from the generating set, and it is certain that the non-zero vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ generate the row space of R.

We need only to prove linear independence of the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$.

Let $\alpha_i = (a_{i1}, a_{i2}, \ldots, a_{in})$.

Since R is a row reduced echelon matrix, there are positive integers $k_1, k_2, \ldots, k_r$ satisfying the following conditions.

(i) the leading 1 of α_i occurs in column k_i,
(ii) $k_1 < k_2 < \cdots < k_r$,
(iii) $a_{ik_j} = \delta_{ij}$,
(iv) $a_{ij} = 0$ if $j < k_i$.

Let us consider the relation $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r = \theta$, where c_i's are scalars.

Then $c_1(a_{11}, a_{12}, \ldots, a_{1n}) + c_2(a_{21}, a_{22}, \ldots, a_{2n}) + \cdots + c_r(a_{r1}, a_{r2}, \ldots, a_{rn}) = (0, 0, \ldots, 0)$.

Equating k_1th, k_2th, $\ldots$, k_rth components only and noting that

$$\begin{aligned} a_{1k_1} &= 1, a_{2k_1} = 0, \ldots, a_{rk_1} = 0 \\ a_{1k_2} &= 0, a_{2k_2} = 1, \ldots, a_{rk_2} = 0 \\ &\ldots \\ a_{1k_r} &= 0, a_{2k_r} = 0, \ldots, a_{rk_r} = 1, \end{aligned}$$

we have $c_1 = c_2 = \cdots = c_r = 0$.

This proves that the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is linearly independent and therefore it is a basis of the row space of R.

Since R is row equivalent to A, the row space of A is same as that of R and therefore $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is a basis of the row space of A.

Corollary. The row rank of a row reduced echelon matrix R is the number of non-zero rows of R.

Theorem 4.9.4. The row rank of an $m \times n$ matrix A is equal to its determinant rank.

Proof. Let A be row equivalent to a row reduced echelon matrix R having r non-zero rows. Then the row rank of A is r, by Theorem 4.9.3. Again since A is row equivalent to a row reduced echelon matrix having r non-zero rows the determinant rank of A is r, by Theorem 3.6.5.
Thus r = row rank of A = determinant rank of A.

Corollary 1. The column rank of a matrix A is equal to its determinant rank

Proof. By the theorem,
the row rank of A^t = the determinant rank of A^t.
But the row rank of A^t = the column rank of A; and
the determinant rank of A^t = the determinant rank of A.

Therefore the column rank of A is equal to its determinant rank.

Corollary 2. For an $m \times n$ matrix A, the row rank of A = the column rank of A.

Theorem 4.9.5. For an $m \times n$ matrix A, the row rank of A = the column rank of A.

Independent Proof.

Proof. Let $A = (a_{ij})_{m,n}$ and $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the row vectors, $\bar{\alpha_1}, \bar{\alpha_2}, \ldots, \bar{\alpha_n}$ be the column vectors of A.

Let the row rank of $A = r$ and $S = \{\beta_1, \beta_2, \ldots, \beta_r\}$ be a basis of the row space of A where $\beta_i = (b_{i1}, b_{i2}, \ldots, b_{in})$ and $b_{ij} = a_{kj}$ for some k.

Since S is a basis,
$$\begin{aligned} \alpha_1 &= c_{11}\beta_1 + c_{12}\beta_2 + \cdots + c_{1r}\beta_r \\ \alpha_2 &= c_{21}\beta_1 + c_{22}\beta_2 + \cdots + c_{2r}\beta_r \\ &\cdots \\ \alpha_m &= c_{m1}\beta_1 + c_{m2}\beta_2 + \cdots + c_{mr}\beta_r, \end{aligned}$$
where c_{ij} are suitable scalars.

The jth component of α_i is a_{ij} and the jth component of $c_{i1}\beta_1 + c_{i2}\beta_2 + \cdots + c_{ir}\beta_r$ is $c_{i1}b_{1j} + c_{i2}b_{2j} + \cdots + c_{ir}b_{rj}$). This holds for $1, 2, \ldots, m$. Therefore
$$\begin{aligned} a_{1j} &= c_{11}b_{1j} + c_{12}b_{2j} + \cdots + c_{1r}b_{rj}, \\ a_{2j} &= c_{21}b_{1j} + c_{22}b_{2j} + \cdots + c_{2r}b_{rj}, \\ &\cdots \\ a_{mj} &= c_{m1}b_{1j} + c_{m2}b_{2j} + \cdots + c_{2r}b_{rj}. \end{aligned}$$

Let $\begin{pmatrix} c_{11} \\ c_{21} \\ \cdot \\ \cdot \\ c_{m1} \end{pmatrix} = \gamma_1, \begin{pmatrix} c_{12} \\ c_{22} \\ \cdot \\ \cdot \\ c_{m2} \end{pmatrix} = \gamma_2, \ldots, \begin{pmatrix} c_{1r} \\ c_{2r} \\ \cdot \\ \cdot \\ c_{mr} \end{pmatrix} = \gamma_r.$ Then

$$\begin{aligned} \bar{\alpha_1} &= b_{11}\gamma_1 + b_{21}\gamma_2 + \cdots + b_{r1}\gamma_r, \\ \bar{\alpha_2} &= b_{12}\gamma_1 + b_{22}\gamma_2 + \cdots + b_{r2}\gamma_r, \\ &\cdots \\ \bar{\alpha_n} &= b_{1n}\gamma_1 + b_{2n}\gamma_2 + \cdots + b_{rn}\gamma_r. \end{aligned}$$

This shows that each of the column vectors $\bar{\alpha_1}, \bar{\alpha_2}, \ldots, \bar{\alpha_n}$ belongs to the linear span of r vectors $\gamma_1, \gamma_2, \ldots, \gamma_r$ and therefore
the column rank of $A \le r$. ... (i)

Now r = the row rank of A = the column rank of A^t.

Also the column rank of $A^t \le$ the row rank of A^t by just what we have done and the row rank of A^t = the column rank of A. Therefore
$r \le$ column rank of A. ... (ii)

Combining (i) and (ii), the row rank of A = the column rank of A.

Theorem 4.9.6. Let A and B be two matrices over the same field F such that AB is defined. Then rank of $AB \le \min\{$rank of A, rank of $B\}$.

Proof. Let $A = (a_{ij})_{m,n}, B = (b_{ij})_{n,p}$. Let $\beta_1, \beta_2, \ldots, \beta_n$ be the row vectors of B. Let $\rho_1, \rho_2, \ldots, \rho_m$ be the row vectors of AB. Then
$$\begin{aligned} \rho_1 &= a_{11}\beta_1 + a_{12}\beta_2 + \cdots + a_{1n}\beta_n, \\ \rho_2 &= a_{21}\beta_1 + a_{22}\beta_2 + \cdots + a_{2n}\beta_n, \\ &\cdots \\ \rho_m &= a_{m1}\beta_1 + a_{m2}\beta_2 + \cdots + a_{mn}\beta_n. \end{aligned}$$

Each ρ_i is a linear combination of the vectors $\beta_1, \beta_2, \ldots, \beta_n$.
So $L\{\rho_1, \rho_2, \ldots, \rho_m\} \subset L\{\beta_1, \beta_2, \ldots, \beta_n\}$, by Theorem 4.3.7.

Therefore the row space of AB is a subspace of the row space of B.

It follows that row rank of $AB \le$ row rank of B. This implies that rank of $AB \le$ rank of B ... (i)

Considering the product B^tA^t, we deduce that rank of $B^tA^t \le$ rank of A^t. That is, rank of $(AB)^t \le$ rank of A^t. This implies that rank of $AB \le$ rank of A ... (ii)

Combining (i) and (ii), rank of $AB \le \min\{$rank of A, rank of $B\}$.

Corollary 1. If A be non-singular, rank of AB = rank of B.

Proof. Since A is non-singular, A^{-1} exists. Considering the matrix $A^{-1}(AB)$, we have rank of $A^{-1}(AB) \le$ rank of AB.
That is, rank of $B \le$ rank of AB. But rank of $AB \le$ rank of B.

Hence it follows that rank of AB = rank of B.

Corollary 2. If B be non-singular, rank of AB = rank of A.

Theorem 4.9.7. Let A and B be $m \times n$ matrices over a field F. Then rank of $(A+B) \leq$ rank of $A+$ rank of B.

Proof. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the row vectors of A; $\beta_1, \beta_2, \ldots, \beta_m$ be the row vectors of B. Then the row vectors of $A+B$ are $\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_m+\beta_m$.

The vectors $\alpha_1, \alpha_2, \ldots, \alpha_m$ generate $R(A)$, the row space of A; the vectors $\beta_1, \beta_2, \ldots, \beta_m$ generate $R(B)$, the row space of B and the vectors $\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_m+\beta_m$ generate $R(A+B)$, the row space of $A+B$.

Because $R(A)+R(B) = \{u+v : u \in R(A), v \in R(B)\}$, the vectors $\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_m+\beta_m$ lie in the subspace $R(A)+R(B)$.

But $R(A+B)$ is the smallest subspace containing the vectors $\alpha_1+\beta_1, \alpha_2+\beta_2, \ldots, \alpha_m+\beta_m$. So $R(A+B) \subset R(A)+R(B)$.

Hence $\dim R(A+B) \leq \dim [R(A)+R(B)]$ ... (i)

As $R(A)$, $R(B)$ are subspces of a finite dimensional vector space F^n, $\dim [R(A)+R(B)] = \dim R(A) + \dim R(B) - \dim [R(A) \cap R(B)]$.

Therefore $\dim [R(A)+R(B)] \leq \dim R(A) + \dim R(B)$ ... (ii)

From (i) and (ii) it follows that $\dim R(A+B) \leq \dim R(A) + \dim R(B)$, i.e., row rank of $A+B \leq$ row rank of $A+$ row rank of B.

Hence rank of $(A+B) \leq$ rank of $A+$ rank of B.

This completes the proof.

Theorem 4.9.8. Factorisation theorem

An $m \times n$ matrix of rank r can be expressed as the product of two matrices, each of rank r.

Proof. Let A be an $m \times n$ matrix of rank r. hen there exist non-singular matrices P and Q of order m and n respectively such that $PAQ = \begin{pmatrix} I_r & O_{r,n-r} \\ O_{m-r,r} & O_{m-r,n-r} \end{pmatrix} = R$, say.

R is the fully reduced normal form of A. R can be expressed as the product ST where $S = \begin{pmatrix} I_r \\ O_{m-r,r} \end{pmatrix}$, an $m \times r$ matrix of rank r and $T = (I_r, O_{r,n-r})$, an $r \times n$ matrix of rank r.

Since P and Q are non-singular, P^{-1} and Q^{-1} both exist and are non-singular. $PAQ = ST \Rightarrow A = (P^{-1}S)(TQ^{-1})$.

Since P^{-1} is non-singular and the rank of S is r, the row rank of $P^{-1}S$ is r. That is, the rank of $P^{-1}S$ is r.

Since Q^{-1} is non-singular and the rank of T is r, the column rank of TQ^{-1} is r. That is, the rank of TQ^{-1} is r.

Thus A is the product of two matrices $P^{-1}S$ and TQ^{-1}, each of rank r. This completes the proof.

Worked Examples.

1. Determine the row rank and the column rank of the matrix A and verify that the row rank of A = column rank of A, where

$$A = \begin{pmatrix} 2 & 1 & 4 & 3 \\ 3 & 2 & 6 & 9 \\ 1 & 1 & 2 & 6 \end{pmatrix}.$$

Let us apply elementary row operations on A to reduce it to a row echelon matrix.

$$A \xrightarrow{R_{13}} \begin{pmatrix} 1 & 1 & 2 & 6 \\ 3 & 2 & 6 & 9 \\ 2 & 1 & 4 & 3 \end{pmatrix} \xrightarrow[R_3-2R_1]{R_2-3R_1} \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & -1 & 0 & -9 \\ 0 & -1 & 0 & -9 \end{pmatrix}$$

$$\xrightarrow{-R_2} \begin{pmatrix} 1 & 1 & 2 & 6 \\ 0 & 1 & 0 & 9 \\ 0 & -1 & 0 & -9 \end{pmatrix} \xrightarrow[R_3+R_2]{R_1-R_2} \begin{pmatrix} 1 & 0 & 2 & -3 \\ 0 & 1 & 0 & 9 \\ 0 & 0 & 0 & 0 \end{pmatrix} = R, \text{ say.}$$

R is a row echelon matrix. The non-zero row vectors of R are (1,0,2,-3), (0,1,0,9). These form a basis of the row space of A. Therefore the row rank of $A = 2$.

To determine the column rank of A let us apply elementary row operations on the matrix A^t.

$$A^t = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \\ 4 & 6 & 2 \\ 3 & 9 & 6 \end{pmatrix} \xrightarrow{R_{12}} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \\ 4 & 6 & 2 \\ 3 & 9 & 6 \end{pmatrix} \xrightarrow[\substack{R_3-4R_1 \\ R_4-3R_1}]{R_2-2R_1} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -2 & -2 \\ 0 & 3 & 3 \end{pmatrix}$$

$$\xrightarrow{-R_2} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \\ 0 & 3 & 3 \end{pmatrix} \xrightarrow[\substack{R_3+2R_2 \\ R_4-3R_2}]{R_1-2R_2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = B, \text{ say.}$$

B is a row echelon matrix. The non-zero row vectors of B are (1,0,-1), (0,1,1). These form a basis of the row space of A^t, i.e., of the column space of A and consequently the column rank of $A = 2$.

Therefore the row rank of A = the column rank of A.

2. Examine linear dependence of the set of vectors $\{(1,-1,2,4), (2,-1,5,7), (-1,3,1,-2)\}$ in $\mathbb{R}^4$.

Let $\alpha = (1,-1,2,4), \beta = (2,-1,5,7), \gamma = (-1,3,1,-2)$.

Let us consider the matrix A whose row vectors are α, β, γ and apply elementary row operations on A to reduce it to a row echelon matrix.

$$A = \begin{pmatrix} 1 & -1 & 2 & 4 \\ 2 & -1 & 5 & 7 \\ -1 & 3 & 1 & -2 \end{pmatrix} \xrightarrow[R_3+R_1]{R_2-2R_1} \begin{pmatrix} 1 & -1 & 2 & 4 \\ 0 & 1 & 1 & -1 \\ 0 & 2 & 3 & 2 \end{pmatrix}$$

$$\xrightarrow[R_3-2R_2]{R_1+R_2} \begin{pmatrix} 1 & 0 & 3 & 3 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 4 \end{pmatrix} \xrightarrow[R_2-R_3]{R_1-3R_3} \begin{pmatrix} 1 & 0 & 0 & -9 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 4 \end{pmatrix} = R, \text{ say.}$$

R is a row echelon matrix row equivalent to A. There are 3 non-zero rows in R. So the row rank of R, and consequently the row rank of A is 3.

Therefore the set $\{\alpha, \beta, \gamma\}$ generates a vector space of dimension 3 and hence the set $\{\alpha, \beta, \gamma\}$ is linearly independent.

3. Show that the rank of the matrix $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 1 & 0 \end{pmatrix}$ is 2. Express A as the product of two matrices, each of rank 2.

Let us apply elementary operations on A to reduce it to the fully reduced normal form.

$$A \xrightarrow[R_3-2R_1]{R_2-R_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & -1 & -2 \end{pmatrix} \xrightarrow[R_3+R_2]{R_1-R_2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[C_3-2C_2]{C_3-C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$= \begin{pmatrix} I_2 & O_{2,1} \\ O_{1,2} & O_{1,1} \end{pmatrix} = R$, say. R is the fully reduced normal form of A.

$R = E_{32}(1)E_{12}(-1)E_{31}(-2)E_{21}(-1)A\{E_{31}(-1)\}^t\{E_{32}(-2)\}^t = PAQ$,
where $P = E_{32}(1)E_{12}(-1)E_{31}(-2)E_{21}(-1)$, and
$Q = \{E_{31}(-1)\}^t\{E_{32}(-2)\}^t$.

R can also be expressed as the product ST, where $S = \begin{pmatrix} I_2 \\ O_{1,2} \end{pmatrix}$, a 3×2 matrix of rank 2 and $T = (I_2, O_{2,1})$, a 2×3 matrix of rank 2.

$$P = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -3 & 1 & 1 \end{pmatrix}, \quad P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 2 & 1 & 1 \end{pmatrix}.$$

$$Q = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

$PAQ = ST$ gives $A = (P^{-1}S)(TQ^{-1})$.

$$P^{-1}S = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

$$TQ^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Hence $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}$.

Exercises 9

1. Determine the row rank of the matrix having each set as the row vectors and hence examine the linear dependence of the set.
   (i) $\{(4,2,5), (3,0,1), (5,4,9)\}$,
   (ii) $\{(-2,0,0,3), (1,5,3,0), (3,2,1,6), (3,5,3,-3)\}$,
   (iii) $\{(4,1,3,2), (2,4,4,3), (1,2,0,0), (0,1,1,1), (1,3,1,1)\}$.
2. Find a basis for the row space of the matrix.
   (i) $\begin{pmatrix} 0 & 3 & 7 \\ 2 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 0 & 3 & 2 \\ 4 & 0 & 2 & 2 \\ 0 & 3 & 1 & 4 \end{pmatrix}$, (iii) $\begin{pmatrix} 2 & 1 & 3 & 5 \\ 3 & 4 & 1 & 2 \\ 0 & 3 & 1 & 1 \\ 5 & 5 & 4 & 7 \end{pmatrix}$.
3. Find a basis for the column space of the matrix.
   (i) $\begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, (ii) $\begin{pmatrix} 2 & 3 & 1 & 0 \\ 4 & 0 & 6 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$, (iii) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 2 \end{pmatrix}$.
4. $A = \begin{pmatrix} 1 & 2 & 5 \\ 3 & 0 & 7 \\ -1 & 4 & 3 \end{pmatrix}$. Examine if $(1,1,1)$, $(1,-1,1)$ are in
   (i) the row space of A, (ii) the column space of A.
5. If A be a singular matrix prove that
   (i) the row vectors of A are linearly dependent;
   (ii) the column vectors of A are linearly dependent.
6. If A be a non singular matrix prove that
   (i) the row vectors of A are linearly independent;
   (ii) the column vectors of A are linearly independent.
7. A is an $m \times r$ matrix and B is an $r \times n$ matrix.
   (i) If the rank of the matrix AB is m, prove that rank of A is m;
   (ii) if the rank of the matrix AB is n, prove that rank of B is n.

8. If A be a rectangular matrix, prove that either its row vectors or its column vectors or both the sets are linearly dependent.

9. If B is a non-null $m \times 1$ matrix and C is a non-null $1 \times n$ matrix prove that the rank of BC is 1.

10. If an $m \times n$ matrix A be of rank 1 prove that A can be expressed the product BC, where B is a non-null $m \times 1$ matrix and C is a non-null $1 \times n$ matrix.

11. Show that the rank of $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ is 2.

Express A as the product of two matrices, each of rank 2.

4.10. System of Linear Equations.

A system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ is of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots \quad \cdots \quad \cdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \quad \ldots \quad \text{(i)} \end{aligned}$$

where a_{ij}'s and b_i's are elements of a field F, called the field of scalars. a_{ij}'s are called coefficients of the system. In particular, a_{ij}'s and b_i's are real (or complex) numbers when F is the field $\mathbb{R}$ (or $\mathbb{C}$).

An ordered set $(c_1, c_2, \ldots, c_n)$ where $c_i \in F$, is said to be a *solution* of the system (i) if each equation of the system is satisfied by

$$x_1 = c_1, x_2 = c_2, \ldots, x_n = c_n.$$

Therefore a solution of the system can be considered as an n-tuple vector of $V_n(F)$. In particular, if the field of scalars be $\mathbb{R}$, a solution of the system is a vector in $\mathbb{R}^n$.

A system of equations is said to be *consistent* if it has a solution. Otherwise, it is said to be *inconsistent*.

Examples.

1. $x_1 + 2x_2 = 3$
$3x_1 + x_2 = 4.$

(1,1) is a solution. There is no other solution of the system.

2. $x_1 + 2x_2 = 3$
$3x_1 + 6x_2 = 7.$

This system has no solution. This is not a consistent system.

3. $x_1 + 2x_1 - x_3 = 0$
$2x_1 + x_2 - 2x_3 = 0.$

(1, 0, 1) is a solution of the system. (2, 0, 2), (3, 0, 3) are also solutions. In fact $k(1, 0, 1)$ is a solution for each real number k. Thus the system has many solutions.

Matrix representation.

Let $A = (a_{ij})_{m,n}$, $X = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$, $B = \begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$.

Then the system (i) can be expressed as $AX = B$. The matrix A is said to be the *coefficient matrix* of the system and the matrix

$$\bar{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \ldots & & \ldots & \ldots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix},$$ also denoted by (A, b), is said to be the *augmented matrix* of the system.

The system $AX = B$ is said to be a *homogeneous system* if $B = O$; otherwise, a *non-homogeneous* system.

Two systems $AX = B$ and $CX = D$ are said to be *equivalent systems* if the augmented matrices (A, b) and (C, d) be row equivalent.

Theorem 4.10.1. Let $AX = B$ and $RX = S$ be two equivalent systems and α be a solution of $AX = B$. Then α is also a solution of $RX = S$.

Proof. Let the equations of the system $AX = B$ be

$$\begin{aligned} f_1 &\equiv a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n - b_1 = 0 \\ f_2 &\equiv a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n - b_2 = 0 \\ &\cdots \quad \cdots \\ f_m &\equiv a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n - b_m = 0 \end{aligned}$$

and let $\alpha = (c_1, c_2, \ldots, c_n)$ be a solution of the system.

Let us apply elementary row operation R_{ij} on the augmented matrix (A, b) of the system. Then the ith and the jth equations f_i and f_j, are interchanged and the others remain unchanged. Therefore $(c_1, c_2, \ldots, c_n)$ is also a solution of the new system.

This implies that if $R_{ij}(A, b) = (C, d)$, then $(c_1, c_2, \ldots, c_n)$ is also a solution of the system $CX = D$.

Let us apply elementary row operation cR_i on the augmented matrix (A, b) of the system. Then the ith equation f_i, is multiplied by $c(\neq 0)$ and the other equations remain unchanged. Therefore $(c_1, c_2, \ldots, c_n)$ is also a solution of the new system.

This implies that if $cR_i(A, b) = (C, d)$, then $(c_1, c_2, \ldots, c_n)$ is also a solution of the system $CX = D$.

Let us apply elementary row operation $R_i + cR_j$ on the augmented matrix (A, b) of the system. Then the ith equation f_i is replaced by $f_i + cf_j$ and the other equations remain unchanged. Therefore $(c_1, c_2, \ldots, c_n)$ is also a solution of the new system.

This implies that if $R_i + cR_j(A, b) = (C, d)$, then $(c_1, c_2, \ldots, c_n)$ is also a solution of the system $CX = D$.

Since (R, s) can be obtained from (A, b) by a finite number of elementary row operations of the above types, a solution of the system $AX = B$ is also a solution of the system $RX = S$.

Corollary. If one of the two equivalent systems be inconsistent, the other is also so.

To examine the solvability of the system $AX = B$, or to determine the solutions (or solution) of the system, when it is consistent, the obvious procedure is to apply such elementary row operations on the augmented matrix (A, b) as will reduce it to a row reduced echelon matrix.

Worked Examples.

1. Solve the system of equations

$$\begin{aligned} x_1 + x_2 &= 4 \\ x_2 - x_3 &= 1 \\ 2x_1 + x_2 + 4x_3 &= 7. \end{aligned}$$

This is a non-homogeneous system. The coefficient matrix of the system is $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 2 & 1 & 4 \end{pmatrix}$ and the augmented matrix is $\bar{A} = \begin{pmatrix} 1 & 1 & 0 & 4 \\ 0 & 1 & -1 & 1 \\ 2 & 1 & 4 & 7 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$ to reduce it to a row reduced echelon matrix.

$$\bar{A} \xrightarrow{R_3 - 2R_1} \begin{pmatrix} 1 & 1 & 0 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 4 & -1 \end{pmatrix} \xrightarrow[R_3 + R_2]{R_1 - R_2} \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 3 & 0 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{3}R_3} \begin{pmatrix} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow[R_2 + R_3]{R_1 - R_3} \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Hence the system is equivalent to

$$\begin{aligned} x_1 &= 3 \\ x_2 &= 1 \\ x_3 &= 0 \end{aligned}$$

and therefore the solution is $(3, 1, 0)$.

Note. The coefficient matrix A is row equivalent to the identity matrix I_3 and so it is non-singular. This also suggests that the system admits of a unique solution.

2. Solve the system of equations

$$\begin{aligned} x_1 + 3x_2 + x_2 &= 0 \\ 2x_1 - x_2 + x_3 &= 0. \end{aligned}$$

This is a non-homogeneous system. The coefficient matrix of the system is $A = \begin{pmatrix} 1 & 3 & 1 \\ 2 & -1 & 1 \end{pmatrix}$ and the augmented matrix is $\bar{A} = \begin{pmatrix} 1 & 3 & 1 & 0 \\ 2 & -1 & 1 & 0 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$ to reduce it to a row reduced echelon matrix.

$$\begin{pmatrix} 1 & 3 & 1 & 0 \\ 2 & -1 & 1 & 0 \end{pmatrix} \xrightarrow{R_2 - 2R_1} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & -7 & -1 & 0 \end{pmatrix}.$$

$$\xrightarrow{-\frac{1}{7}R_2} \begin{pmatrix} 1 & 3 & 1 & 0 \\ 0 & 1 & \frac{1}{7} & 0 \end{pmatrix} \xrightarrow{R_1 - 3R_2} \begin{pmatrix} 1 & 0 & \frac{4}{7} & 0 \\ 0 & 1 & \frac{1}{7} & 0 \end{pmatrix}.$$

The given system is equivalent to

$$\begin{aligned} x_1 + \tfrac{4}{7}x_3 &= 0 \\ x_2 + \tfrac{1}{7}x_3 &= 0. \end{aligned}$$

Assigning to x_3 an arbitrary real number c, we have the solution $x_1 = -\frac{4}{7}c$, $x_2 = -\frac{1}{7}c$, $x_3 = c$.

Therefore the solution is $(-\frac{4}{7}c, -\frac{1}{7}c, c)$, i.e., $c(-\frac{4}{7}, -\frac{1}{7}, 1)$, where c is an arbitrary real number. The solution can be also equivalently expressed as $k(4, 1, -7)$, where k is an arbitrary real number.

Note 1. Instead of considering the augmented matrix we can consider only the coefficient matrix A in the case of a homogeneous system $AX = 0$, since the last column of the augmented matrix is the zero column and this column remains unchanged under elementary row operations.

Note 2. Since k is an arbitrary real number, the number of solutions of the system is infinite. $\alpha = (4, 1, -7)$ is a solution and all solutions are of the form $k\alpha$, where $k \in \mathbb{R}$. So the solutions form a subspace of $\mathbb{R}^3$ and the dimension of this subspace is 1.

3. Solve, if possible, the system of equations

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 10 \\ -x_1 + x_2 + 2x_3 &= 2 \\ 2x_1 + x_2 - 3x_3 &= 2. \end{aligned}$$

This is a non-homogeneous system. The coefficient matrix of the system is $A = \begin{pmatrix} 1 & 2 & -1 \\ -1 & 1 & 2 \\ 2 & 1 & -3 \end{pmatrix}$ and the augmented matrix is $\bar{A} = \begin{pmatrix} 1 & 2 & -1 & 10 \\ -1 & 1 & 2 & 2 \\ 2 & 1 & -3 & 2 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$ to reduce it to a row reduced echelon matrix.

$$\begin{pmatrix} 1 & 2 & -1 & 10 \\ -1 & 1 & 2 & 2 \\ 2 & 1 & -3 & 2 \end{pmatrix} \xrightarrow[R_3-2R_1]{R_2+R_1} \begin{pmatrix} 1 & 2 & -1 & 10 \\ 0 & 3 & 1 & 12 \\ 0 & -3 & -1 & -18 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{3}R_2} \begin{pmatrix} 1 & 2 & -1 & 10 \\ 0 & 1 & \frac{1}{3} & 4 \\ 0 & -3 & -1 & -18 \end{pmatrix} \xrightarrow[R_3+3R_2]{R_1-2R_2} \begin{pmatrix} 1 & 0 & -\frac{5}{3} & 2 \\ 0 & 1 & \frac{1}{3} & 4 \\ 0 & 0 & 0 & -6 \end{pmatrix}.$$

The given system is equivalent to

$$\begin{aligned} x_1 - \tfrac{5}{3}x_3 &= 2 \\ x_2 + \tfrac{1}{3}x_3 &= 4 \\ 0 &= -6. \end{aligned}$$

The last equation disallows the existence of any solution of the equivalent system. Therefore the given system is inconsistent.

4. Solve, if possible, the system of equations

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 10 \\ -x_1 + x_2 + 2x_3 &= 2 \\ 2x_1 + x_2 - 3x_3 &= 8. \end{aligned}$$

This is a non-homogeneous system. Let us apply elementary row operations on the augmented matrix of the system to reduce it to a row reduced echelon matrix.

$$\begin{pmatrix} 1 & 2 & -1 & 10 \\ -1 & 1 & 2 & 2 \\ 2 & 1 & -3 & 8 \end{pmatrix} \xrightarrow[R_3-2R_1]{R_2+R_1} \begin{pmatrix} 1 & 2 & -1 & 10 \\ 0 & 3 & 1 & 12 \\ 0 & -3 & -1 & -12 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{3}R_2} \begin{pmatrix} 1 & 2 & -1 & 10 \\ 0 & 1 & \frac{1}{3} & 4 \\ 0 & -3 & -1 & -12 \end{pmatrix} \xrightarrow[R_3+3R_2]{R_1-2R_2} \begin{pmatrix} 1 & 0 & -\frac{5}{3} & 2 \\ 0 & 1 & \frac{1}{3} & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The given system is equivalent to

$$\begin{aligned} x_1 - \tfrac{5}{3}x_3 &= 2 \\ x_2 + \tfrac{1}{3}x_3 &= 4. \end{aligned}$$

Assigning to x_3 an arbitrary real number c, the solution is $(\frac{5}{3}c + 2, -\frac{1}{3}c + 4, c)$. This can be expressed as $c(\frac{5}{3}, -\frac{1}{3}, 1) + (2, 4, 0)$, where c is an arbitrary real number.

Note 1. Since c is arbitrary, the number of solutions is infinite.

Note 2: $(2, 4, 0)$ is a *particular solution* of the system. $c(\frac{5}{3} - \frac{1}{3}, 1)$ is the general solution of the associated homogeneous system.

Homogeneous System.

Let the system of equations be $AX = O$, where A is an $m \times n$ matrix over a field F. The system is necessarily a consistent one, since $(0, 0, \ldots, 0)$ is always a solution of the system. This solution (the zero solution) is said to be the *trivial solution* of the system.

Our main interest will be in the non-zero solutions, if there be any, of the system.

Theorem 4.10.2. The solutions of a homogeneous system $AX = O$ in n unknowns where A is an $m \times n$ matrix over a field F, form a subspace of $V_n(F)$, the vector space of all ordered n-tuples $\{(a_1, a_2, \ldots, a_n) : a_i \in F\}$.

Proof. The system being always a consistent system has a solution which is an n-tuple vector in $V_n(F)$. Let S be the set of all solutions of the system.

Case 1. The zero solution is the only solution of the system.

Then $S = \{\theta\}$ and this is a subspace of $V_n(F)$.

Case 2. The system has many solutons.

Let $A = (a_{ij})_{m,n}$, $\alpha = (c_1, c_2, \ldots, c_n) \in S$ and $c \in F$.

Since α is a solution of the system, we have

$a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n = 0$ for $i = 1, 2, \ldots, m$.

Therefore $a_{i1}(cc_1) + a_{i2}(cc_2) + \cdots + a_{in}(cc_n)$
$= c(a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n) = 0$ for $i = 1, 2, \ldots, m$.

This shows that $(cc_1, cc_2, \ldots, cc_n)$ is a solution of the system.

Therefore $\alpha \in S \Rightarrow c\alpha \in S$ ... (i)

Let $\alpha = (c_1, c_2, \ldots, c_n)$, $\beta = (d_1, d_2, \ldots, d_n) \in S$.

Since α, β are solutions of the system,

$$a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n = 0 \text{ and}$$
$$a_{i1}d_1 + a_{i2}d_2 + \cdots + a_{in}d_n = 0 \text{ for } i = 1, 1, 2, \ldots, m.$$

Therefore $a_{i1}(c_1 + d_1) + a_{i2}(c_2 + d_2) + \cdots + a_{in}(c_n + d_n)$

$$= (a_{i1}c_1 + a_{i2}c_2 + \cdots + a_{in}c_n) + (a_{i1}d_1 + a_{i2}d_2 + \cdots + a_{in}d_n)$$
$$= 0 \text{ for } i = 1, 2, \ldots, m.$$

This shows that $(c_1 + d_1, c_2 + d_2, \ldots, c_n + d_n)$ is a solution of the system. Therefore $\alpha \in S, \beta \in S \Rightarrow \alpha + \beta \in S \quad \ldots \quad$ (ii)

From (i) and (ii) it follows that S is a subspace of $V_n(F)$.

This completes the proof.

Note. The subspace of solutions of the homogeneous system $AX = O$ is denoted by $X(A)$.

Theorem 4.10.3. Let $AX = O$ be a homogeneous system of n unknowns and $X(A)$ be the solution space of the system. Then

rank of A+ rank of $X(A) = n$.

Proof. Let rank of $A = r$. Then A has r independent column vectors. Without loss of generality, we can assume that first r column vectors $\alpha_1, \alpha_2, \ldots, \alpha_r$ of A are linearly independent. Then the remaining column vectors $\alpha_{r+1}, \ldots, \alpha_n$ can be expressed as

$$\begin{aligned}
\alpha_{r+1} &= e_{11}\alpha_1 + e_{12}\alpha_2 + \cdots + e_{1r}\alpha_r \\
\alpha_{r+2} &= e_{21}\alpha_1 + e_{22}\alpha_2 + \cdots + e_{2r}\alpha_r \\
&\ldots \quad \ldots \\
\alpha_n &= e_{n-r1}\alpha_1 + e_{n-r2}\alpha_2 + \cdots + e_{n-rr}\alpha_r,
\end{aligned}$$

for suitable scalars e_{ij}.

Equivalently,
$$\begin{aligned}
e_{11}\alpha_1 + e_{12}\alpha_2 + \cdots + e_{1r}\alpha_r - \alpha_{r+1} &= 0 \\
e_{21}\alpha_1 + e_{22}\alpha_2 + \cdots + e_{2r}\alpha_r - \alpha_{r+2} &= 0 \\
\ldots \quad &\ldots \\
e_{n-r1}\alpha_1 + e_{n-r2}\alpha_2 + \cdots + e_{n-rr}\alpha_r - \alpha_n &= 0.
\end{aligned}$$

The relations show that

$$\begin{aligned}
\xi_1 &= (e_{11}, e_{12}, \ldots, e_{1r}, -1, 0, 0, \ldots, 0), \\
\xi_2 &= (e_{21}, e_{22}, \ldots, e_{2r}, 0, -1, 0, \ldots, 0), \\
&\ldots \quad \ldots \\
\xi_{n-r} &= (e_{n-r1}, e_{n-r2}, \ldots, e_{n-rr}, 0, 0, \ldots, -1)
\end{aligned}$$

are solutions of the system.

But $\xi_1, \xi_2, \ldots, \xi_{n-r}$ are linearly independent, because $c_1\xi_1 + c_2\xi_2 + \cdots + c_{n-r}\xi_{n-r} = \theta$ implies $c_1 = c_2 = \cdots = c_{n-r} = 0$.

Let $\xi = (d_1, d_2, \ldots, d_r, \ldots, d_n)$ be any solution of the system.

Then $d_1\alpha_1 + d_2\alpha_2 + \cdots + d_r\alpha_r + d_{r+1}(e_{11}\alpha_1 + e_{12}\alpha_2 + \cdots + e_{1r}\alpha_r) + \cdots + d_n(e_{n-r1}\alpha_1 + e_{n-r2}\alpha_2 + \cdots + e_{n-rr}\alpha_r) = \theta$

or, $(d_1 + d_{r+1}e_{11} + d_{r+2}e_{21} + \cdots + d_n e_{n-r1})\alpha_1 + (d_2 + d_{r+1}e_{12} + d_{r+2}e_{22} + \cdots + d_n e_{n-r2})\alpha_2 + \cdots\cdots + (d_r + d_{r+1}e_{1r} + \cdots + d_n e_{n-rr})\alpha_r = \theta$.

Since $\alpha_1, \alpha_2, \ldots, \alpha_r$ are linearly independent,

$$\begin{aligned}
d_1 &= -d_{r+1}e_{11} - d_{r+2}e_{21} - \cdots - d_n e_{n-r1} \\
d_2 &= -d_{r+1}e_{12} - d_{r+2}e_{22} - \cdots - d_n e_{n-r2} \\
&\ldots \quad \ldots \quad \ldots \\
d_r &= -d_{r+1}e_{1r} - d_{r+2}e_{2r} - \cdots - d_n e_{n-rr}.
\end{aligned}$$

$$\begin{aligned}
\text{Therefore } \xi &= (d_1, d_2, \ldots, d_r, d_{r+1}, \ldots, d_n) \\
&= -d_{r+1}(e_{11}, e_{12}, \ldots, e_{1r}, -1, 0, 0, \ldots, 0) \\
&\quad -d_{r+2}(e_{21}, e_{22}, \ldots, e_{2r}, 0, -1, \ldots, 0) \\
&\quad - \cdots - d_n(e_{n-r1}, e_{n-r2}, e_{n-rr}, 0, 0, \ldots, -1) \\
&= -d_{r+1}\xi_1 - d_{r+2}\xi_2 - \cdots - d_n\xi_{n-r}.
\end{aligned}$$

This shows that any solution vector ξ is a linear combination of $\xi_1, \xi_2, \ldots, \xi_{n-r}$ and since these solution vectors $\xi_1, \xi_2, \ldots, \xi_{n-r}$ are linearly independent, the rank (dimension) of the solution space is $n - r$.

Therefore rank of A +rank of $X(A) = r + (n - r) = n$ and the theorem is done.

Note. $\xi_1, \xi_2, \ldots, \xi_{n-r}$ are the basis vectors of the solution space of the homogeneous system . So any solution can be expressed as $c_1\xi_1 + c_2\xi_2 + \cdots + c_{n-r}\xi_{n-r}$, where c_i's are arbitrary scalars. This is called the *general solution* of the homogeneous system.

Corollary. If the number of equations be less than the number of unknowns in a homogeneous system $AX = O$, then the system admits of a non-zero solution.

Proof. Let the order of A be $m \times n$. Then $m < n$ and rank of $A < n$.

As rank of A + rank of $X(A) = n$, we have rank of $X(A) > 0$ and this proves that there is a non-zero solution of the system.

Theorem 4.10.4. The homogeneous system $AX = O$ containing n equations in n unknowns has a non-zero solution if and only if rank of $A < n$.

Proof. Let $(c_1, c_2, \ldots, c_n)$ be a non-zero solution of the system.

Let $A = (a_{ij})_{n,n}$. Then
$$\begin{aligned}
a_{11}c_1 + a_{12}c_2 + \cdots + a_{1n}c_n &= 0 \\
a_{21}c_1 + a_{22}c_2 + \cdots + a_{2n}c_n &= 0 \\
\ldots \quad &\ldots \\
a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nn}c_n &= 0.
\end{aligned}$$

Since $(c_1, c_2, \ldots, c_n)$ is non-zero, at least one of the components, say c_j, is non-zero.

$$c_j.\det A = \begin{vmatrix} a_{11} & a_{12} & \cdots & c_j a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & c_j a_{2j} & \cdots & a_{2n} \\ & & \cdots & & \cdots & \\ a_{n1} & a_{n2} & \cdots & c_j a_{nj} & \cdots & a_{nn} \end{vmatrix} =$$

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & c_1a_{11} + c_2a_{12} + \cdots + c_ja_{ij} + \cdots + c_na_{1n} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & c_1a_{21} + c_2a_{22} + \cdots + c_ja_{2j} + \cdots + c_na_{2n} & \cdots & a_{2n} \\ & \cdots & & & \cdots & \\ a_{n1} & a_{n2} & \cdots & c_1a_{n1} + c_2a_{n2} + \cdots + c_ja_{nj} + \cdots + c_na_{nn} & \cdots & a_{nn} \end{vmatrix}$$

$$[C_j' = c_1C_1 + c_2C_2 + \cdots + c_{j-1}C_{j-1} + C_j + c_{j+1}C_{j+1} + \cdots + c_nC_n]$$

$= 0$, since the jth column is the zero column.

Since $c_j \neq 0$, $\det A = 0$ and this implies that rank of $A < n$.

Conversely, let rank of $A < n$. Let $X(A)$ be the solution space of the homogeneous system. Then rank of $X(A)$ + rank of $A = n$. Since rank of $A < n$, it follows that rank of $X(A) > 0$. This proves that there is a non-zero solution of the system.

Note. The number of solutions in this case is infinite.

Method of solution of a homogeneous system.

Let the system of equations be $AX = O$. The matrix A can be reduced to a row-echelon matrix R by elementary row operations on A. If the rank of A be r, then R has r non-zero rows. The leading 1's of the non-zero rows appear in columns $k_1, k_2, \ldots, k_r$ where $k_1 < k_2 < \cdots < k_r$. By suitable interchange of columns (i.e., by suitably renaming the unknowns), R takes the form

$$\begin{bmatrix} 1 & 0 & \ldots & 0 & b_{1r+1} & b_{1r+2} & \ldots & b_{1n} \\ 0 & 1 & \ldots & 0 & b_{2r+1} & b_{2r+2} & \ldots & b_{2n} \\ & & \ldots & & \ldots & & \ldots & \\ 0 & 0 & \ldots & 1 & b_{rr+1} & b_{rr+2} & \ldots & b_{rn} \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ & & \ldots & & \ldots & & \ldots & \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Therefore, by suitable adjustment of renaming the unknowns, if necessary, the equivalent system is

$$\begin{aligned} x_1 + b_{1r+1}x_{r+1} + b_{1r+2}x_{r+2} + \cdots + b_{1n}x_n &= 0 \\ x_2 + b_{2r+1}x_{r+1} + b_{2r+2}x_{r+2} + \cdots + b_{2n}x_n &= 0 \\ \ldots \quad \ldots \quad \ldots & \\ x_r + b_{rr+1}x_{r+1} + b_{rr+2}x_{r+2} + \ldots + b_{rn}x_n &= 0. \end{aligned}$$

A solution can be obtained by choosing arbitrary scalars for $x_{r+1}, x_{r+2}, \ldots, x_n$.

Let $x_{r+1} = -c_1, x_{r+2} = -c_2, \ldots, x_n = -c_{n-r}$.

Then the general solution of the system is

$(c_1b_{1r+1} + c_2b_{1r+2} + \cdots + c_{n-r}b_{1n}, c_1b_{2r+1} + c_2b_{2r+1} + \cdots + c_{n-r}b_{2n}, \ldots, c_1b_{rr+1} + c_2b_{rr+2} + \cdots + c_{n-r}b_{rn}, -c_1, -c_2, \ldots, -c_{n-r})$

$= c_1(b_{1r+1}, b_{2r+1}, \ldots, b_{rr+1}, -1, 0, 0, \ldots, 0)$
$+ c_2(b_{1r+2}, b_{2r+2}, \ldots, b_{rr+2}, 0, -1, 0, \ldots, 0)$
$\ldots \quad \ldots \quad \ldots$
$+$
$+ c_{n-r}(b_{1n}, b_{2n}, \ldots, b_{rn}, 0, 0, 0, \ldots, -1)$, where $c_1, c_2, \ldots, c_{n-r}$ are arbitrary scalars.

Worked Example (continued).

5. Solve the system of equations

$$\begin{aligned} x + 2y + z - 3w &= 0 \\ 2x + 4y + 3z + w &= 0 \\ 3x + 6y + 4z - 2w &= 0. \end{aligned}$$

This is a homogeneous system. Let A be the coefficient matrix of the system.

Let us apply elementary row operations on A to reduce it to a row reduced echelon matrix.

$$A \xrightarrow[R_3 - 3R_1]{R_2 - 2R_1} \begin{pmatrix} 1 & 2 & 1 & -3 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 1 & 7 \end{pmatrix} \xrightarrow[R_3 - R_2]{R_1 - R_2} \begin{pmatrix} 1 & 2 & 0 & -10 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The given system is equivalent to

$$\begin{aligned} x + 2y - 10w &= 0 \\ z + 7w &= 0. \end{aligned}$$

Choosing $y = c, w = d$, where c, d are arbitrary real numbers, the solution is $(-2c + 10d, c, -7d, d) = c(-2, 1, 0, 0) + d(10, 0, -7, 1)$.

Note. The solutions form a vector space generated by the vectors $(-2, 1, 0, 0)$ and $(10, 0, -7, 1)$ which are linearly independent. So the rank of the solution space is 2. The rank of the matrix A is 2. Therefore rank of A + rank of $X(A) = 4$ (the number of unknowns).

Non-homogeneous system.

We have seen that a non-homogeneous system may not have a solution.

First of all, we discuss the solvability of a non-homogeneous system.

Theorem 4.10.5. A necessary and sufficient condition for a non-homogeneous system $AX = B$ to be consistent is

$$\text{rank of } A = \text{rank of } \bar{A},$$

where $\bar{A}$ is the augmented matrix of the system.

Proof. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the column vectors of A and $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta$ be those of the augmented matrix (A, b).

Suppose that there exists a solution $(c_1, c_2, \ldots, c_n)$ of the system. Then $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \beta$ ... ... (i)

Let $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$, $T = \{\alpha_1, \alpha_2, \ldots, \alpha_n, \beta\}$.

Since $S \subset T$, we have $L(S) \subset L(T)$.

Using (i), we can say that each element of T is a linear combination of the vectors of S. Therefore $L(T) \subset L(S)$.

Consequently, $L(S) = L(T)$, i.e., the column space of A = the column space of $\bar{A}$.

Therefore the column rank of A = the column rank of $\bar{A}$.

Consequently, rank of A = rank of $\bar{A}$.

Conversely, let rank of $\bar{A}$ = rank of $A = r$.

Then r columns of A are linearly independent. Without loss of generality, we can assume $\alpha_1, \alpha_2, \ldots, \alpha_r$ are linearly independent. Since rank of $\bar{A} = r$, these column vectors are also linearly independent column vectors of $\bar{A}$ and β is a linear combination of $\alpha_1, \alpha_2, \ldots, \alpha_r$.

Hence $\beta = d_1\alpha_1 + d_2\alpha_2 + \cdots + d_r\alpha_r$ for some scalars d_i.

It follows that $(d_1, d_2, \ldots, d_r, 0, 0, \ldots, 0)$ is a solution of the system. Hence the system is consistent.

This completes the proof.

Note. The solutions of a consistent non-homogeneous system $AX = B$ do not form a vector space because $(0, 0, \ldots, 0)$ is not a solution.

Theorem 4.10.6. If the non-homogeneous system $AX = B$ possesses a solution X_0 then all solutions of the system are obtained by adding X_0 to the general solution of the associated homogeneous system $AX = 0$.

Proof. Let S be the set of all solutions of the system $AX = B$ and T be the set of all solutions of the associated homogeneous system $AX = 0$.

Let $Y \in T$. Then $AY = O$.

$A(Y + X_0) = AY + AX_0 = B$. This shows that $Y + X_0 \in S$.

Let $Z \in S$. Then $AZ = B$ and also $AX_0 = B$.

$A(Z - X_0) = AZ - AX_0 = O$. This shows that $Z - X_0 \in T$.

Thus $Y \in T \Rightarrow Y + X_0 \in S$, and $Z \in S \Rightarrow Z - X_0 \in T$.

This completes the proof.

Corollary. If the non-homogeneous system $AX = B$ be consistent, the system possesses only one solution or infinitely many solutions according as the associated homogeneous system possesses only the zero solution or infinitely many solutions.

Existence and number of solutions of the non-homogeneous system $AX = B$, where A is an $m \times n$ matrix.

Case 1. $m = n$.

The system is consistent if and only if rank of A = rank of $\bar{A}$. For a consistent system, two cases arise.

Subcase (i). Rank of A = rank of $\bar{A} = n$.

Here A is non singular and so A^{-1} exists.

The system possesses the unique solution $X = A^{-1}B$.

Subcase (ii). Rank of A = rank of $\bar{A} < n$.

The associated homogeneous system $AX = O$ has infinitely many solutions and therefore the system $AX = B$ possesses infinitely many solutions.

Case 2. $m < n$.

The system is consistent if and only if rank of A = rank of $\bar{A} \leq m$. If consistent, rank of A = rank of $\bar{A} < n$.

In the consistent case the homogeneous system $AX = O$ has infinitely many solutions and therefore the system $AX = B$ possesses infinitely many solutions.

Case 3. $m > n$.

The system is consistent if and only if rank of A = rank of $\bar{A} \leq n$. For a consistent system, two cases arise.

Subcase (i). Rank of A = rank of $\bar{A} = n$.

Let $X(A)$ be the solution space of the homogeneous system $AX = O$.

Then rank of $X(A)$ + rank of $A = n$ gives rank of $X(A) = 0$. The system $AX = O$ possesses only the zero solution and therefore the system $AX = B$ possesses only one solution.

Subcase (ii). Rank of A = rank of $\bar{A} < n$.

The associated homogeneous system $AX = O$ possesses infinitely many solutions and therefore the system $AX = B$ possesses infinitely many solutions.

Method of solution of a non-homogeneous system.

Let the system of equations be $AX = B$ where A is an $m \times n$ matrix and let rank of A = rank of $\bar{A} = r$.

Let $\bar{A}$ be reduced by elementary row operations to the row echelon matrix $\bar{R}$ which by suitable interchange of columns takes the form

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & b_{1r+1} & b_{1r+2} & \cdots & b_{1n} & d_1 \\ 0 & 1 & \cdots & 0 & b_{2r+1} & b_{2r+2} & \cdots & b_{2n} & d_2 \\ \cdots & & \cdots & & \cdots & & \cdots & & \\ 0 & 0 & \cdots & 1 & b_{rr+1} & b_{rr+2} & \cdots & b_{rn} & d_r \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & & \cdots & & \cdots & & \cdots & & \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Therefore, by suitable adjustment of renaming the unknowns, if necessary, the equivalent system is

$$\begin{aligned} x_1 + b_{1r+1}x_{r+1} + b_{1r+2}x_{r+2} + \cdots + b_{1n}x_n &= d_1 \\ x_2 + b_{2r+1}x_{r+1} + b_{2r+2}x_{r+2} + \cdots + b_{2n}x_n &= d_2 \\ \cdots \quad \cdots \quad \cdots \quad \cdots & \quad \cdots \\ x_r + b_{rr+1}x_{r+1} + b_{rr+2}x_{r+2} + \cdots + b_{rn}x_n &= d_r. \end{aligned}$$

A solution can be obtained by choosing arbitrary scalars for $x_{r+1}, x_{r+2}, \ldots, x_n$. Let $x_{r+1} = -c_1, x_{r+2} = -c_2, \cdots, x_n = -c_{n-r}$.

Then the general solution of the system is

$$\begin{aligned} &(d_1 + c_1b_{1r+1} + c_2b_{1r+2} + \cdots + c_{n-r}b_{1n}, \\ &d_2 + c_1b_{2r+1} + c_2b_{2r+2} + \cdots + c_{n-r}b_{2n}, \ldots, \\ &d_r + c_1b_{rr+1} + c_2b_{rr+2} + \cdots + c_{n-r}b_{rn}, -c_1, -c_2, \ldots, -c_{n-r}) \\ &= (d_1, d_2, \ldots, d_r, 0, 0, \ldots, 0) \\ &+c_1(b_{1r+1}, b_{2r+1}, \ldots, b_{rr+1}, -1, 0, 0, \ldots, 0) \\ &+c_2(b_{1r+2}, b_{2r+2}, \ldots, b_{rr+2}, 0, -1, 0, \ldots, 0) \\ &+ \quad \cdots \quad \cdots \quad \cdots \\ &+c_{n-r}(b_{1n}, b_{2n}, \ldots, b_{rn}, 0, 0, 0, \ldots, -1), \end{aligned}$$

where $c_1, c_2, \ldots, c_{n-r}$ are arbitrary scalars.

Note. The solution $(d_1, d_2, \ldots, d_r, 0, 0, \ldots, 0)$ is a particular solution of the system, obtained by taking $x_{r+1} = x_{r+2} = \cdots = x_n = 0$.

Worked Examples (continued).

6. Solve, if possible.

(i) $$\begin{aligned} x + 2y + z - 3w &= 1 \\ 2x + 4y + 3z + w &= 3 \\ 3x + 6y + 4z - 2w &= 5. \end{aligned}$$

(ii) $$\begin{aligned} x + 2y + z - 3w &= 1 \\ 2x + 4y + 3z + w &= 3 \\ 3x + 6y + 4z - 2w &= 4. \end{aligned}$$

(i) This is a non-homogeneous system.

The coefficient matrix of the system is $A = \begin{pmatrix} 1 & 2 & 1 & -3 \\ 2 & 4 & 3 & 1 \\ 3 & 6 & 4 & -2 \end{pmatrix}$ and the augmented matrix is $\bar{A} = \begin{pmatrix} 1 & 2 & 1 & -3 & 1 \\ 2 & 4 & 3 & 1 & 3 \\ 3 & 6 & 4 & -2 & 5 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$.

$$\bar{A} \xrightarrow[R_3-3R_1]{R_2-2R_1} \begin{pmatrix} 1 & 2 & 1 & -3 & 1 \\ 0 & 0 & 1 & 7 & 1 \\ 0 & 0 & 1 & 7 & 2 \end{pmatrix} \xrightarrow[R_3-R_2]{R_1-R_2} \begin{pmatrix} 1 & 2 & 0 & -10 & 0 \\ 0 & 0 & 1 & 7 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Here rank of $\bar{A} = 3$ and rank of $A = 2$. Since rank of $\bar{A} \neq$ rank of A, the system is inconsistent.

(ii) This is a non-homogeneous system. The coefficient matrix of the system is $A = \begin{pmatrix} 1 & 2 & 1 & -3 \\ 2 & 4 & 3 & 1 \\ 3 & 6 & 4 & -2 \end{pmatrix}$

and the augmented matrix is $\bar{A} = \begin{pmatrix} 1 & 2 & 1 & -3 & 1 \\ 2 & 4 & 3 & 1 & 3 \\ 3 & 6 & 4 & -2 & 4 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$ to reduce it to a row reduced echelon matrix.

$$\bar{A} \xrightarrow[R_3-3R_1]{R_2-2R_1} \begin{pmatrix} 1 & 2 & 1 & -3 & 1 \\ 0 & 0 & 1 & 7 & 1 \\ 0 & 0 & 1 & 7 & 1 \end{pmatrix} \xrightarrow[R_3-R_2]{R_1-R_2} \begin{pmatrix} 1 & 2 & 0 & -10 & 0 \\ 0 & 0 & 1 & 7 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Here rank of A = rank of $\bar{A} = 2$. So the system is consistent.

The given system is equivalent to $$\begin{aligned} x + 2y - 10w &= 0 \\ z + 7w &= 1. \end{aligned}$$

Choosing $y = c, w = d$, where c, d are arbitrary real numbers, the solution is $(-2c + 10d, c, 1 - 7d, d)$
$= (0, 0, 1, 0) + c(-2, 1, 0, 0) + d(10, 0, -7, 1)$.

7. Solve the system of equations $\begin{aligned} x+2y+z &= 1 \\ 3x+y+2z &= 3 \\ x+7y+2z &= 1 \end{aligned}$ in integers.

This is a non-homogeneous system.

The augmented matrix of the system is $\bar{A} = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 3 & 1 & 2 & 3 \\ 1 & 7 & 2 & 1 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$ to reduce it to a row reduced echelon matrix.

$$\bar{A} \xrightarrow[R_3-R_1]{R_2-3R_1} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & -5 & -1 & 0 \\ 0 & 5 & 1 & 0 \end{pmatrix} \xrightarrow{-\frac{1}{5}R_2} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 5 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[R_3-5R_2]{R_1-2R_2} \begin{pmatrix} 1 & 0 & \frac{3}{5} & 1 \\ 0 & 1 & \frac{1}{5} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The given system is equivalent to

$$x + \tfrac{3}{5}z = 1$$
$$y + \tfrac{1}{5}z = 0.$$

Choosing $z = c$, the solution is $(1 - \frac{3}{5}c, -\frac{1}{5}c, c)$, where $c \in \mathbb{R}$.

Since the solutions are to be in integers, $c = 5k$ where k is an arbitrary integer. Hence the solution is $(1-3k, -k, 5k) = (1,0,0) + k(-3,-1,5)$, k being an integer.

8. Determine the conditions for which the system of equations

$$\begin{aligned} x+y+z &= 1 \\ x+2y-z &= b \\ 5x+7y+az &= b^2 \end{aligned}$$

admits of (i) only one solution, (ii) no solution , (iii) many solutions.

The system has a unique solution if the coefficient determinant be non-zero.

The coefficient determinant$= \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 5 & 7 & a \end{vmatrix} = a - 1.$

If $a - 1 \neq 0$, i.e., if $a \neq 1$, the system has only one solution.

If $a = 1$, the system has either no solution or many solutions.

When $a = 1$, the coefficient matrix of the system is

$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 5 & 7 & 1 \end{pmatrix}$ and the augmented matrix of the system is

$\bar{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & b \\ 5 & 7 & 1 & b^2 \end{pmatrix}$.

Let us apply elementary row operations on $\bar{A}$.

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & b \\ 5 & 7 & 1 & b^2 \end{pmatrix} \xrightarrow[R_3-5R_1]{R_2-R_1} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & b-1 \\ 0 & 2 & -4 & b^2-5 \end{pmatrix}$$

$$\xrightarrow[R_3-2R_2]{R_1-R_2} \begin{pmatrix} 1 & 0 & 3 & -b+2 \\ 0 & 1 & -2 & b-1 \\ 0 & 0 & 0 & b^2-2b-3 \end{pmatrix}.$$

If $b^2 - 2b - 3 = 0$, then rank of $\bar{A} =$ rank of $A = 2$ and therefore the system is consistent.

If $b^2 - 2b - 3 \neq 0$, then rank of $\bar{A} = 3$, rank of $A = 2$ and since rank of $A \neq$ rank of $\bar{A}$, the system is inconsistent.

Therefore if $a = 1, b \neq -1, 3$, the system has no solution; and if $a = 1, b = -1$ or if $a = 1, b = 3$, the system has many solutions.

Exercises 10

1. Solve the equations

(i) $\begin{aligned} x+y+3z &= 0 \\ 2x+y+z &= 0 \\ 3x+2y+4z &= 0, \end{aligned}$ (ii) $\begin{aligned} x+y-z-w &= 0 \\ x-y+z-w &= 0. \end{aligned}$

2. Find the solution of the system of equations in rational numbers.

(i) $2x+3y+z = 0,$ (ii) $\begin{aligned} x+4y+z &= 0 \\ 4x+y-z &= 0. \end{aligned}$

3. Find the solution of the system of equations in integers.

(i) $\begin{aligned} x+2y+z &= 0 \\ 3x+y+2z &= 0, \end{aligned}$ (ii) $\begin{aligned} x-3y+4z &= 0 \\ 3x+y-2z &= 0. \end{aligned}$

4. Find a linear homogeneous equation in x_1, x_2, x_3, x_4 such that $x_1 = 1, x_2 = 1, x_3 = 1, x_4 = 1$; $x_1 = 1, x_2 = -1, x_3 = -1, x_4 = 1$ and $x_1 = 2, x_2 = 3, x_3 = 3, x_4 = 2$ are solutions of the equation.

5. Find a linear homogeneous system of two independent equations in x_1, x_2, x_3, x_4 such that $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$ and $x_1 = 2, x_2 = 3, x_3 = 4, x_4 = 1$ are solutions of the system.

6. Examine the solvability of the system of equations and solve, if possible.

(i) $$\begin{aligned} x+y+z &= 1\\ 2x+y+2z &= 1\\ x+2y+3z &= 0, \end{aligned}$$ (ii) $$\begin{aligned} x+y+z &= 1\\ 2x+y+2z &= 2\\ 3x+2y+3z &= 5. \end{aligned}$$

7. For what values of a the system of equations is consistent? Solve completely in each consistent case.

(i) $$\begin{aligned} x-y+z &= 1\\ x+2y+4z &= a\\ x+4y+6z &= a^2, \end{aligned}$$ (ii) $$\begin{aligned} x+y+z &= 1\\ 2x+3y-z &= a+1\\ 2x+y+5z &= a^2+1. \end{aligned}$$

8. For what values of k the system of equations has a non-trivial solution? Solve in each case.

(i) $$\begin{aligned} x+y+z &= kx\\ x+y+z &= ky\\ x+y+z &= kz, \end{aligned}$$ (ii) $$\begin{aligned} x+2y+3z &= kx\\ 2x+y+3z &= ky\\ 2x+3y+z &= kz. \end{aligned}$$

9. Determine the conditions for which the system of equations has (a) only one solution, (b) no solution, (c) many solutions.

(i) $$\begin{aligned} x+2y+z &= 1\\ 2x+y+3z &= b\\ x+ay+3z &= b+1; \end{aligned}$$ (ii) $$\begin{aligned} x+y+z &= b\\ 2x+y+3z &= b+1\\ 5x+2y+az &= b^2. \end{aligned}$$

10. Solve the system of equations $$\begin{aligned} x_2+x_3 &= a\\ x_1+x_3 &= b\\ x_1+x_2 &= c \end{aligned}$$ and use the solution to find the inverse of the matrix $A=\begin{pmatrix} 0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0 \end{pmatrix}$.

11. Solve the system of equations $$\begin{aligned} -x_1+x_2+x_3 &= a\\ x_1-x_2+x_3 &= b\\ x_1+x_2-x_3 &= c \end{aligned}$$ and use the solution to find the inverse of the matrix $A=\begin{pmatrix} -1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1 \end{pmatrix}$.

12. Solve the systems $AX=E_1, AX=E_2, AX=E_3$,

where $E_1=\begin{pmatrix} 1\\ 0\\ 0 \end{pmatrix}$, $E_2=\begin{pmatrix} 0\\ 1\\ 0 \end{pmatrix}$, $E_3=\begin{pmatrix} 0\\ 0\\ 1 \end{pmatrix}$. Hence find A^{-1}.

(i) $A=\begin{pmatrix} 1 & 1 & 0\\ 0 & 1 & -1\\ 2 & 1 & 2 \end{pmatrix}$, (ii) $A=\begin{pmatrix} 1 & 1 & 1\\ 8 & 3 & 4\\ 2 & 1 & 1 \end{pmatrix}$.

4.11. Application to Geometry.

4.11.1. Intersection of two lines in Euclidean plane.

Let the lines be
$$a_{11}x_1 + a_{12}x_2 = b_1,$$
$$a_{21}x_1 + a_{22}x_2 = b_2.$$

Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \end{pmatrix}$.

Let $\alpha_i = (a_{i1}, a_{i2})$, $i = 1, 2$, $\beta_i = (a_{i1}, a_{i2}, b_i)$, $i = 1, 2$.

In order to investigate the nature of solutions of the given non homogeneous system of equations, we are to consider the following cases.

Case 1. Rank of $A = 2$.

There is a unique solution of the system since det $A \neq 0$. Therefore, the lines intersect in a point.

Case 2. Rank of $A = 1$, rank of $B = 2$.

The system is inconsistent and therefore there is no solution of the system. The lines are parallel.

Case 3. Rank of $A = 1$, rank of $B = 1$.

The system is consistent and there are infinite number of solutions since rank of $A < 2$.

Since rank of $B = 1$, β_1, β_2 are linearly dependent. Therefore $\beta_2 = c\beta_1$ for some non-zero real number c.

This shows that the two equations are identical and therefore the lines are coincident.

Examples.

1. The lines $2x + 3y = 3$ and $x + 2y = 1$ intersect in a point, since the rank of the matrix $\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$ is 2.

2. The lines $2x + 3y = 3$ and $4x + 6y = 7$ are parallel, since the rank of the matrix $\begin{pmatrix} 2 & 3 \\ 4 & 6 \end{pmatrix}$ is 1 and the rank of $\begin{pmatrix} 2 & 3 & 3 \\ 4 & 6 & 7 \end{pmatrix}$ is 2.

3. The lines $2x + 3y = 3$ and $6x + 9y = 9$ are identical, since the rank of the matrix $\begin{pmatrix} 2 & 3 \\ 6 & 9 \end{pmatrix}$ = the rank of $\begin{pmatrix} 2 & 3 & 3 \\ 6 & 9 & 9 \end{pmatrix}$ =1.

4.11.2. Intersection of two planes in Euclidean 3-space.

Let the planes be
$$a_{11}x_1 + a_{12}x_2 a_{13}x_3 = b_1,$$
$$a_{21}x_1 + a_{22}x_2 a_{23}x_3 = b_2.$$

Let $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$, $B = \begin{pmatrix} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \end{pmatrix}$,

$\alpha_i = (a_{i1}, a_{i2}, a_{i3})$, $i = 1, 2$; $\beta_i = (a_{i1}, a_{i2}, a_{i3}, b_i)$, $i = 1, 2$.

The following cases come up.

Case 1. Rank of $A = 2$, rank of $B = 2$.

In this case, the solution space of the homogeneous system $AX = O$ is of dimension 1. Let (l_1, l_2, l_3) be the independent generator of the solution space. Then $a_{i1}l_1 + a_{i2}l_2 + a_{i3}l_3 = 0$, $i = 1, 2$.

This shows that the planes are parallel to the direction (l_1, l_2, l_3). Since rank of A = rank of B, the system admits of a solution. Let (p_1, p_2, p_3) be a solution.

Then the general solution of the system is $r(l_1, l_2, l_3) + (p_1, p_2, p_3)$ $= (l_1r + p_1, l_2r + p_2, l_3r + p_3)$, where r is an arbitrary real number.

Therefore the planes intersect along the line

$$\frac{x_1 - p_1}{l_1} = \frac{x_2 - p_2}{l_2} = \frac{x_3 - p_3}{l_3} = r.$$

Case 2. Rank of $A = 1$.

In this case, α_1, α_2 are linearly dependent. Then $\alpha_2 = c\alpha_1$, where c is a non-zero real number. This shows that the planes are perpendicular to the direction (a_{11}, a_{12}, a_{13}).

Sub case (i). Rank of $A = 1$, rank of $B = 2$.

In this case, the system is inconsistent and therefore admits of no solution. The planes are parallel.

Sub case (ii). Rank of $A = 1$, rank of $B = 1$.

In this case, $\beta_2 = c\beta_1$, where $c \neq 0$. This shows that the equations are identical. Therefore the planes are coincident.

4.11.3. Intersection of three planes in Euclidean 3-space.

Let the planes be
$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \quad \ldots \text{(i)}$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \quad \ldots \text{(ii)}$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3 \quad \ldots \text{(iii)}$$

Let $A = (a_{ij})$, $B = (A, b)$, $\alpha_i = (a_{i1}, a_{i2}, a_{i3})$,
$\beta_i = (a_{i1}, a_{i2}, a_{i3}, b_i)$, $i = 1, 2, 3$.

The following cases come up.

Case 1. Rank of $A = 1$.

In this case, only one of $\alpha_1, \alpha_2, \alpha_3$, say α_1 is independent. Therefore $\alpha_2 = c\alpha_1, \alpha_3 = d\alpha_1$, c, d being non-zero real numbers.

The planes are perpendicular to a common direction, the direction vector being (a_{11}, a_{12}, a_{13}).

Sub case (i). Rank of $A = 1$, rank of $B = 1$.

In this case, $\beta_2 = c\beta_1$, $\beta_3 = d\beta_1$.

This shows that the equations are identical and therefore the planes are coincident.

Sub case (ii). Rank of $A = 1$, rank of $B = 2$.

The system of equations is inconsistent and therefore the planes have no common point.

Since rank of $B = 2$, only two of $\beta_1, \beta_2, \beta_3$, say β_1, β_2 are independent. Therefore the planes (i) and (ii) are parallel, by case 2 (i) of 4.11.2.

Now $\beta_3 = c_1\beta_1 + c_2\beta_2$ where $c_1, c_2 \in \mathbb{R}$ and $(c_1, c_2) \neq (0, 0)$.

If $c_1 = 0$, $\beta_3 = c_2\beta_2$. Therefore, the planes (ii) and (iii) are identical.

If $c_2 = 0$, $\beta_3 = c_1\beta_1$. Therefore, the planes (i) and (iii) are identical.

If $c_1 \neq 0$, $c_2 \neq 0$, the three planes are parallel.

Therefore, in this subcase, the planes are parallel, two of which may be coincident.

Case 2. Rank of $A = 2$.

In this case, the solution space of the homogeneous system $AX = O$ is of dimension 1. Let (l_1, l_2, l_3) be the independent solution of the homogeneous system.

Then $a_{i1}l_1 + a_{i2}l_2 + a_{i3}l_3 = 0$, $i = 1, 2, 3$.

Therefore the planes are all parallel to the common direction (l_1, l_2, l_3).

Sub case (i). Rank of $A = 2$, rank of $B = 2$.

The system is consistent. Let (p_1, p_2, p_3) be a solution of the system. The general solution is $r(l_1, l_2, l_3) + (p_1, p_2, p_3), r \in \mathbb{R}$.

That is, $(x_1, x_2, x_3) = (l_1r + p_1, l_2r + p_2, l_3r + p_3)$.

Therefore the planes intersect along the line

$$\frac{x_1 - p_1}{l_1} = \frac{x_2 - p_2}{l_2} = \frac{x_3 - p_3}{l_3} = r.$$

Sub case (ii). Rank of $A = 2$, rank of $B = 3$.

The system is inconsistent and therefore admits of no solution. The vectors $\beta_1, \beta_2, \beta_3$ are linearly independent. Two of $\alpha_1, \alpha_2, \alpha_3$, say α_1, α_2 are linearly independent.

Therefore the planes (i) and (ii) intersect along a line parallel to the direction (l_1, l_2, l_3), by case 1 of 4.11.2.

Let $\alpha_3 = c\alpha_1 + d\alpha_2$ where $(c, d) \neq (0, 0)$.

If $c = 0, \alpha_3 = d\alpha_2, d \neq 0$; but $\beta_3 \neq d\beta_2$. Therefore the planes (ii) and (iii) are parallel.

If $d = 0, \alpha_3 = c\alpha_2, c \neq 0$; but $\beta_3 \neq c\beta_2$. Therefore the planes (i) and (iii) are parallel.

If $c \neq 0, d \neq 0$ then α_1, α_3 and also α_2, α_3 are linearly independent pairs. Therefore the planes (i) and (iii) intersect along a line parallel to the direction (l_1, l_2, l_3) and also the planes (ii) and (iii) intersect along a line parallel to the direction (l_1, l_2, l_3), by case 1 of 4.11.2.

Therefore either

(i) two of the planes are parallel and the third intersects them, or

(ii) three planes intersect in pairs along three parallel lines and the planes form a prism, the axis of the prism being parallel to (l_1, l_2, l_3).

Case 3. Rank of $A = 3$.

In this case, the system of equations admits of a unique solution, since $\det A \neq 0$. The planes intersect in one point only.

Examples (continued).

4. Let the planes be
$$2x_1 + 3x_2 - x_3 = 0$$
$$3x_1 + 3x_2 + x_3 = 2$$
$$x_1 - x_2 + 2x_3 = 5.$$

Here $\det A \neq 0$ and therefore the planes intersect in one point only. The point is $(4, -3, -1)$.

5. Let the planes be
$$x_1 + x_2 - 2x_3 = 3$$
$$x_1 - 2x_2 + x_3 = 3$$
$$x_1 - x_3 = 1.$$

Here rank of $A = 2$, rank of $B = 3$. The solution of the homogeneous system $AX = O$ is $c(1, 1, 1)$, where c is an arbitrary real number. The planes intersect in pairs along lines which are parallel to the direction (1, 1, 1). Therefore the planes form a prism, the axis of the prism is parallel to the direction (1, 1, 1).

6. Let the planes be
$$x_1 + x_2 - 2x_3 = 3$$
$$x_1 - 2x_2 + x_3 = 3$$
$$2x_1 + 2x_2 - 4x_3 = 1.$$

Here rank of $A = 2$, rank of $B = 3$. The solution of the homogeneous system $AX = O$ is $c(1, 1, 1)$. The planes are all parallel to the direction (1, 1, 1). The first and the second planes are parallel. The third plane intersects the other two along lines parallel to the direction (1, 1, 1).

7. Let the planes be
$$5x_1 + 3x_2 + 7x_3 = 4$$
$$2x_1 + x_2 + 3x_3 = 1$$
$$7x_1 + 3x_2 + 11x_3 = 2.$$

Here rank of $A = 2$, rank of $B = 2$.

The system of equations is consistent. It is equivalent to
$$x_1 + 2x_3 = -1$$
$$x_2 - x_3 = 3.$$

Taking $x_3 = r$, the general solution is $(-1 - 2r, 3 + r, r)$.

Therefore $x_1 = -1 - 2r, x_2 = 3 + r, x_3 = r$.

Hence the planes intersect along the line
$$\frac{x_1+1}{-2} = \frac{x_2-3}{1} = \frac{x_3-0}{1} = r.$$

Exercises 11

1. Examine the nature of intersection of the triad of planes

(i) $2x - y + z = 5, x + 2y + 4z = 7, 5x + 3y - z = 0$;

(ii) $x - 2y = 0, 3x + y + z = 8, 2x + 3y + z = 1$;

(iii) $x + y - z = 3, 5x + 2y + z = 1, 2x + 2y - 2z = 1$;

(iv) $2x + y + 2z = 6, y + z = 4, 4x + y + 3z = 8$;

(v) $2x - y = 1, 3y - 2z = 3, 3x - z = 3$.

2. Show that planes $bx - ay = n, cy - bz = l, az - cx = m$ intersect in a line if $al + bm + cn = 0$ and the direction of the line is (a, b, c).

3. For what value of k the planes

(i) $x - 4y + 5z = k, x - y + 2z = 3$ and $2x + y + z = 0$,

(ii) $x + y + z = 2, 3x + y - 2z = k$ and $2x + 4y + 7z = k + 2$

intersect in a line? Find the equations of the line in that case.

4. For what values of k the planes

(i) $x + y + z = 2, 3x + y - 2z = k$ and $2x + 4y + 7z = k + 1$;

(ii) $x + y + 1 = 0, 4x + y - z = k$ and $5x - y - 2z = k^2$

form a triangular prism ?

4.12. Euclidean Spaces.

Real inner product. Let V be a real vector space. A *real inner product* on V is a mapping $f : V \times V \to \mathbb{R}$ that assigns to each ordered pair of vectors (α, β) of V a real number $f(\alpha, \beta)$, generally denoted by $\alpha.\beta$ or by (α, β), satisfying the following properties–

(1) $(\alpha, \beta) = (\beta, \alpha)$ for all $\alpha, \beta \in V$, (symmetry)

(2) $(\alpha, \beta + \gamma) = (\alpha, \beta) + (\alpha, \gamma)$ for all $\alpha, \beta, \gamma \in V$, (linearity)

(3) $(c\alpha, \beta) = c(\alpha, \beta) = (\alpha, c\beta)$ for all $\alpha, \beta \in V$ and all $c \in \mathbb{R}$, (homogeneity)

(4) $(\alpha, \alpha) > 0$ if $\alpha \neq \theta$. (positivity)

If $\alpha = \theta$, then $(\alpha, \alpha) = 0$. This follows from the property (3) by taking $\alpha = \beta = \theta$. That is, $(\theta, \theta) = 0$.

The mapping f is also denoted by $(,)$ and the f-image of the ordered pair (α, β) is denoted by (α, β).

Definition. A real vector space V together with a real inner product defined on it, is said to be a **Euclidean space.**

Complex inner product. Let V be a complex vector space. A complex inner product is a mapping $f : V \times V \to \mathbb{C}$ that assigns to each ordered pair of vectors (α, β) of V a complex number $f(\alpha, \beta)$, generally denoted by (α, β), satisfying the following properties–

(1) $(\alpha, \beta) = \overline{(\beta, \alpha)}$, where $\overline{(\beta, \alpha)}$ is the conjugate of the complex number (β, α),

(2) $(c\alpha + d\beta, \gamma) = c(\alpha, \gamma) + d(\beta, \gamma)$,

(3) $(\alpha, \alpha) > 0$ if $\alpha \neq \theta$ and $(\theta, \theta) = 0$.

From (1) it follows that $(\alpha, \alpha) = \overline{(\alpha, \alpha)}$ showing that (α, α) is a real number. The property (3) says that the complex inner product satisfies the positivity condition as in the case of a real inner product.

Deduction.

$$\begin{aligned}(\alpha, c\beta + d\gamma) = \overline{(c\beta + d\gamma, \alpha)} &= \overline{c(\beta, \alpha) + d(\gamma, \alpha)} \\ &= \bar{c}\overline{(\beta, \alpha)} + \bar{d}\overline{(\gamma, \alpha)} = \bar{c}(\alpha, \beta) + \bar{d}(\alpha, \gamma),\end{aligned}$$

for all $\alpha, \beta, \gamma \in V$; $c, d \in \mathbb{C}$.

Definition. A complex vector space V together with a complex inner product defined on it, is said to be a **Unitary Space.**

A Euclidean space (or a unitary space) is also called an *inner product space.*

We shall be mainly concerned with Euclidean spaces in this treatise.

Examples.

1. In the real vector space $\mathbb{R}^n$, let $\alpha = (a_1, a_2, \ldots, a_n)$, $\beta = (b_1, b_2, \ldots, b_n)$ be two vectors and let us define

$$(\alpha, \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n.$$

Then $(,)$ satisfies all the conditions for a real inner product.

This inner product is called the *standard inner product* and is often called the *dot product*, denoted by $\alpha.\beta$.

The vector space $\mathbb{R}^n$ equipped with this inner product becomes a Euclidean space.

2. In $\mathbb{R}^2$, the standard inner product is defined by
$(\alpha, \beta) = a_1b_1 + a_2b_2$ where $\alpha = (a_1, a_2) \in \mathbb{R}^2, \beta = (b_1, b_2) \in \mathbb{R}^2$.

In $\mathbb{R}^2$, let us define $(\alpha, \beta) = 2a_1b_1 + a_1b_2 + a_2b_1 + a_2b_2$. Then $(,)$ satisfies the conditions 1, 2, 3 of a real inner product.

$$\begin{aligned}(\alpha, \alpha) &= 2a_1^2 + 2a_1a_2 + a_2^2 \\ &= a_1^2 + (a_1 + a_2)^2 > 0 \text{ except when } a_1 = a_2 = 0.\end{aligned}$$

Therefore $(\alpha, \alpha) > 0$ for $\alpha \neq \theta$, showing that the positivity condition is also satisfied by $(,)$.

Thus $(\alpha, \beta) = 2a_1b_1 + a_1b_2 + a_2b_1 + a_2b_2$ defines a real inner product in $\mathbb{R}^2$. Therefore the vector space $\mathbb{R}^2$ becomes a Euclidean space under this inner product.

This example shows that the a real vector space can be made a Euclidean space in many ways.

3. In $\mathbb{R}^n$, let $B = (\beta_1, \beta_2, \cdots, \beta_n)$ be an ordered basis and let the co-ordinates of two vectors α, β of $\mathbb{R}^n$ relative to the ordered basis B be $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ respectively.

If we define $(\alpha, \beta) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$ then $(,)$ satisfies all the conditions of a real inner product.

Therefore the vector space $\mathbb{R}^n$ becomes a Euclidean space under this inner product.

As there may be many ordered bases of $\mathbb{R}^n$, $\mathbb{R}^n$ can be made a Euclidean space in many ways.

4. In the real vector space P_n, let us define for polynomials $f, g \in P_n$,

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

Then $(,)$ satisfies all the conditions for a real inner product.

Therefore the vector space P_n equipped with this real inner product becomes a Euclidean space.

5. In P_n, let us define $(f,g) = \int_{-1}^{1} f(t)g(t)dt$ for $f, g \in P_n$.
Then (,) also satisfies all the conditions for a real inner product.

Therefore the vector space P_n equipped with this real inner product becomes a Euclidean space different from the space described in Ex.4.

Definition. Norm of a vector.

If α be a vector in a Euclidean space V with the inner product (,), the *norm* of α, denoted by $\|\alpha\|$, is defined by $\|\alpha\| = \sqrt{(\alpha,\alpha)}$.

Theorem 4.12.1. Let α be a vector in a Euclidean space V and $\|\alpha\|$ be its norm.

Then (i) $\|c\alpha\| = |c|\|\alpha\|$, c being a real number;

(ii) $\|\alpha\| > 0$ unless $\alpha = \theta$ and $\|\theta\| = 0$.

Proof. (i) $\|c\alpha\| = \sqrt{(c\alpha, c\alpha)} = \sqrt{c^2(\alpha,\alpha)} = |c|\sqrt{(\alpha,\alpha)} = |c|\|\alpha\|$.

(ii) $\alpha \neq \theta$ implies $(\alpha,\alpha) > 0$ and therefore $\|\alpha\| > 0$.

If $\alpha = \theta$, then $(\alpha,\alpha) = (\theta,\theta) = 0$ and therefore $\|\alpha\| = 0$.

4.12.2. Schwarz's inequality.

For any two vectors α, β in a Euclidean space V,
$$|(\alpha,\beta)| \leq \|\alpha\|\|\beta\|,$$
the equality holds when α, β are linearly dependent.

Proof. **Case 1.** Let one or both of α, β be null. Then both sides being zero, the equality holds.

Case 2. Let α, β be non-null and linearly dependent. Then there exists a non-zero real number k such that $\alpha = k\beta$.

Then $\|\alpha\| = |k|\|\beta\|$ and $(\alpha,\beta) = (k\beta,\beta) = k(\beta,\beta) = k\|\beta\|^2$.
Therefore $|(\alpha,\beta)| = |k|\|\beta\|^2 = \|\alpha\|\|\beta\|$.

Case 3. Let α, β be not linearly dependent. Then $\alpha - k\beta \neq \theta$ for all real k.

Therefore $(\alpha - k\beta, \alpha - k\beta) > 0$ for all real k
or, $(\alpha,\alpha) - 2k(\alpha,\beta) + k^2(\beta,\beta) > 0$ for all real k.

Since $(\alpha,\alpha), (\alpha,\beta), (\beta,\beta)$ are all real and $(\beta,\beta) \neq 0$, the left-hand side is a real quadratic polynomial in k and since it is positive for all real values of k, the discriminant of the quadratic polynomial must be negative, for otherwise the polynomial would be zero for some real k.

Thus $(\alpha,\beta)^2 - (\alpha,\alpha)(\beta,\beta) < 0$, whence $|(\alpha,\beta)| < \|\alpha\|\|\beta\|$.
This completes the proof.

Note. In particular, if the Euclidean space be $\mathbb{R}^n$ with standard inner product, taking $\alpha = (a_1, a_2, \ldots, a_n), \beta = (b_1, b_2, \ldots, b_n)$ the inequality takes the form
$$(a_1b_1 + a_2b_2 + \cdots + a_nb_n)^2 \leq (a_1^2 + a_2^2 + \cdots + a_n^2)(b_1^2 + b_2^2 + \cdots + b_n^2),$$
the equality holds when
either (i) $a_i = 0$, or $b_i = 0$, or both $a_i = 0$ and $b_i = 0$ for $i = 1, 2, \ldots, n$;
or (ii) $a_i = kb_i$ for some non-zero real $k, i = 1, 2, \ldots, n$.

Theorem 4.12.3. In a Euclidean space V, two vectors α, β are linearly dependent if and only if $|(\alpha,\beta)| = \|\alpha\|\|\beta\|$.

Proof. Let α, β be linearly dependent.
If one or both of α, β be null, then the equality holds.

If α, β be both non-null, then $\alpha = k\beta$ for some non-zero real k.
In this case, $\|\alpha\| = |k|\|\beta\|$ and $(\alpha,\beta) = (k\beta,\beta) = k(\beta,\beta) = k\|\beta\|^2$.
Therefore $|(\alpha,\beta)| = |k|\|\beta\|^2 = \|\alpha\|\|\beta\|$.

Conversely, let $|(\alpha,\beta)| = \|\alpha\|\|\beta\|$.
α, β are linearly independent implies $|(\alpha,\beta)| < \|\alpha\|\|\beta\|$, by case 3 of the previous theorem.
Contrapositively, $|(\alpha,\beta)| \not< \|\alpha\|\|\beta\|$ implies α, β are linearly dependent. But by Schwarz's inequality, $|(\alpha,\beta)| \leq \|\alpha\|\|\beta\|$ for all α, β in V.

Therefore $|(\alpha,\beta)| = \|\alpha\|\|\beta\|$ implies α, β are linearly dependent.
This completes the proof.

Note. Linear dependence of α, β implies $|(\alpha,\beta)| = \|\alpha\|\|\beta\|$.
But (i) linear dependence of α, β may not imply $(\alpha,\beta) = \|\alpha\|\|\beta\|$. For example, let $\alpha = (1,2,3), \beta = (-2,-4,-6)$;
(ii) linear dependence of α, β may not imply $(\alpha,\beta) = -\|\alpha\|\|\beta\|$. For example, let $\alpha = (1,2,3), \beta = (2,4,6)$.

4.12.4. Triangle inequality.

If α, β be any two vectors in a Euclidean space V, then
$$\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|.$$

Proof. From the properties of an inner product,
$$\begin{aligned} \|\alpha+\beta\|^2 &= (\alpha+\beta, \alpha+\beta) = (\alpha,\alpha) + 2(\alpha,\beta) + (\beta,\beta) \\ &= \|\alpha\|^2 + 2(\alpha,\beta) + \|\beta\|^2 \\ &\leq \|\alpha\|^2 + 2\|\alpha\|\,\|\beta\| + \|\beta\|^2, \text{ by Schwarz's inequality} \\ &= (\|\alpha\| + \|\beta\|)^2. \end{aligned}$$
Therefore $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.
This completes the proof.

Note. $\|\alpha+\beta\| = \|\alpha\| + \|\beta\|$ implies $(\alpha,\beta) = \|\alpha\|\|\beta\|$ and this again implies α, β are linearly dependent.

Since linear dependence of α, β may not imply $(\alpha,\beta) = \|\alpha\|\|\beta\|$, $\|\alpha + \beta\|$ may not be equal to $\|\alpha\| + \|\beta\|$ if α, β are linearly dependent. For example, let $\alpha = (1,2,3), \beta = (-1,-2,-3)$.

Definitions.

A vector α in a Euclidean space V is said to be a *unit* vector if $\|\alpha\| = 1$. If α be a non-zero vector in V, then $\frac{1}{\|\alpha\|}\alpha$ is a unit vector.

In a Euclidean space, a vector α is said to be *orthogonal* to a vector β if $(\alpha,\beta) = 0$. Since $(\alpha,\beta) = (\beta,\alpha)$, if α be orthogonal to β then β is orthogonal to α. In this case, α, β are said to be orthogonal.

The null vector θ is orthogonal to any non-null vector α and also it is orthogonal to itself. This follows from the property of an inner product.

4.12.5. Pythagoras theorem.

If α, β be two orthogonal vectors in a Euclidean space V, then

$$\|\alpha+\beta\|^2 = \|\alpha\|^2 + \|\beta\|^2.$$

Proof. $\|\alpha+\beta\|^2 = (\alpha+\beta, \alpha+\beta) = (\alpha,\alpha) + 2(\alpha,\beta) + (\beta,\beta)$
$= \|\alpha\|^2 + \| b\|^2$, since $(\alpha,\beta) = 0$.

This completes the proof.

4.12.6. Parallelogram law.

If α, β be any two vectors in a Euclidean space V, then

$$\|\alpha+\beta\|^2 + \|\alpha-\beta\|^2 = 2\|\alpha\|^2 + 2\|\beta\|^2.$$

Proof. $\|\alpha+\beta\|^2 = (\alpha+\beta,\alpha+\beta) = (\alpha,\alpha) + 2(\alpha,\beta) + (\beta,\beta)$.

$\|\alpha-\beta\|^2 = (\alpha-\beta,\alpha-\beta) = (\alpha,\alpha) - 2(\alpha,\beta) + (\beta,\beta)$.

Therefore $\|\alpha+\beta\|^2 + \|\alpha-\beta\|^2 = 2(\alpha,\alpha) + 2(\beta,\beta) = 2\|\alpha\|^2 + 2\|\beta\|^2$.

This completes the proof.

Definitions.

A set of vectors $\{\beta_1, \beta_2, \ldots, \beta_r\}$ in a Euclidean space is said to be *orthogonal* if $(\beta_i, \beta_j) = 0$ whenever $i \neq j$.

A set of vectors $\{\beta_1, \beta_2, \ldots, \beta_r\}$ in a Euclidean space is said to be *orthonormal* if $(\beta_i, \beta_j) = 0$ for $i \neq j$
$= 1$ for $i = j$.

Note. An orthogonal set of vectors may contain the null vector θ but an orthonormal set contains only non-null vectors.

Example 6. In an $n \times n$ real orthogonal matrix, the row vectors form an orthonormal set and the column vectors form another orthonormal set in $\mathbb{R}^n$ with standard inner product.

Let A be a real $n \times n$ orthogonal matrix. Then $AA^t = I_n$. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be the row vectors of A.

Then the column vectors of A^t are $\alpha_1, \alpha_2, \ldots, \alpha_n$.

The ijth element of AA^t

$=$ the inner product of the ith row vector of A and the jth column vector of A^t

$= (\alpha_i, \alpha_j)$.

Since $AA^t = I_n$, $(\alpha_i, \alpha_j) = 0$ if $i \neq j$
$= 1$ if $i = j$.

This proves that $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an orthonormal set.

In a similar manner it can be shown that the column vectors of A form an orthonormal set.

Note. The row vectors of I_n, i.e., the vectors $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ form an orthonormal set in $\mathbb{R}^n$ with standard inner product.

Theorem 4.12.7. An orthogonal set of non-null vectors in a Euclidean space V is linearly independent.

Proof. Let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthogonal set of non-null vectors. Let us consider the relation

$c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r = \theta$, where c_i are real numbers.

Then $(c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r, \beta_i) = (\theta, \beta_i) = 0$ for $i = 1, 2, \ldots, r$.

or, $c_1(\beta_1, \beta_i) + c_2(\beta_2, \beta_i) + \cdots + c_i(\beta_i, \beta_i) + \cdots + c_r(\beta_r, \beta_i) = 0$

or, $c_i(\beta_i, \beta_i) = 0$, since $(\beta_i, \beta_j) = 0, j \neq i$

Since β_i is non-null, $(\beta_i, \beta_i) > 0$ and therefore $c_i = 0$.

This proves that the set $\{\beta_1, \beta_2, \ldots, \beta_r\}$ is linearly independent.

Corollary. An orthonormal set of vectors in a Euclidean space is linearly independent.

Definitions.

Let β be a fixed non-zero vector in a Euclidean space V. Then for a non-zero vector α in V there exists a *unique* real number c such that $\alpha - c\beta$ is orthogonal to β.

c is determined by $(\alpha - c\beta, \beta) = 0$. Therefore $(\alpha, \beta) = c(\beta, \beta)$, giving $c = \frac{(\alpha,\beta)}{(\beta,\beta)}$.

c is said to be the *scalar component* (or *component*) of α along β and $c\beta$ is said to be the *projection* of α upon β.

Theorem 4.12.8. If $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthogonal set of non-null vectors in a Euclidean space V then any vector β in $L\{\beta_1, \beta_2, \ldots, \beta_r\}$ has the unique representation $\beta = c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$, where c_i is the scalar component of β along β_i.

Proof. Since $\{\beta_1, \beta_2, \ldots, \beta_r\}$ is an orthogonal set of non-null vectors, it is linearly independent and therefore it is a basis of the subspace $L\{\beta_1, \beta_2, \ldots, \beta_r\}$.

Therefore β can be expressed as $\beta = c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$, where c_i are unique real numbers.

Now $(\beta, \beta_i) = c_i(\beta_i, \beta_i)$, since $(\beta_j, \beta_i) = 0$ for $j \neq i$.

So $c_i = \frac{(\beta, \beta_i)}{(\beta_i, \beta_i)}$, which is the scalar component of β along β_i.

Note. Every vector β in a Euclidean space V is the sum of its projections along the vectors of an orthogonal basis of V.

Corollary. If $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthonormal set of vectors in a Euclidean space V, any vector β in $L\{\beta_1, \beta_2, \ldots, \beta_r\}$ can be expressed as $\beta = (\beta, \beta_1)\beta_1 + (\beta, \beta_2)\beta_2 + \cdots + (\beta, \beta_r)\beta_r$.

Worked Example.

1. Prove that the set of vectors $\{(1,2,2), (2,-2,1), (2,1,-2)\}$ is an orthogonal basis of the Euclidean space $\mathbb{R}^3$ with standard inner product. Express $(4,3,2)$ as a linear combination of these basis vectors.

Let $\beta_1 = (1,2,2)$, $\beta_2 = (2,-2,1)$, $\beta_3 = (2,1,-2)$.

Then $(\beta_1, \beta_2) = 0$, $(\beta_2, \beta_3) = 0$, $(\beta_3, \beta_1) = 0$.

So $\{\beta_1, \beta_2, \beta_3\}$ is an orthogonal set of non-zero vectors and therefore it is linearly independent.

Since $\mathbb{R}^3$ is a vector space of dimension 3, $\{\beta_1, \beta_2, \beta_3\}$ is a basis of $\mathbb{R}^3$.

Let $\beta = (4,3,2)$. So $\beta = c_1\beta_1 + c_2\beta_2 + c_3\beta_3$, where c_i is the component of β along β_i.

$$c_1 = \frac{(\beta, \beta_1)}{(\beta_1, \beta_1)} = \frac{4.1+3.2+2.2}{9} = \frac{14}{9},$$
$$c_2 = \frac{(\beta, \beta_2)}{(\beta_2, \beta_2)} = \frac{4.2+3.-2+2.1}{9} = \frac{4}{9},$$
$$c_3 = \frac{(\beta, \beta_3)}{(\beta_3, \beta_3)} = \frac{4.2+3.1+2.-2}{9} = \frac{7}{9}.$$

Therefore $\beta = \frac{14}{9}\beta_1 + \frac{4}{9}\beta_2 + \frac{7}{9}\beta_3$.

Theorem 4.12.9. Bessel's inequality.

If $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthonormal set of vectors in a Euclidean space V, then for any vector α in V,

$$\|\alpha\|^2 \geq c_1^2 + c_2^2 + \cdots + c_r^2,$$

where c_i is the scalar component of α along β_i, $i = 1, 2, \ldots, r$.

Proof. For all i $(i = 1, 2, \ldots, r)$, $c_i = (\alpha, \beta_i)$, since $(\beta_i, \beta_i) = 1$.

$\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r$ is orthogonal to each β_i, $1 \leq i \leq r$, since $(\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r, \beta_i) = (\alpha, \beta_i) - c_i(\beta_i, \beta_i) = 0$.

It follows that $\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r$ is orthogonal to $c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$.

$\alpha = (\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r) + (c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r)$.

By Pythagoras Theorem,

$$\begin{aligned} \|\alpha\|^2 &= \|\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r\|^2 + \|c_1\beta_1 + \cdots + c_r\beta_r\|^2 \\ &\geq \|c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r\|^2, \text{ since a norm is non-negative.} \end{aligned}$$

But $\|c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r\|^2$

$= (c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r, c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r)$

$= c_1^2 + c_2^2 + \cdots + c_r^2$, since $\{\beta_1, \beta_2, \ldots, \beta_r\}$ is an orthonormal set.

Consequently, $\|\alpha\|^2 \geq c_1^2 + c_2^2 + \cdots + c_r^2$.

This completes the proof.

Note. The equality occurs if $\|\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r\|^2 = 0$, i.e., if $\alpha = c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$, i.e., if $\alpha \in L\{\beta_1, \beta_2, \ldots, \beta_r\}$.

Theorem 4.12.10. Parseval's theorem.

If $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be an orthonormal basis of a Euclidean space V, then for any vector α in V,

$$\|\alpha\|^2 = c_1^2 + c_2^2 + \cdots + c_n^2,$$

where c_i is the scalar component of α along β_i, $i = 1, 2, \ldots, n$.

Proof. Since $\{\beta_1, \beta_2, \ldots, \beta_n\}$ is a basis of V, any vector $\alpha \in V$ can be expressed as $\alpha = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$, where c_i is the scalar component of α along β_i, $i = 1, 2, \ldots, r$.

$$\begin{aligned} \text{So } \|\alpha\|^2 = (\alpha, \alpha) &= (c_1\beta_1 + \cdots + c_n\beta_n, c_1\beta_1 + \cdots + c_n\beta_n) \\ &= c_1^2 + c_2^2 + \cdots + c_n^2, \text{ since } \{\beta_1, \beta_2, \ldots, \beta_r\} \\ &\quad \text{is an orthonormal set.} \end{aligned}$$

This completes the proof.

Theorem 4.12.11. Let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthogonal set of non-null vectors in a Euclidean space V and α be a vector in $V - L\{\beta_1, \beta_2, \ldots, \beta_r\}$. If the scalar component of α along β_i be c_i, then

(i) $\beta = \alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r$ is orthogonal to each β_i; and

(ii) $L\{\beta_1, \beta_2, \ldots, \beta_r, \alpha\} = L\{\beta_1, \beta_2, \ldots, \beta_r, \beta\}$.

Proof. (i)
$$\begin{aligned}(\beta, \beta_1) &= (\alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r, \beta_1) \\ &= (\alpha, \beta_1) - c_1(\beta_1, \beta_1) = 0, \text{ since } c_1 = \tfrac{(\alpha,\beta_1)}{(\beta_1,\beta_1)}.\end{aligned}$$

Similarly, $(\beta, \beta_2) = 0, \ldots, (\beta, \beta_r) = 0$.
Therefore β is orthogonal to each β_i.

(ii) Let $S = \{\beta_1, \beta_2, \ldots, \beta_r, \alpha\}$, $T = \{\beta_1, \beta_2, \ldots, \beta_r, \beta\}$.
We have $\beta = \alpha - c_1\beta_1 - c_2\beta_2 - \cdots - c_r\beta_r$ ... (i)
and $\alpha = \beta + c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$ ... (ii)

Using (i) we can say that each vector in T is a linear combination of the vectors in S. Therefore $L(T) \subset L(S)$.

Using (ii) we can say that each vector in S is a linear combination of the vectors in T. Therefore $L(S) \subset L(T)$.
It follows that $L(S) = L(T)$.
This completes the proof.

Note 1. The theorem is also valid when $\alpha \in L\{\beta_1, \beta_2, \ldots, \beta_r\}$. In this case, $\alpha = c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$, where c_i is the scalar component of α along β_i. Clearly, $\beta = \theta$ and therefore

(i) β is orthogonal to each β_i, and

(ii) $L\{\beta_1, \beta_2, \ldots, \beta_r, \beta\} = L\{\beta_1, \beta_2, \ldots, \beta_r, \theta\} = L\{\beta_1, \beta_2, \ldots, \beta_r\} = L\{\beta_1, \beta_2, \ldots, \beta_r, \alpha\}$, since α is a linear combination of $\beta_1, \beta_2, \ldots, \beta_r$.

2. An orthogonal set of r non-null vectors $\{\beta_1, \beta_2, \ldots, \beta_r\}$ can be extended to an orthogonal set of $r+1$ non-null vectors with the help of a vector α lying in $V - L\{\beta_1, \beta_2, \ldots, \beta_r\}$. The theorem gives a clue to the extension.

Theorem 4.12.12. An orthogonal set of non-null vectors in a finite dimensional Euclidean space V, if not a basis of V, can be extended to an orthogonal basis of V.

Proof. Let $\dim V = n$ and let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthogonal set of non-null vectors in V where $1 \le r \le n$. So $\{\beta_1, \beta_2, \ldots, \beta_r\}$ is a linearly independent set.

If $r = n$, then $\{\beta_1, \beta_2, \ldots, \beta_r\}$ is an orthogonal basis of V.

If $r < n$, then $L\{\beta_1, \beta_2, \ldots, \beta_r\}$ is a proper subspace of V and so there exists a vector α_{r+1} in V such that $\alpha_{r+1} \notin L\{\beta_1, \beta_2, \ldots, \beta_r\}$. We assert that the set $\{\beta_1, \beta_2, \ldots, \beta_r, \alpha_{r+1}\}$ is linearly independent.

Let us consider the relation $c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r + c_{r+1}\alpha_{r+1} = \theta$, where $c_i \in \mathbb{R}$. Then c_{r+1} must be zero, for otherwise α_{r+1} would belong to $L\{\beta_1, \beta_2, \ldots, \beta_r\}$. The linear independence of $\{\beta_1, \beta_2, \ldots, \beta_r\}$ and $c_{r+1} = 0$ together imply that $c_1 = c_2 = \cdots = c_{r+1} = 0$ and this proves our assertion.

Let $\beta_{r+1} = \alpha_{r+1} - d_1\beta_1 - d_2\beta_2 - \cdots - d_r\beta_r$ where d_i is the scalar component of α_{r+1} along β_i. Then $\beta_{r+1} \neq \theta$ and is orthogonal to $\beta_1, \beta_2, \ldots, \beta_r$ and thus an orthogonal set of $r+1$ non-null vectors $\{\beta_1, \beta_2, \ldots, \beta_{r+1}\}$ is obtained in V as an extension of the set $\{\beta_1, \beta_2, \ldots, \beta_r\}$.

If $r + 1 = n$, then $\{\beta_1, \beta_2, \ldots, \beta_{r+1}\}$ is an orthogonal basis of V.

If $r + 1 < n$, then by repeated application of the procedure described above we obtain in a finite number of steps an orthogonal set of n vectors $\{\beta_1, \beta_2, \ldots, \beta_{r+1}, \ldots, \beta_n\}$ in V.

This set being an orthogonal set of non-null vectors, is linearly independent. Furthermore, this being a linearly independent set of n vectors in V is a basis of V.

This completes the proof.

Corollary. An orthonormal set of vectors in a finite dimensional Euclidean space V, if not a basis of V, can be extended to an orthonormal basis of V.

Worked Example (continued).

2. Extend the set of vectors $\{(2, 3, -1), (1, -2, -4)\}$ to an orthogonal basis of the Euclidean space $\mathbb{R}^3$ with standard inner product and then find the associated orthonormal basis.

Let $\alpha_1 = (2, 3, -1)$, $\alpha_2 = (1, -2, -4)$.

α_1, α_2 are orthogonal vectors. Let $\alpha_3 = (0, 0, 1)$. Then $\{\alpha_1, \alpha_2, \alpha_3\}$ is linearly independent because $\begin{vmatrix} 2 & 3 & -1 \\ 1 & -2 & -4 \\ 0 & 0 & 1 \end{vmatrix} \neq 0$.

So $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis of $\mathbb{R}^3$.

Let $\beta = \alpha_3 - c_1\alpha_1 - c_2\alpha_2$, where $c_1 = \frac{(\alpha_3, \alpha_1)}{(\alpha_1, \alpha_1)}$, $c_2 = \frac{(\alpha_3, \alpha_2)}{(\alpha_2, \alpha_2)}$.

Then β is orthogonal to α_1 and α_2 and $L\{\alpha_1, \alpha_2, \alpha_3\} = L\{\alpha_1, \alpha_2, \beta\}$.

Therefore $\{\alpha_1, \alpha_2, \beta\}$ is an orthogonal basis of $\mathbb{R}^3$.

$c_1 = -\frac{1}{14}, c_2 = -\frac{4}{21}$ and therefore

$\beta = (0,0,1) + \frac{1}{14}(2,3,-1) + \frac{4}{21}(1,-2,-4) = (\frac{1}{3}, -\frac{1}{6}, \frac{1}{6})$.

Hence an extended orthogonal basis is $\{(2,3,-1), (1,-2,-4), (\frac{1}{3}, -\frac{1}{6}, \frac{1}{6})\}$ and the associated orthonormal basis is

$$\left\{\frac{1}{\sqrt{14}}(2,3,-1), \frac{1}{\sqrt{21}}(1,-2,-4), \frac{1}{\sqrt{6}}(2,-1,1)\right\}.$$

Theorem 4.12.13. Every non-null subspace W of a finite dimensional Euclidean space V possesses an orthonormal basis.

Proof. Let $\{\alpha_1, \alpha_2, ..., \alpha_r\}$ be a basis of W. An orthogonal basis will be obtained by a method known as **Gram-Schmidt process** of orthogonalisation. Since the basis vectors are none zero, we pick up one of them, say α_1, and consider as the first member of the new basis. For convenience we rename it β_1.

$$\beta_1 = \alpha_1.$$

Let $\beta_2 = \alpha_2 - c_1\beta_1$, where $c_1\beta_1$ is the projection of α_2 upon β_1.

Then β_2 is orthogonal to β_1 and $L\{\beta_1, \beta_2\} = L\{\beta_1, \alpha_2\} = L\{\alpha_1, \alpha_2\}$.

$$\beta_2 = \alpha_2 - \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)}\beta_1.$$

$\alpha_3 \notin L\{\beta_1, \beta_2\}$. Let $\beta_3 = \alpha_3 - d_1\beta_1 - d_2\beta_2$, where $d_1\beta_1, d_2\beta_2$ are the projections of α_3 upon β_1, β_2 respectively.

Then β_3 is orthogonal to β_1, β_2 and $L\{\beta_1, \beta_2, \beta_3\} = L\{\beta_1, \beta_2, \alpha_3\} = L\{\alpha_1, \alpha_2, \alpha_3\}$.

$$\beta_3 = \alpha_3 - \frac{(\alpha_3, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)}\beta_2.$$

$\alpha_4 \notin L\{\beta_1, \beta_2, \beta_3\}$. Let $\beta_4 = \alpha_4 - r_1\beta_1 - r_2\beta_2 - r_3\beta_3$, where $r_1\beta_1, r_2\beta_2, r_3\beta_3$ are the projections of α_4 upon $\beta_1, \beta_2, \beta_3$ respectively.

Then β_4 is orthogonal to $\beta_1, \beta_2, \beta_3$ and $L\{\beta_1, \beta_2, \beta_3, \beta_4\} = L\{\beta_1, \beta_2, \beta_3, \alpha_4\} = L\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$.

$$\beta_4 = \alpha_4 - \frac{(\alpha_4, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_4, \beta_2)}{(\beta_2, \beta_2)}\beta_2 - \frac{(\alpha_4, \beta_3)}{(\beta_3, \beta_3)}\beta_3.$$

This process terminates after a finite number of steps because at every step one vector of the given basis is replaced by a vector in the desired orthogonal basis. Finally we obtain

$$\beta_r = \alpha_r - \frac{(\alpha_r, \beta_1)}{(\beta_1, \beta_1)}\beta_1 - \frac{(\alpha_r, \beta_2)}{(\beta_2, \beta_2)}\beta_2 - \cdots - \frac{(\alpha_r, \beta_{r-1})}{(\beta_{r-1}, \beta_{r-1})}\beta_{r-1}, \text{ and}$$

$\{\beta_1, \beta_2, \ldots, \beta_r\}$ is an orthogonal basis of the subspace W.

This completes the proof.

Worked Examples (continued).

3. Use Gram-Schmidt process to obtain an orthogonal basis from the basis set $\{(1,0,1), (1,1,1), (1,3,4)\}$ of the Euclidean space $\mathbb{R}^3$ with standard inner product.

Let $\alpha_1 = (1,0,1), \alpha_2 = (1,1,1), \alpha_3 = (1,3,4)$.

Let $\beta_1 = \alpha_1$ and $\beta_2 = \alpha_2 - c_1\beta_1$, where c_1 is the scalar component of α_2 along β_1. Then β_2 is orthogonal to β_1 and $L\{\beta_1, \beta_2\} = L\{\beta_1, \alpha_2\} = L\{\alpha_1, \alpha_2\}$.

$c_1 = \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} = 1$. Therefore $\beta_2 = \alpha_2 - \beta_1 = (0,1,0)$.

Let $\beta_3 = \alpha_3 - d_1\beta_1 - d_2\beta_2$, where d_1, d_2 are scalar components of α_3 along β_1, β_2 respectively.

Then β_3 is orthogonal to β_1, β_2 and $L\{\beta_1, \beta_2, \beta_3\} = L\{\alpha_1, \alpha_2, \alpha_3\}$.

$d_1 = \frac{(\alpha_3, \beta_1)}{(\beta_1, \beta_1)} = \frac{(\alpha_3, \alpha_1)}{(\alpha_1, \alpha_1)} = \frac{5}{2}$, $d_2 = \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)} = \frac{3}{1} = 3$.

Therefore $\beta_3 = (1,3,4) - \frac{5}{2}(1,0,1) - 3(0,1,0) = \frac{3}{2}(-1,0,1)$.

Therefore an orthogonal basis is $\{(1,0,1), (0,1,0), \frac{3}{2}(-1,0,1)\}$.

4. Use Gram-Schmidt process to obtain an orthonormal basis of the subspace of the Euclidean space $\mathbb{R}^4$ with standard inner product, generated by the linearly independent set $\{(1,1,0,1), (1,1,0,0), (0,1,0,1)\}$.

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

Let $\beta_1 = \alpha_1$ and $\beta_2 = \alpha_2 - c_1\beta_1$, where $c_1\beta_1$ is the projection of α_2 upon β_1.

Then β_2 is orthogonal to β_1 and $L\{\beta_1, \beta_2\} = L\{\beta_1, \alpha_2\} = L\{\alpha_1, \alpha_2\}$.

$c_1 = \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} = \frac{(\alpha_2, \alpha_1)}{(\alpha_1, \alpha_1)} = \frac{2}{3}$.

$\beta_2 = \alpha_2 - \frac{2}{3}\alpha_1 = (1,1,0,0) - \frac{2}{3}(1,1,0,1) = \frac{1}{3}(1,1,0,-2)$.

Let $\beta_3 = \alpha_3 - d_1\beta_1 - d_2\beta_2$, where $d_1\beta_1, d_2\beta_2$ are projections of α_3 upon β_1 and β_2 respectively.

Then β_3 is orthogonal to β_1, β_2 and $L\{\beta_1, \beta_2, \beta_3\} = L\{\alpha_1, \alpha_2, \alpha_3\}$.

$d_1 = \frac{(\alpha_3, \beta_1)}{(\beta_1, \beta_1)} = \frac{2}{3}, d_2 = \frac{(\alpha_3, \beta_2)}{(\beta_2, \beta_2)} = -\frac{1}{2}$.

$\beta_3 = (0,1,0,1) - \frac{2}{3}(1,1,0,1) + \frac{1}{6}(1,1,0,-2) = \frac{1}{2}(-1,1,0,0)$.

Therefore an orthogonal basis of the subspace is

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

and the corresponding orthonormal basis is

$$\left\{\frac{1}{\sqrt{3}}(1,1,0,1), \frac{1}{\sqrt{6}}(1,1,0,-2), \frac{1}{\sqrt{2}}(-1,1,0,0)\right\}.$$

4.13. Orthogonal complement of a subspace.

Theorem 4.13.1. In a Euclidean space V if a vector be orthogonal to a set of vectors, then it is orthogonal to every vector belonging to the subspace spanned by the set of vectors.

Proof. Let $\alpha \in V$ and α be orthogonal to the vectors $\beta_1, \beta_2, \ldots, \beta_r$ in V. Then $(\alpha, \beta_1) = 0$, $(\alpha, \beta_2) = 0, \ldots, (\alpha, \beta_r) = 0$.

Let $P = L\{\beta_1, \beta_2, \ldots, \beta_r\}$ and $\xi \in P$. Then $\xi = c_1\beta_1 + c_2\beta_2 + \cdots + c_r\beta_r$ for $c_i \in \mathbb{R}$.

$$\begin{aligned}(\alpha, \xi) &= (\alpha, c_1\beta_1) + \cdots + (\alpha, c_r\beta_r)\\ &= c_1(\alpha, \beta_1) + c_2(\alpha, \beta_2) + \cdots + c_r(\alpha, \beta_r)\\ &= 0, \text{ since } (\alpha, \beta_i) = 0 \text{ for } i = 1, 2, \ldots, r.\end{aligned}$$

This proves that α is orthogonal to every vector of P.

Note. In this case α is said to be *orthogonal to the subspace P*.

Theorem 4.13.2. Let P be a subspace of a finite dimensional Euclidean space V. A vector α in V is orthogonal to P if and only if α is orthogonal to every vector of a generating set of P.

Proof. Let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be a set of generators of P.

Let α be orthogonal to P. Then α is orthogonal to every vector of P and therefore α is orthogonal to each β_i, $i = 1, 2, \ldots, r$.

Conversely, let α be orthogonal to each β_i. By the previous theorem, α is orthogonal to $L\{\beta_1, \beta_2, \ldots, \beta_r\}$, i.e., α is orthogonal to P.

Theorem 4.13.3. Let P be a subspace of a finite dimensional Euclidean space V. The set of all vectors in V which are orthogonal to P is a subspace of V.

Proof. Let S be the set. Since θ in V is orthogonal to every vector in P, $\theta \in S$ and therefore S is non-empty.

Case 1. $S = \{\theta\}$. Then S is a subspace of V.

Case 2. Let $S \neq \{\theta\}$ and let $\alpha \in S$.

Let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be a set of generators of P. Since α is orthogonal to P, α is orthogonal to each of $\beta_1, \beta_2, \ldots, \beta_r$.

Then $(\alpha, \beta_i) = 0$ for $i = 1, 2, \ldots, r$ and this implies $(c\alpha, \beta_i) = 0$ for all $c \in \mathbb{R}$ and for $i = 1, 2, \ldots, r$.

This shows $c\alpha$ is orthogonal to $L\{\beta_1, \beta_2, \ldots, \beta_r\}$, i.e., $c\alpha$ is orthogonal to P for all $c \in \mathbb{R}$.

Therefore $c\alpha \in S$ for all $c \in \mathbb{R}$... ... (i)

Let $\alpha, \beta \in S$. Then $(\alpha, \beta_i) = 0$ and $(\beta, \beta_i) = 0$ for $i = 1, 2, \ldots, r$. This implies $(\alpha + \beta, \beta_i) = 0$ for $i = 1, 2, \ldots, r$.

This shows $\alpha + \beta$ is orthogonal to $L\{\beta_1, \beta_2, \ldots, \beta_r\}$, i.e., $\alpha + \beta$ is orthogonal to P. Therefore $\alpha \in S, \beta \in S$ implies $\alpha + \beta \in S$... ... (ii)

From (i) and (ii) it follows that S is a subspace of V.

Note. This subspace S is denoted by $P^\perp$.

Theorem 4.13.4. Let P be a subspace of a finite dimensional Euclidean space V. Then $V = P \oplus P^\perp$.

Proof. **Case 1.** Let $P = \{\theta\}$. Then $P^\perp = V$ and the theorem is obvious.

Case 2. Let $P \neq \{\theta\}$ and let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthogonal basis of P. This can be extended to an orthogonal basis $\{\beta_1, \beta_2, \ldots, \beta_r, \beta_{r+1}, \ldots, \beta_n\}$ of V.

β_{r+1} is orthogonal to each of $\beta_1, \beta_2, \ldots, \beta_r$.

Therefore $\beta_{r+1} \in P^\perp$. Similarly, $\beta_{r+2}, \ldots, \beta_n \in P^\perp$.

Since $\{\beta_{r+1}, \beta_{r+2}, \ldots, \beta_n\}$ is an orthogonal set, it is linearly independent in V and this being therefore linearly independent in $P^\perp$ is either a basis of $P^\perp$, or can be extended to a basis of $P^\perp$.

Therefore $n - r \leq \dim P^\perp < n$.

P and $P^\perp$ being subspaces of V, $P + P^\perp$ is a subspace of V and furthermore $P \cap P^\perp = \{\theta\}$, since θ is the only vector in V orthogonal to itself.

The relation $\dim (P + P^\perp) = \dim P + \dim P^\perp - \dim (P \cap P^\perp)$ gives $\dim (P + P^\perp) = \dim P + \dim P^\perp \geq r + (n - r)$, i.e., $\geq n$ ... ... (i)

Again, $P + P^\perp$ being a subspace of V, $\dim (P + P^\perp) \leq n$ ... ... (ii)

From (i) and (ii), $\dim P + P^\perp = n$ and this implies $P + P^\perp = V$.

Hence $V = P \oplus P^\perp$ since $P + P^\perp = V$ and $P \cap P^\perp = \{\theta\}$.

This completes the proof.

Note. $\{\beta_{r+1}, \beta_{r+2}, \ldots, \beta_n\}$ is an orthogonal basis of $P^\perp$.

Definition. The subspace $P^\perp$ is said to be the *orthogonal complement* of P in V.

Since an orthogonal set of non null vectors in a Euclidean space V can be extended to an orthogonal basis of V in a unique way, $P^\perp$ is uniquely determined by P. Thus although a given subspace P may have many complements in V, its orthogonal complement $P^\perp$ is unique.

Thus a finite dimensional Euclidean space V is the direct sum of any subspace of V and its orthogonal complement in V.

Worked Examples.

1. In Euclidean space $\mathbb{R}^3$ with standard inner product, let P be the subspace generated by the vectors (1,1,0) and (0,1,1). Find $P^\perp$.

Let $\alpha = (1,1,0), \beta = (0,1,1)$ and let $\gamma = (a_1, a_2, a_3) \in P^\perp$.

Then $(\alpha, \gamma) = 0$ and $(\beta, \gamma) = 0$. Therefore $a_1 + a_2 = 0, a_2 + a_3 = 0$.

Taking $a_2 = k$, we have $a_1 = a_3 = -k$ and therefore $\gamma = k(-1, 1, -1)$, where $k \in \mathbb{R}$.

So $P^\perp$ is the subspace generated by the vector $(-1, 1, -1)$.

2. A is a real $m \times n$ matrix. Show that the solution space of the system of equations $AX = O$ is the orthogonal complement of the row space of A in the Euclidean space $\mathbb{R}^n$ with standard inner product.

Let $A = (a_{ij})_{m,n}$, $a_{ij} \in \mathbb{R}$ and $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the row vectors of A. The row space $P = L\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, a subspace of $\mathbb{R}^n$.

Let Q be the solution space of $AX = O$ and let $\xi = (t_1, t_2, \ldots, t_n)$ be a solution of the system.

Then $a_{i1}t_1 + a_{i2}t_2 + \cdots + a_{in}t_n = 0$ for $i = 1, 2, \ldots, m$, i.e., $(\alpha_i, \xi) = 0$ for $i = 1, 2, \ldots, m$.

Therefore ξ is orthogonal to each α_i and therefore ξ is orthogonal to P, i.e., $\xi \in P^\perp$. Thus $\xi \in Q \Rightarrow \xi \in P^\perp$. Therefore $Q \subset P^\perp$.... (i)

Let $\eta = (u_1, u_2, \ldots, u_n) \in P^\perp$. Then η is orthogonal to each of the generators $\alpha_1, \alpha_2, \ldots, \alpha_m$ of P.

Therefore $a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n = 0$ for $i = 1, 2, \ldots, m$.

This shows that η is a solution of the system, i.e., $\eta \in Q$.

Thus $\eta \in P^\perp \Rightarrow \eta \in Q$. Therefore $P^\perp \subset Q$. ... (ii)

From (i) and (ii) $P^\perp = Q$. That is, the solution space of the system $AX = O$ is the orthogonal complement of the row space of A.

Note. Since $P \oplus P^\perp = \mathbb{R}^n$, we have $P \oplus Q = \mathbb{R}^n$.

Therefore dim P+dim $Q = n$, i.e., rank of A +rank of $X(A) = n$, where $X(A)$ is the solution space.

3. P is a subspace of a Euclidean space V of finite dimension. Prove that $P^{\perp\perp} = P$.

Let dim $V = n$. Let $\{\beta_1, \beta_2, \ldots, \beta_r\}$ be an orthogonal basis of P and let $\{\beta_1, \ldots, \beta_r, \beta_{r+1}, \ldots, \beta_n\}$ be an extended orthogonal basis of V.

$P = L\{\beta_1, \beta_2, \ldots, \beta_r\}$. Since $P \oplus P^\perp = V$, $P^\perp = L\{\beta_{r+1}, \ldots, \beta_n\}$.

Since $P^{\perp\perp}$ is the orthogonal complment of $P^\perp$ in V and $P^\perp = L\{\beta_{r+1}, \ldots, \beta_n\}$, we have $P^{\perp\perp} = L\{\beta_1, \beta_2, \ldots, \beta_r\}$. Therefore $P^{\perp\perp} = P$.

Exercises 12

1. In $\mathbb{R}^3$, let $\alpha = (a_1, a_2, a_3), \beta = (b_1, b_2, b_3)$. Determine whether $(,)$ is a real inner product for $\mathbb{R}^3$ if $(,)$ be defined by

(i) $(\alpha, \beta) = |\, a_1b_1 + a_2b_2 + a_3b_3 \,|$;

(ii) $(\alpha, \beta) = (a_1 + a_2 + a_3)(b_1 + b_2 + b_3)$;

(iii) $(\alpha, \beta) = a_1b_1 + (a_2 + a_3)(b_2 + b_3)$;

(iv) $(\alpha, \beta) = a_1b_1 + (a_2 + a_3)(b_2 + b_3) + a_3b_3$.

2. Prove that for vectors α, β in a Euclidean space V,

(i) $(\alpha, \beta) = 0$ if and only if $\|\alpha + \beta\| = \|\alpha - \beta\|$;

(ii) $(\alpha, \beta) = 0$ if and only if $\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2$;

(iii) $(\alpha + \beta, \alpha - \beta) = 0$ if and only if $\|\alpha\| = \|\beta\|$.

3. Prove that the set of vectors $\{(2, 3, -1), (1, -2, -4), (2, -1, 1)\}$ is an orthogonal basis of the Euclidean space $\mathbb{R}^3$ with standard inner product. Find the projections of the vector $\alpha = (1, 1, 1)$ along these basis vectors and verify that α is the sum of its projections along these basis vectors.

4. Use Gram-Schmidt process to convert the given basis of the Euclidean space $\mathbb{R}^3$ with standard inner product into an orthogonal basis.

(i) $\{(1, 2, -2), (2, 0, 1), (1, 1, 0)\}$; (ii) $\{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$.

5. Apply Gram-Schmidt process to find an orthonormal basis for the Euclidean space $\mathbb{R}^3$ with standard inner product, that contains the vectors

(i) $(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0)$; (ii) $(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}})$, $(\frac{2}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{1}{\sqrt{6}})$.

6. Apply Gram-Schmidt process to obtain an orthonormal basis of the subspace of the Euclidean space $\mathbb{R}^4$ with standard inner product, spanned by the vectors

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

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

7. Find an orthonormal basis of the row space of the matrix:

(i) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 0 \\ 2 & 1 & 2 & 3 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 3 & 1 & 1 \\ 1 & 2 & 3 & 1 \end{pmatrix}$.

8. Find the orthogonal complement of the row space of the matrix:

(i) $\begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 3 & 2 & 2 \\ 0 & 1 & 1 & 2 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 1 & 2 \\ 2 & 3 & 5 \\ 3 & 4 & 7 \end{pmatrix}$.

[Hint. The solution space of the system of equations $AX = O$ with the given matrix A as the coefficient matrix, is the orthogonal complement of the row space of A.]

4.14. Matric polynomials.

Let us consider a 2×2 matrix $A = \begin{pmatrix} x^2+x+1 & x^3+2x \\ 3x^3+x & 4x^2+3 \end{pmatrix}$

whose elements are real polynomials in x. A can be expressed as the polynomial in x

$$\begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} x^3 + \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} x^2 + \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

whose coefficients are real matrices of order 2×2.

Such a polynomial is said to be a *matric polynomial*. The *degree* of the matric polynomial is the degree of the constituent polynomial of highest degree appearing in the matrix A.

In general, if A be an $n \times n$ matrix whose elements are real (complex) polynomials in x, then A can be expressed as a matric polynomial whose coefficients are $n \times n$ real (complex) matrices.

Two matric polynomials $F(x)$ and $G(x)$ whose coefficients are matrices of the same order over the same field are said to be *equal* if they have the same degree and the coefficients of like powers of x be equal matrices.

Let $F(x) = A_0 + A_1x + \cdots + A_nx^n$,

$G(x) = B_0 + B_1x + \cdots + B_mx^m$

be two matric polynomials whose coefficients are square matrices of the same order over the same field. Then the sum $F(x) + G(x)$ and the product $F(x)G(x)$ are defined by

$$\begin{aligned} F(x) + G(x) &= (A_0 + B_0) + (A_1 + B_1)x + \cdots + (A_m + B_m)x^m \\ &\quad + A_{m+1}x^{m+1} + \cdots + A_nx^n, \text{ if } m < n \\ &= (A_0 + B_0) + (A_1 + B_1)x + \cdots + (A_n + B_n)x^n \\ &\quad + B_{n+1}x^{n+1} + \cdots + B_mx^m, \text{ if } n < m \\ &= (A_0 + B_0) + (A_1 + B_1)x + \cdots + (A_n + B_n)x^n, \\ &\qquad \text{if } m = n. \end{aligned}$$

$F(x)G(x) = C_0 + C_1x + \cdots + C_{m+n}x^{m+n}$, where

$C_r = A_0B_r + A_1B_{r-1} + \cdots + A_rB_0$, $r = 1, 2, \ldots, m+n$

taking $A_{n+1} = A_{n+2} = \cdots = A_{n+m} = 0$, $B_{m+1} = B_{m+2} = \cdots = B_{m+n} = 0$.

Note. $G(x)F(x)$ can be defined in a similar manner. $F(x)G(x) \neq G(x)F(x)$, in general, because matrix multiplication is not commutative.

Examples.

1. Let $F(x) = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 5 \\ 3 & 0 \end{pmatrix} x + \begin{pmatrix} 0 & 4 \\ 1 & 6 \end{pmatrix} x^2$,

$G(x) = \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix} x$.

Then $F(x) + G(x) = \begin{pmatrix} 3 & 3 \\ 4 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 6 \\ 7 & 2 \end{pmatrix} x + \begin{pmatrix} 0 & 4 \\ 1 & 6 \end{pmatrix} x^2$.

$F(x)G(x) =$

$$\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix} + \left[\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 5 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}\right] x$$

$$+ \left[\begin{pmatrix} 1 & 5 \\ 3 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 4 \\ 1 & 6 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}\right] x^2 + \begin{pmatrix} 0 & 4 \\ 1 & 6 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix} x^3$$

$$= \begin{pmatrix} 11 & 6 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 28 & 18 \\ 7 & 3 \end{pmatrix} x + \begin{pmatrix} 32 & 19 \\ 19 & 15 \end{pmatrix} x^2 + \begin{pmatrix} 16 & 8 \\ 24 & 13 \end{pmatrix} x^3.$$

$G(x)F(x) =$

$$\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} + \left[\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 5 \\ 3 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}\right] x$$

$$+ \left[\begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 0 & 4 \\ 1 & 6 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 1 & 5 \\ 3 & 0 \end{pmatrix}\right] x^2 + \begin{pmatrix} 0 & 1 \\ 4 & 2 \end{pmatrix}\begin{pmatrix} 0 & 4 \\ 1 & 6 \end{pmatrix} x^3$$

$$= \begin{pmatrix} 2 & 3 \\ 8 & 11 \end{pmatrix} + \begin{pmatrix} 2 & 6 \\ 19 & 29 \end{pmatrix} x + \begin{pmatrix} 3 & 4 \\ 12 & 44 \end{pmatrix} x^2 + \begin{pmatrix} 1 & 6 \\ 2 & 28 \end{pmatrix} x^3.$$

2. Let $F(x) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}$.

Express adj $F(x)$ as a matric polynomial and verify that $F(x)$. adj $F(x) = \det F(x) I_2$.

$F(x) = \begin{pmatrix} x & 2x+1 \\ x+1 & 3 \end{pmatrix}$.

$\operatorname{adj} F(x) = \begin{pmatrix} 3 & -2x-1 \\ -x-1 & x \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ -1 & 1 \end{pmatrix} x + \begin{pmatrix} 3 & -1 \\ -1 & 0 \end{pmatrix}$.

$F(x)$. adj $F(x) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -2 \\ -1 & 1 \end{pmatrix} x^2 +$

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 0 & -2 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -1 & 0 \end{pmatrix}\right] x +$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ -1 & 0 \end{pmatrix} =$$

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} x^2 + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} x + \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$= -2I_2x^2 + Ox - I_2 = (-2x^2 - 1)I_2$ and $\det F(x) = -2x^2 - 1$.

Therefore $F(x)$. adj $F(x) = \det F(x).I_2$ (verified).

4.15. Characteristic equation.

Let A be an $n \times n$ matrix over a field F. Then det $(A - xI_n)$ is said to be the *characteristic polynomial* of A and is denoted by $\Psi_A(x)$. The equation $\psi_A(x) = 0$ is said to be the *characteristic equation* of A.

Let $A = (a_{ij})$. Then $\Psi_A(x) =$

$$\begin{vmatrix} a_{11}-x & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-x & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-x \end{vmatrix}$$

$= c_0x^n + c_1x^{n-1} + ... + c_n$, where $c_0 = (-1)^n$ and $c_r = (-1)^{n-r}$.[sum of the principal minors of A of order r].

In particular, $c_1 = (-1)^{n-1}(a_{11} + a_{22} + ... + a_{nn}) = (-1)^{n-1}$ trace A, and $c_n = \det A$.

The degree of the characteristic equation is same as the order of the matrix A and the coefficients are scalars belonging to F.

[Note that the determinant of the submatrix of an $n \times n$ matrix A obtained by deleting $i_1th, i_2th, \ldots, i_{n-r}th$ rows and $i_1th, i_2th, \ldots, i_{n-r}th$ columns, is a *principal minor* of order r of A.]

Theorem 4.15.1. Cayley-Hamilton theorem.

Every square matrix satisfies its own characteristic equation.

The theorem states that if A be an $n \times n$ matrix and the characteristic polynomial of A be $c_0x^n + c_1x^{n-1} + \cdots + c_n$, then

$$c_0A^n + c_1A^{n-1} + \cdots + c_nI_n = O.$$

Proof. Let A be an $n \times n$ matrix. Then det $(A - xI_n) = c_0x^n + c_1x^{n-1} + \cdots + c_n$. $A - I_nx$ is a matric polynomial in x of degree 1 and adj $(A - I_nx)$ is a matric polynomial in x of degree $n-1$, since each element of adj $(A - xI_n)$ (i.e., a cofactor of an element of the matrix $A - I_nx$) is a polynomial in x of degree $n-1$ at most.

Let adj $(A - I_nx) = B_0x^{n-1} + B_1x^{n-2} + \cdots + B_{n-1}$, where each B_i is an $n \times n$ matrix.

$(A - xI_n)$. adj $(A - xI_n)$ =[det $(A - xI_n)$] I_n gives

$(A - I_nx)(B_0x^{n-1} + B_1x^{n-2} + \cdots + B_{n-1}) = (c_0x^n + c_1x^{n-1} + \cdots + c_n)I_n$

or, $A(B_0x^{n-1} + B_1x^{n-2} + \cdots + B_{n-1}) - (B_0x^n + B_1x^{n-1} + \cdots + B_{n-1}x)$
$= (c_0I_n)x^n + (c_1I_n)x^{n-1} + \cdots + (c_nI_n)$.

Equating coefficients of like powers of x, we have

$$\begin{aligned} -B_0 &= c_0I_n, \\ AB_0 - B_1 &= c_1I_n, \\ AB_1 - B_2 &= c_2I_n, \\ \cdots & \quad \cdots \\ AB_{n-2} - B_{n-1} &= c_{n-1}I_n, \\ AB_{n-1} &= c_nI_n. \end{aligned}$$

Pre-multiplying the relations by $A^n, A^{n-1}, A^{n-2}, \ldots, A, I_n$ respectively and adding, we have $c_0A^n + c_1A^{n-1} + \cdots + c_{n-1}A + c_nI_n = O$. This completes the proof.

Cayley-Hamilton theorem gives a method of computing A^{-1} when A is a non-singular matrix.

Let the characteristic equation of A be $c_0x^n + c_1x^{n-1} + \cdots + c_n = 0$. By Cayley-Hamilton theorem, $c_0A^n + c_1A^{n-1} + \cdots + c_{n-1}A + c_nI_n = O$.

Since $c_n = \det A \neq 0$, c_n^{-1} exists in F. Multiplying by $-c_n^{-1}$, we have
$-c_n^{-1}(c_0A^n + c_1A^{n-1} + \cdots + c_{n-1}A) - I_n = O$
or, $-c_n^{-1}(c_0A^{n-1} + c_1A^{n-2} + \cdots + c_{n-1}I_n)A = I_n$.

From the definition and uniqueness of an inverse it follows that
$A^{-1} = -c_n^{-1}(c_0A^{n-1} + c_1A^{n-2} + \cdots + c_{n-1}I_n)$.

Thus A^{-1} is expressed as a polynomial in A with scalar coefficients.

Worked Examples.

1. Use Cayley-Hamilton theorem to find A^{-1}, where $A = \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix}$.

The characteristic equation of A is $\begin{vmatrix} 2-x & 1 \\ 3 & 5-x \end{vmatrix} = 0$

or, $x^2 - 7x + 7 = 0$.

By Cayley-Hamilton theorem, $A^2 - 7A + 7I_2 = O$

or, $A(A - 7I_2) = -7I_2$ or, $-\frac{1}{7}A(A - 7I_2) = I_2$.

This gives $A^{-1} = -\frac{1}{7}(A - 7I_2) = \frac{1}{7}\begin{pmatrix} 5 & -1 \\ -3 & 2 \end{pmatrix}$.

2. Use Cayley-Hamilton theorem to find A^{50}, where $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

The characteristic equation of A is $x^2 - 2x + 1 = 0$.

By Cayley-Hamilton theorem, $A^2 - 2A + I_2 = O$ or, $A^2 - A = A - I_2$.

Therefore $A^3 - A^2 = A^2 - A = A - I_2, \cdots, A^{50} - A^{49} = A - I_2$.

Adding, we have $A^{50} = 50A - 49I_2 = \begin{pmatrix} 1 & 50 \\ 0 & 1 \end{pmatrix}$.

4.16. Eigen value of a matrix.

A root of the characteristic equation of a square matrix A is said to be an *eigen value* (or a *characteristic value*) of A.

Although the coefficients of $\Psi_A(x)$ are elements of F, the eigen values of A may not be all elements of F. But they all belong to a suitable algebraic extension of the field F.

For example, if the ground field of A be $\mathbb{R}$ then $\Psi_A(x)$ is a real polynomial but the roots of $\Psi_A(x) = 0$ may not be all real. They are all elements of the field $\mathbb{C}$ which is an algebraic extension of the field $\mathbb{R}$.

A root of $\Psi_A(x) = 0$ of multiplicity r is said to be an *r-fold* eigen value of A.

Theorem 4.16.1. The product of the eigen values of a square matrix A is det A.

Proof. Let A be an $n \times n$ matix and let the characteristic equation of A be $c_0x^n + c_1x^{n-1} + c_2x^{n-2} + \cdots + c_n = 0$.

Then $c_0 = (-1)^n$, $c_n =$det A.

The product of the roots of the equation is $(-1)^n \frac{c_n}{c_0} = c_n =$ det A.

Hence the product of the eigen values of A is det A.

Theorem 4.16.2. If A be a singular matrix, 0 is an eigen value of A.

Proof. Since A is singular, det $A = 0$.

Let $\Psi_A(x) = c_0x^n + c_1x^{n-1} + c_2x^{n-2} + \cdots + c_n$.

Then $c_n =$ det $A = 0$. Consequently, 0 is a root of the characteristic equation of A and therefore 0 is an eigen value of A.

Theorem 4.16.3. The eigen values of a diagonal matrix are its diagonal elements.

Proof. Let $A = \text{diag}\,(d_1, d_2, ..., d_n)$.

Then det $(A - xI_n) = (d_1 - x)(d_2 - x)...(d_n - x)$.

So the roots of the characteristic equation of A are $d_1, d_2, ..., d_n$ and hence the eigen values of A are $d_1, d_2, ..., d_n$.

Theorem 4.16.4. If λ be an r-fold eigen value of A, 0 is an r-fold eigen value of the matrix $A - \lambda I_n$.

Proof. Let $\Psi_A(x) = \det\,(A - xI_n) = (x-\lambda)^r\phi(x)$, where ϕ is a polynomial of degree $n - r$ and $\phi(\lambda) \neq 0$.

The characteristic polynomial of $A - \lambda I_n$ is det $(A - \lambda I_n - xI_n)$.

$$\begin{aligned} \det\,(A - \lambda I_n - xI_n) &= \det\,[A - (\lambda + x)I_n] \\ &= (\lambda + x - \lambda)^r\phi(\lambda + x) \\ &= x^r\mu(x), \text{ where } \mu(0) = \phi(\lambda) \neq 0. \end{aligned}$$

This proves that 0 is a root of multiplicity r of the characteristic equation of the matrix $A - \lambda I_n$.

Theorem 4.16.5. If λ be an eigen value of a non-singular matrix A, then λ^{-1} is an eigen value of A^{-1}.

Proof. Let the order of A be n. Since A is non-singular, A^{-1} exists. Also $\lambda \neq 0$. Therefore λ^{-1} exists.

Since λ is an eigen value of A, det $(A - \lambda I_n) = 0$.

$$\begin{aligned} \det\,(A^{-1} - \lambda^{-1}I_n) &= (\det A)^{-1} \det\,(AA^{-1} - \lambda^{-1}A) \\ &= (\det A)^{-1}(\lambda^{-1})^n \det\,(\lambda I_n - A) \\ &= (\det A)^{-1}(\lambda^{-1})^n(-1)^n \det\,(A - \lambda I_n) \\ &= 0, \text{ since } \det\,(A - \lambda I_n) = 0. \end{aligned}$$

This proves that λ^{-1} is an eigen value of A^{-1}.

Theorem 4.16.6. If A and P be both $n \times n$ matrices and P be non-singular, then A and $P^{-1}AP$ have the same eigen values.

Proof. The characteristic polynomial of $P^{-1}AP$ is det $(P^{-1}AP - xI_n)$.

$$\begin{aligned} \det\,(P^{-1}AP - xI_n) &= \det\,[P^{-1}AP - P^{-1}(xI_n)P], \\ &\qquad \text{since } P^{-1}(xI_n)P = xI_n \\ &= \det\,[P^{-1}(A - xI_n)P] \\ &= \det P^{-1} \det\,(A - xI_n) \det P \\ &= \det\,(A - xI_n) \det\,(P^{-1}P) \\ &= \det\,(A - xI_n) \det\,(I_n) \\ &= \det\,(A - xI_n). \end{aligned}$$

Therefore the matrices $P^{-1}AP$ and A have the same characteristic polynomial and so they have the same eigen values.

Examples.

1. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

The characteristic equation of A is $\begin{vmatrix} -x & -1 \\ 1 & -x \end{vmatrix} = 0$

or, $x^2 + 1 = 0$.

The eigen values of A are $i, -i$.

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

The characteristic equation of A is $\begin{vmatrix} 1-x & -1 & 0 \\ 1 & 2-x & -1 \\ 3 & 2 & -2-x \end{vmatrix} = 0$

or, $(1-x)(x^2-2)+(1-x)=0$
or, $(1-x)(x^2-1)=0$.
The eigen values of A are $1, 1, -1$.

4.17. Eigen vectors of a matrix.

Definition. Let A be $n \times n$ matrix over a field F. A non-null vector X belonging to $V_n(F)$ is said to be an *eigen vector* or a *characteristic vector* of A if there exists a scalar λ belonging to F such that $AX = \lambda X$ holds.

Let there exist an eigen vector X of the matrix. Then for some suitable scalar λ, $AX = \lambda X$ holds. That is, $(A - \lambda I_n)X = O$.

This is a homogeneous system of n equations in n unknowns. Since there exists a non-null solution of the system, det $(A - \lambda I_n) = 0$.

This implies that λ is an eigen value of A. Thus for an eigen vector, if it exists, there corresponds an eigen value of the matrix.

Theorem 4.17.1. Let A be an $n \times n$ matrix over a field F. To an eigen vector of A there corresponds a unique eigen value of A.

Proof. Let there be two distinct eigen values λ_1, λ_2 of A corresponding to an eigen vector X. Then $AX = \lambda_1 X$ and $AX = \lambda_2 X$

Therefore $(\lambda_1 - \lambda_2)X = O$. But this is a contradiction, since X is a non-null vector and $\lambda_1 - \lambda_2 \neq 0$. Hence the theorem.

Theorem 4.17.2. Let A be an $n \times n$ matrix over a field F and λ be an eigen value belonging to F. To each such eigen value of A there corresponds at least one eigen vector.

Proof. Since λ is an eigen value, det $(A - \lambda I_n) = 0$. Therefore the homogeneous system of equations $(A - \lambda I_n)X = O$ has a non-null solution, say $X = X_1$ where $X_1 \in V_n(F)$.

Then $(A - \lambda I_n)X_1 = O$ or, $AX_1 = \lambda X_1$.

This shows that X_1 is an eigen vector of A corresponding to λ. This proves the theorem.

Note. In fact, there are many eigen vectors of A corresponding to an eigen value λ belonging to F, because det $(A - \lambda I_n) = 0$ implies that there are infinite number of non-null solutions of the homogeneous system $(A - \lambda I_n)X = 0$ and each such non-null solution gives an eigen vector of A corresponding to λ.

Examples.

1. Let $A = \begin{pmatrix} 1 & 3 \\ 4 & 5 \end{pmatrix}$. The eigen values of A are $-1, 7$.

Let $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be an eigen vector corresponding to -1.

Then $AX = -X$ and this gives $\begin{aligned} 2x_1 + 3x_2 &= 0 \\ 4x_1 + 6x_2 &= 0. \end{aligned}$

The equivalent system is $x_1 + \frac{3}{2}x_2 = 0$.
The solution of the system is $k(-\frac{3}{2}, 1)$, where $k \in \mathbb{R}$.

The eigen vectors are $k\begin{pmatrix} -\frac{3}{2} \\ 1 \end{pmatrix}$ or equivalently, $c\begin{pmatrix} 3 \\ -2 \end{pmatrix}$, where c is a *non-zero* real number.

Let $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be an eigen vector corresponding to 7.

Then $AX = 7X$ and this gives $\begin{aligned} -6x_1 + 3x_2 &= 0 \\ 4x_1 - 2x_2 &= 0. \end{aligned}$

The system is equivalent to $x_1 - \frac{1}{2}x_2 = 0$.
The solution of the system is $c(1, 2)$, where $c \in \mathbb{R}$.

Therefore the eigen vectors are $c\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, where c is a *non-zero* real number.

2. Let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. The eigen values of A are $i, -i$.

A is a real matrix and the eigen values of A are not real numbers. Therefore the real matrix A has no eigen vector.

But if A be considered as a complex matrix, then the eigen vectors of A corresponding to the eigen values i and $-i$ can be obtained.

Let $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ be an eigen vector corresponding to i.

Then $AX = iX$ and this gives $\begin{aligned} -ix_1 - x_2 &= 0 \\ x_1 - ix_2 &= 0. \end{aligned}$

The equivalent system is $x_1 - ix_2 = 0$.
The solution is $k(i, 1)$, where $k \in \mathbb{C}$.

The eigen vectors are $k\begin{pmatrix} i \\ 1 \end{pmatrix}$, where k is a *non-zero* complex number.

The eigen vectors corresponding to $-i$ are $c\begin{pmatrix} 1 \\ i \end{pmatrix}$, where c is a *non-zero* complex number.

Theorem 4.17.3. Two eigen vectors of a square matrix A over a field F corresponding to two distinct eigen values of A are linearly independent.

Proof. Let X_1, X_2 be the eigen vectors of A corresponding to two distinct eigen values λ_1, λ_2 respectively. Then $AX_1 = \lambda_1 X_1,\ AX_2 = \lambda_2 X_2$.

Let us consider the relation $c_1X_1 + c_2X_2 = 0$, where $c_1, c_2 \in F$.

Then $c_1AX_1 + c_2AX_2 = 0$, or, $c_1\lambda_1X_1 + c_2\lambda_2X_2 = 0$.

Therefore $(c_1X_1\ c_2X_2)\begin{pmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{pmatrix} = (0\ 0)$ ... (i)

Let $P = \begin{pmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{pmatrix}$. Since $\lambda_1 \neq \lambda_2$, P is non-singular. Postmultiplying both sides of (i) by P^{-1}, we have $(c_1X_1\ c_2X_2) = (0\ 0)$.

This implies $c_1 = 0, c_2 = 0$, since X_1, X_2 are both non-zero.

This proves that X_1, X_2 are linearly independent.

Theorem 4.17.4. If X_1, X_2, X_3 be three eigen vectors of a square matrix A over a field F corresponding to three distinct eigen values $\lambda_1, \lambda_2, \lambda_3$ respectively, then X_1, X_2, X_3 are linearly independent.

Proof. $AX_1 = \lambda_1X_1, AX_2 = \lambda_2X_2, AX_3 = \lambda_3X_3$.

Let us consider the relation $c_1X_1 + c_2X_2 + c_3X_3 = 0$, where $c_1, c_2, c_3 \in F$.

Then $c_1AX_1 + c_2AX_2 + c_3AX_3 = 0$
or, $c_1\lambda_1X_1 + c_2\lambda_2X_2 + c_3\lambda_3X_3 = 0$.

Also we have $c_1A^2X_1 + c_2A^2X_2 + c_3A^2X_3 = 0$
or, $c_1{\lambda_1}^2X_1 + c_2{\lambda_2}^2X_2 + c_3{\lambda_3}^2X_3 = 0$.

Therefore $(c_1X_1\ c_2X_2\ c_3X_3)\begin{pmatrix} 1 & \lambda_1 & {\lambda_1}^2 \\ 1 & \lambda_2 & {\lambda_2}^2 \\ 1 & \lambda_3 & {\lambda_3}^2 \end{pmatrix} = (0\ 0\ 0)$ ... (i)

Let $P = \begin{pmatrix} 1 & \lambda_1 & {\lambda_1}^2 \\ 1 & \lambda_2 & {\lambda_2}^2 \\ 1 & \lambda_3 & {\lambda_3}^2 \end{pmatrix}$. Since $\lambda_1, \lambda_2, \lambda_3$ are all different, P is non-singular. Postmultiplying both sides of (i) by P^{-1}, we have $(c_1X_1\ c_2X_2\ c_3X_3) = (0\ 0\ 0)$.

This implies $c_1 = 0, c_2 = 0, c_3 = 0$, since X_1, X_2, X_3 are all non-zero.

This proves that X_1, X_2, X_3 are linearly independent.

Corollary. If $X_1, X_2, \ldots, X_r$ be r eigen vectors of an $n \times n$ matrix A corresponding to r distinct eigen values $\lambda_1, \lambda_2, \ldots, \lambda_r$ respectively, then $X_1, X_2, \ldots, X_r$ are linearly independent.

Theorem 4.17.5. The eigen vectors of an $n \times n$ matrix A over a field F corresponding to an eigen value $\lambda \in F$, together with the null- vector, form a vector space, a subspace of $V_n(F)$.

Proof. To an eigen value $\lambda \in F$, there corresponds an eigen vector of A. Let S be the set of all eigen vectors of A corresponding to λ. Then S is the set of all non-null solutions of the homogeneous system of equations $(A - \lambda I_n)X = O$.

The null vector $\theta \in V_n(F)$ is also a solution of the system.

As the solutions of a homogeneous system with an $n \times n$ matrix over F as the coefficient matrix form a subspace of $V_n(F)$, it follows that $S \cup \{\theta\}$ is a subspace of $V_n(F)$.

In other words, the eigen vectors corresponding to λ, together with the null-vector, form a non-null subspace of $V_n(F)$.

This completes the proof.

Definition. Characteristic subspace.

The non-null vector space formed by the eigen vectors of a matrix A corresponding to an eigen value λ, together with the null-vector, is said to be the *characteristic subspace* corresponding to λ.

Theorem 4.17.6. If λ be an r-fold eigen value of an $n \times n$ matrix A, then rank of $(A - \lambda I_n) \geq n - r$.

Proof. Since λ is an r-fold eigen value of A, 0 is an r- fold eigen value of $A - \lambda I_n$, by Theorem 4.16.4.

Let det $(A - \lambda I_n - xI_n) = a_0x^n + a_1x^{n-1} + \cdots + a_n$.

Then $a_0 = (-1)^n, a_k = (-1)^{n-k}$ [sum of all principal minors of order k of $A - \lambda I_n$], for $k = 1, 2, \ldots, n$.

Since 0 is an r-fold eigen value of $A - \lambda I_n, a_n = a_{n-1} = \ldots = a_{n-r+1} = 0, a_{n-r} \neq 0$. So the sum of all principal minors of order $n-r$ of $A - \lambda I_n$ is not zero and hence there exists at least one non-zero minor of order $n - r$ of the matrix $A - \lambda I_n$.

It follows that rank of $(A - \lambda I_n) \geq n - r$.

Corollary. If λ be a simple eigen value (i.e., of multiplicity 1) of A, then rank of $(A - \lambda I_n) = n - 1$.

Because, by the theorem, rank of $(A - \lambda I_n) \geq n - 1$ and since λ is an eigen value, det $(A - \lambda I_n) = 0$ and so rank of $(A - \lambda I_n) \leq n - 1$.

Theorem 4.17.7. If λ be an r-fold eigen value of A, the rank of the characteristic subspace corresponding to λ does not exceed r.

Proof. The characteristic subspace is the subspace of solutions of the homogeneous system $(A - \lambda I_n)X = O$.

It follows from the Theorem 4.10.3 that the rank of the characteristic subspace +rank of $(A - \lambda I_n) = n$.

Since rank of $(A - \lambda I_n) \geq n - r$, the theorem follows.

Corollary. The rank of the characteristic subspace corresponding to a simple eigen value λ is 1, since in this case, rank of $(A - \lambda I_n) = n - 1$.

Definition. For an r-fold eigen value λ, r is called the *algebraic multiplicity of* λ and the rank of the characteristic subspace corresponding to λ is called the *geometric multiplicity of* λ.

Since the characteristic subspace is always a non-null subspace, it follows that for an eigen value λ,

$$1 \leq \text{geometric multiplicity} \leq \text{algebraic multiplicity}.$$

Definition. An eigen value λ is said to be *regular* if the geometric multiplicity of λ is equal to its algebraic multiplicity.

Example (continued).

3. Let $A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

The characteristic equation of A is $x^2(1-x) = 0$ and therefore the eigen values of A are 0, 0, 1.

0 is an eigen value of algebraic multiplicity 2; and 1 is a simple eigen value of A (i.e., of algebraic multiplicity 1).

The eigen vectors corresponding to the eigen value 0 are the non-null solutions of the system

$$\begin{aligned} x + y + z &= 0 \\ z &= 0. \end{aligned}$$

The eigen vectors are $c \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, where c is a non-zero real number.

The rank of the characteristic subspace is 1.

Therefore the geometric multiplicity of the eigen value 0 is 1. So in this case,

the geometric multiplicity is less than the algebraic multiplicity.

The eigen vectors corresponding to the eigen value 1 are the non-null solutions of the system

$$\begin{aligned} y + z &= 0 \\ x + 2y + z &= 0. \end{aligned}$$

The solution of the system is $c(1, -1, 1)$, where $c \in \mathbb{R}$.

So the eigen vectors are $c \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, where c is a non-zero real number.

The rank of the characteristic subspace is 1 and therefore the geometric multiplicity of the eigen value 1 is 1. In this case, the geometric multiplicity = the algebraic multiplicity.

Theorem 4.17.8. The eigen values of a real symmetric matrix are all real.

Proof. Let A be an $n \times n$ real symmetric matrix. The characteristic equation of A is an equation with real coefficients. So the eigen values of A are complex numbers, some or all of which may be purely real.

Let λ be an eigen value of A. Then $\det(A - \lambda I_n) = 0$. Therefore there exist non-null solutions of the homogeneous system $(A - \lambda I_n)X = O$. Let X_1 be one such solution.

Then $(A - \lambda I_n)X_1 = O$. That is, $AX_1 = \lambda X_1$.

[Note that this X_1 is not an eigen vector of A unless λ is purely real.]

Taking transpose of the conjugate, we have

$(\overline{AX_1})^t = (\overline{\lambda X_1})^t$

or, $(\bar{X}_1)^t(\bar{A})^t = \bar{\lambda}(\bar{X}_1)^t$, since λ is a scalar

or, $(\bar{X}_1)^t A = \bar{\lambda}(\bar{X}_1)^t$, since $\bar{A}^t = A^t = A$.

Multiplying by X_1 from the right, we have

$(\bar{X}_1)^t A X_1 = \bar{\lambda}(\bar{X}_1)^t X_1$

or, $(\bar{X}_1)^t \lambda X_1 = \bar{\lambda}(\bar{X}_1)^t X_1$

or, $\lambda(\bar{X}_1)^t X_1 = \bar{\lambda}(\bar{X}_1)^t X_1$

or, $(\lambda - \bar{\lambda})(\bar{X}_1)^t X_1 = 0$.

But $(\bar{X}_1)^t X_1 \neq 0$, since X_1 is non-null.

It follows that $\lambda = \bar{\lambda}$ and therefore λ is purely real.

This proves the theorem.

Theorem 4.17.9. The eigen values of a real skew symmetric matrix are purely imaginary or zero.

Proof. Let A be an $n \times n$ real skew symmetric matrix. Following the same argument as in the previous theorem, we have

$(\lambda + \bar{\lambda})(\bar{X}_1)^t X_1 = 0$, since $\bar{A}^t = A^t = -A$.

Since X_1 is non-null, $\lambda + \bar{\lambda} = 0$. That is, $\lambda = -\bar{\lambda}$.

Therefore λ is purely imaginary or zero and the theorem is proved.

Note. The eigen values of a Hermitian matrix are all real.

Let λ be an eigen value of a Hermitian matrix A and X be an eigen vector corresponding to λ. Then $AX = \lambda X$. Using $\bar{A}^t = A$ and proceeding in a similar manner as in the theorem 4.17.8, the assertion can be established.

Theorem 4.17.10. The eigen vectors corresponding to two distinct eigen values of a real symmetric matrix are orthogonal.

Proof. Let A be a real symmetric matrix. Let X_1, X_2 be two eigen vectors of A corresponding to two distinct eigen values λ_1 and λ_2.

Then $AX_1 = \lambda_1 X_1$ and $AX_2 = \lambda_2 X_2$.

Now $AX_1 = \lambda_1 X_1 \quad \Rightarrow (AX_1)^t = \lambda_1 X_1^t$, since λ_1 is real
or, $X_1^t A = \lambda_1 X_1^t$, since $A = A^t$.

Multiplying by X_2 from the right, we have $X_1^t A X_2 = \lambda_1 X_1^t X_2$

or, $X_1^t \lambda_2 X_2 = \lambda_1 X_1^t X_2$

or, $(\lambda_2 - \lambda_1) X_1^t X_2 = 0$

or, $X_1^t X_2 = 0$, since $\lambda_1 \neq \lambda_2$.

Since $X_1 \neq 0$ and $X_2 \neq 0$, it follows that X_1 is orthogonal to X_2. This proves the theorem.

Theorem 4.17.11. Each eigen value of a real orthogonal matrix has unit modulus.

Proof. Let A be an $n \times n$ real orthogonal matrix. Then $AA^t = I_n$. The eigen values of A are in general, complex numbers, some of which may be purely real. Let λ be an eigen value of A. Then det $(A - \lambda I_n) = 0$.

Therefore there exists a non-null solution of the homogeneous system $(A - \lambda I_n)X = O$. Let X_1 be one such solution.

Then $(A - \lambda I_n)X_1 = O$. That is, $AX_1 = \lambda X_1$.

Taking transpose of the conjugate, we have

$\overline{(AX_1)}^t = \overline{(\lambda X_1)}^t$

or, $(\bar{X}_1)^t (\bar{A})^t = \bar{\lambda}(\bar{X}_1)^t$, since λ is a scalar

or, $(\bar{X}_1)^t A^t = \bar{\lambda}(\bar{X}_1)^t$, since $(\bar{A})^t = A^t$.

Multiplying by AX_1 from the right, we have $(\bar{X}_1)^t A^t (AX_1) = \bar{\lambda}(\bar{X}_1)^t (AX_1)$

or, $(\bar{X}_1)^t (A^t A) X_1 = \bar{\lambda}(\bar{X}_1)^t \lambda X_1$

or, $(\bar{X}_1)^t X_1 = \bar{\lambda}\lambda(\bar{X}_1)^t X_1$, since $AA^t = I_n \Rightarrow A^t A = I_n$.

This implies $(\bar{X}_1)^t X_1 (1 - \bar{\lambda}\lambda) = 0$.

Since X_1 is non-null, $(\bar{X}_1)^t X_1 \neq 0$. It follows that $\bar{\lambda}\lambda = 1$, i.e., $|\lambda| = 1$.

This proves the theorem.

Worked Examples.

1. If λ be an eigen value of a real orthogonal matrix A, prove that $\frac{1}{\lambda}$ is also an eigen value of A.

Let A be an orthogonal matrix of order n. Then $AA^t = I_n$ and A is non-singular. Since A is non-singular, $\lambda \neq 0$.

Since λ is an eigen value of A, det $(A - \lambda I_n) = 0$

or, det $(A - \lambda AA^t) = 0$

or, det A. det $(I_n - \lambda A^t) = 0$

or, det $(I_n - \lambda A^t) = 0$, since det $A \neq 0$

or, $(-1)^n \lambda^n$ det $(A^t - \frac{1}{\lambda} I_n) = 0$

or, $(-1)^n \lambda^n$ det $(A - \frac{1}{\lambda} I_n) = 0$, since det $(A^t - \frac{1}{\lambda} I_n) = \det (A - \frac{1}{\lambda} I_n)^t$
$= \det (A - \frac{1}{\lambda} I_n)$.

This proves that $\frac{1}{\lambda}$ is an eigen value of A.

2. A is a 3×3 real matrix having the eigen values 2, 3, 1.
$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ are the eigen vectors of A corresponding to the eigen values 2, 3, 1 respectively. Find the matrix A.

Let $X_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, X_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, X_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Then $AX_1 = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}, AX_2 = \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix}, AX_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Let $P = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Then $AP = \begin{pmatrix} 2 & 0 & 1 \\ 4 & 3 & 1 \\ 2 & 3 & 1 \end{pmatrix}$.

The column vectors of P are the eigen vectors of A corresponding to three distinct eigen values. Therefore P is non-singular.

$$P^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}.$$

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 4 & 3 & 1 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 3 & 0 \\ -2 & 1 & 2 \end{pmatrix}.$$

3. If S be a real skew symmetric matrix of order n prove that

(i) $I_n + S$ is non-singular,

(ii) $(I_n + S)^{-1}(I_n - S)$ is orthogonal,

(iii) if X be an eigen vector of S with eigen value λ then X is also an eigen vector of the matrix $(I_n + S)^{-1}(I_n - S)$ with eigen value $\frac{1-\lambda}{1+\lambda}$,

(iv) if $\bar{S} = (I_n + S)^{-1}(I_n - S)$ then $I_n + \bar{S}$ is also non-singular and $\bar{\bar{S}} = S$.

(i) Since S is a real skew symmetric matrix, its eigen values are imaginary or zero. Therefore -1 is not an eigen value of S. So -1 is not a root of the characteristic equation det $(S - xI_n) = 0$.

It follows that det $(S + I_n) \neq 0$. That is, $I_n + S$ is non-singular.

(ii) Let $P = (I_n + S)^{-1}(I_n - S)$.

$$\begin{aligned}
\text{Then } PP^t &= (I_n+S)^{-1}(I_n-S)[(I_n+S)^{-1}(I_n-S)]^t \\
&= (I_n+S)^{-1}(I_n-S)(I_n-S)^t\{(I_n+S)^{-1}\}^t \\
&= (I_n+S)^{-1}(I_n-S)(I_n+S)\{(I_n+S)^t\}^{-1} \\
&= (I_n+S)^{-1}\{(I_n+S)(I_n-S)\}(I_n-S)^{-1} \\
&\quad [\text{ since } (I_n-S)(I_n+S) = (I_n+S)(I_n-S)] \\
&= \{(I_n+S)^{-1}(I_n+S)\}\{(I_n-S)(I_n-S)^{-1}\} \\
&= I_n \cdot I_n = I_n.
\end{aligned}$$

This proves that P is orthogonal.

(iii) $SX = \lambda X$.

$$\begin{aligned}
\text{Therefore } (I_n+S)^{-1}(I_n-S)X &= (I_n+S)^{-1}(1-\lambda)X \\
&= (1-\lambda)(I_n+S)^{-1}X. \quad \text{... (i)}
\end{aligned}$$

Again $(I_n + S)X = (1 + \lambda)X$

or, $X = (I_n + S)^{-1}(1 + \lambda)X = (1 + \lambda)(I_n + S)^{-1}X$.

So we have $\frac{1}{1+\lambda}X = (I_n + S)^{-1}X$, since $\lambda + 1 \neq 0$. ... (ii)

From (i) and (ii) $(I_n + S)^{-1}(I_n - S)X = (1 - \lambda)\frac{1}{1+\lambda}X = \frac{1-\lambda}{1+\lambda}X$.

This proves that X is an eigen vector of $(I_n + S)^{-1}(I_n - S)$ with eigen value $\frac{1-\lambda}{1+\lambda}$.

(iv) $\bar{S} = (I_n + S)^{-1}(I_n - S)$.

$$\begin{aligned}
I_n + \bar{S} &= (I_n+S)^{-1}(I_n+S) + (I_n+S)^{-1}(I_n-S) \\
&= (I_n+S)^{-1}\{(I_n+S) + (I_n-S)\} \\
&= 2(I_n+S)^{-1}.
\end{aligned}$$

Therefore $(I_n + \bar{S})^{-1} = \frac{1}{2}(I_n + S)$, proving that $I_n + \bar{S}$ is non-singular.

$$\begin{aligned}
\text{Also } I_n - \bar{S} &= (I_n+S)^{-1}(I_n+S) - (I_n+S)^{-1}(I_n-S) \\
&= (I_n+S)^{-1}\{(I_n+S) - (I_n-S)\} \\
&= 2(I_n+S)^{-1}S.
\end{aligned}$$

$$\begin{aligned}
\text{Therefore } \bar{\bar{S}} &= (I_n+\bar{S})^{-1}(I_n-\bar{S}) \\
&= \tfrac{1}{2}(I_n+S).2(I_n+S)^{-1}.S = S.
\end{aligned}$$

4. If every non-zero vector in $\mathbb{R}^n$ be an eigen vector of a real $n \times n$ matrix A, prove that A is a scalar matrix.

First we show that A has only one repeated eigen value.

Let λ_1, λ_2 be two distinct eigen values of A. Let X_1 be an eigen vector corresponding to λ_1 and X_2 be an eigen vector corresponding to λ_2. Then $AX_1 = \lambda_1 X_1$, $AX_2 = \lambda_2 X_2$ and X_1, X_2 are linearly independent.

Since every non-zero vector in $\mathbb{R}^n$ is an eigen vector of A, $X_1 + X_2$ is also an eigen vector of A. Let λ_3 be the corresponding eigen value.

Then $A(X_1 + X_2) = \lambda_3(X_1 + X_2)$

or, $\lambda_1 X_1 + \lambda_2 X_2 = \lambda_3(X_1 + X_2)$

or, $(\lambda_1 - \lambda_3)X_1 + (\lambda_2 - \lambda_3)X_2 = 0$.

This implies $\lambda_1 = \lambda_3$, $\lambda_2 = \lambda_3$, since X_1, X_2 are linearly independent.

So $\lambda_1 = \lambda_2 = \lambda_3$, a contradiction.

So A has only one repeated eigen value λ, say. Since the rank of the characteristic subspace is n, λ is an n-fold eigen value. [Theorem 4.17.7.]

The solution space of the homogeneous system $(A - \lambda I_n)X = O$ is $\mathbb{R}^n$. Since rank of $(A - \lambda I_n)$ + dimension of the solution space $= n$, it follows that rank of $(A - \lambda I_n)$ is 0. So $A = \lambda I_n$, a scalar matrix.

Exercises 13

1. Verify Cayley-Hamilton theorem for the matrix A. Express A^{-1} as a polynomial in A and then compute A^{-1}.

(i) $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & 3 & 2 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 3 & -1 \end{pmatrix}$.

2. Use Cayley-Hamilton theorem to find A^{100}, where $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

3. (i) Find the eigen values and the corresponding eigen vectors of the matrix I_3. Generalise the result for the matrix I_n.

(ii) Find the eigen values and the corresponding eigen vectors of the scalar matrix $cI_3, c \in \mathbb{R}$. Generalise the result for the matrix cI_n.

(iii) Find the eigen values and the corresponding eigen vectors of the diagonal matrix diag (d_1, d_2, d_3). Generalise the result for the matrix diag $(d_1, d_2, \cdots, d_n)$.

4. If λ be an eigen value of an $n \times n$ matrix A, prove that

(i) λ is also an eigen value of the matrix A^t;

(ii) $k\lambda$ is an eigen value of the matrix kA, k being a scalar;

(iii) λ^2 is an eigen value of the matrix A^2.

5. $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigen values of a non-singular matrix A of order n. Find the eigen values of the matrix (i) A^{-1}, (ii) adj A.

6. If λ be an eigen value of an $n \times n$ idempotent matrix A, prove that λ is either 1 or 0.

[**Hint.** $A^2 = A$. Let $AX = \lambda X$ for some $X \neq 0$. Then $\lambda X = AX = A^2X = \lambda^2 X$.]

7. Find the eigen values and the corresponding eigen vectors of the following real matrices.

(i) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \\ 3 & -3 & 6 \end{pmatrix}$, (iii) $\begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}$.

8. Find the eigen values and the corresponding eigen vectors of the following complex matrices.

(i) $\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$, (ii) $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, (iii) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

9. Find the algebraic and the geometric multiplicities of each eigen value of the matrix $\begin{pmatrix} 1 & -1 & 0 \\ 1 & 2 & -1 \\ 3 & 2 & -2 \end{pmatrix}$.

10. $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ are eigen vectors of the matrix $\begin{pmatrix} 2 & 2 & 1 \\ a & 3 & 1 \\ 1 & b & c \end{pmatrix}$. Find a, b, c.

11. If λ be an eigen value of an $n \times n$ matrix A, prove that λ^m is an eigen value of the matrix A^m, where m is a positive integer.

[**Hint.** $AX = \lambda X$ implies $A^2X = \lambda(AX) = \lambda^2 X$ etc.]

12. If λ be an eigen value of an $n \times n$ non-singular matrix A, prove that λ^{-m} is an eigen value of the matrix A^{-m}, where m is a positive integer.

13. If A and B are square matrices of the same order and one of them is non-singular, prove that the matrices AB and BA have the same eigen values.

14. A and B are symmetric non-singular matrices such that $AB + BA = O$. Prove that each of the matrices A, B and AB have eigen values symmetric about the origin.

15. λ is an eigen value of a real skew symmetric matrix. Prove that $\left| \frac{1-\lambda}{1+\lambda} \right| = 1$.

16. If S be a real skew symmetric matrix of order n, prove that the matrices $I_n + S$ and $I_n - S$ are both non-singular.

17. If A be a real non-singular symmetric matrix, prove that the matrices A and A^{-1} have the same set of eigen vectors.

18. Let X be an eigen vector of an $n \times n$ matrix A associated with an eigen value λ. Prove that $P^{-1}X$ is an eigen vector of the matrix $P^{-1}AP$ associated with λ.

19. Show that the characteristic equation of an orthogonal matrix is a reciprocal equation.

[**Hint.** Let A be orthogonal and $\psi(x) = \det(A - xI_n)$. Then $\psi(x) = \pm x^n \psi(\frac{1}{x})$.]

20. P is a real orthogonal matrix with $\det P = -1$. Prove that -1 is an eigen value of P.

[**Hint.** $\det(P + I) = \det(P + PP^t) = \det P \det(I + P^t) = -\det(P + I)$.]

21. If S be a real skew symmetric matrix of order n prove that

(i) the matrix $S - I_n$ is non-singular,

(ii) the matrix $(S - I_n)^{-1}(S + I_n)$ is orthogonal,

(iii) if X be an eigen vector of S with eigen value λ then X is also an eigen vector of the matrix $(S - I_n)^{-1}(S + I_n)$ with eigen value $\frac{\lambda+1}{\lambda-1}$,

(iv) if $\bar{S} = (S - I_n)^{-1}(S + I_n)$ then $\bar{S} - I_n$ is also non-singular and $\bar{\bar{S}} = S$.

22. A is an $n \times n$ real symmetric matrix and $A^2 = I_n$ but $A \neq \pm I_n$. Prove that 1 and -1 are the only distinct eigen values of A.

23. A is a 3×3 real matrix having the eigen values $1, 2, 0$.

$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ are the eigen vectors of A corresponding to the eigen values $1, 2, 0$ respectively. Find the matrix A.

24. A is a 3×3 real matrix having the eigen values $5, 2, 2$. The eigen vectors of A corresponding to the eigen values 5 and 2 are

$c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, c \neq 0$ and $c\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + d\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, (c, d) \neq (0, 0)$ respectively.

Find the matrix A.

25. If every non-zero vector in $\mathbb{R}^n$ be an eigen vector of a real $n \times n$ matrix A corresponding to a real eigen value λ, prove that A is the scalar matrix λI_n.

[**Hint.** The solution space of the homogeneous system $(A - \lambda I_n)X = O$ is $\mathbb{R}^n$ of dimension n. Therefore the rank of $A - \lambda I_n$ is 0]

4.18. Diagonalisation of matrices.

Let us consider the set of all $n \times n$ matrices over a field F. An $n \times n$ matrix A is said to be *similar* to an $n \times n$ matrix B if there exists a non-singular $n \times n$ matrix P such that $B = P^{-1}AP$.

$B = P^{-1}AP$ implies $A = PBP^{-1} = Q^{-1}BQ$ where $Q(= P^{-1})$ is non-singular. Therefore if A is similar to B then B is similar to A and two matrices A and B are said to be similar.

By the Theorem 4.16.6, two similar matrices have the same eigen values. But the matrices having the same eigen values may not be similar.

For example, let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$.

These matrices have the same characteristic polynomial and hence they have the same eigen values. But A being the matrix I_2 there is no matrix other than itself which is similar to it, because for any non-singular 2×2 matrix P, $P^{-1}I_2P = I_2$. Therefore B is not similar to A.

Definition.

An $n \times n$ matrix A is said to be *diagonalisable* if A is similar to an $n \times n$ diagonal matrix.

If A is similar to a diagonal matrix $D = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$ then $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigen values of A, by Theorem 4.16.3.

We have a necessary and sufficient condition for diagonalisability of an $n \times n$ matrix in the following theorem.

Theorem 4.18.1. An $n \times n$ matrix A over a field F is diagonalisable if and only if there exist n eigen vectors of A which are linearly independent.

Proof. Let A be similar to a diagonal matrix $D = \text{diag}\ (d_1, d_2, \ldots, d_n)$. Then there exists a non-singular matrix P such that $D = P^{-1}AP$, i.e., $AP = PD$. It follows that the ith column vector of AP = the i th column vector of PD for each $i = 1, 2, \ldots, n$.

The ith column vector of $AP = A\begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$, where $\begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$ is the ith column vector of P.

The ith column vector of $PD = d_i \begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$.

As $A \begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix} = d_i \begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$ holds, $\begin{pmatrix} p_{1i} \\ p_{2i} \\ \vdots \\ p_{ni} \end{pmatrix}$ is an eigen vector of A corresponding to d_i as an eigen value of A.

Thus each column vector of P is an eigen vector of A. Since P is non-singular, the column vectors of P are linearly independent.

It follows that A has n linearly independent eigen vectors.

Conversely, let $X_1, X_2, \ldots, X_n$ be n linearly independent eigen vectors corresponding to the respective eigen values $\lambda_1, \lambda_2, \ldots, \lambda_n$ which may not be all distinct. Let P be the $n \times n$ matrix whose ith column vector is X_i. Since the column vectors are linearly independent, P is non-singular.

Now the ith column vector of AP is AX_i.

Let $D = \text{diag}\ (\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then the ith column vector of PD is $\lambda_i X_i$. As $AX_i = \lambda_i X_i$, the ith column vectors of AP and PD are same for each $i = 1, 2, \ldots, n$.

Therefore $AP = PD$. Since P is non-singular, $P^{-1}AP = D$, showing that A is similar to D and therefore A is diagonalisable.

A sufficient condition for diagonalisability of an $n \times n$ matrix A can be deduced from the above theorem

Theorem 4.18.2. Let A be an $n \times n$ matrix over a field F. If the eigen values of A be all distinct and belong to F, then A is diagonalisable.

Proof. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be n distinct eigen values of A and $\lambda_i \in F$.

Let X_i be an eigen vector corresponding to the eigen value λ_i. Then by the corollary of the Theorem 4.17.4, $X_1, X_2, \ldots, X_n$ are n linearly independent eigen vectors of A.

Thus A has n linearly independent eigen vectors and therefore A is diagonalisable by Theorem 4.18.1.

Note. The condition stated in the Theorem is not a necessary condition for a matrix A to be diagonalisable.

For example, let $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Then the characteristic equation of A is $(x-1)^2(x-5) = 0$. The eigen values of A are 1, 1, 5.

The eigen vectors corresponding to the eigen value 1 are the non-null solutions of the system of equations

$$\begin{aligned} 2x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 + x_3 &= 0. \end{aligned}$$

The system is equivalent to $x_1 + x_2 + \frac{1}{2}x_3 = 0$.

The eigen vectors are $c\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + d\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$, where $(c, d) \neq (0, 0)$.

Two linearly independent eigen vectors corresponding to the eigen value 1 are $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$.

The eigen vectors corresponding to the eigen value 5 are the non-null solutions of the system of equations

$$\begin{aligned} -2x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 - 2x_2 + x_3 &= 0 \\ x_3 &= 0. \end{aligned}$$

The system is equivalent to

$$\begin{aligned} x_1 - x_2 &= 0 \\ x_3 &= 0. \end{aligned}$$

The eigen vectors are $c\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ where $c \neq 0$. Thus A has three distinct eigen vcetors

$\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ which are linearly independent.

By Theorem 4.18.1, A is diagonalisable.

If $P = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -2 & 0 \end{pmatrix}$ then $P^{-1}AP = \text{diag}\ (1, 1, 5)$.

The three eigen values of A are not distinct, yet A is diagonalisable.

We state here, without proof, another necessary and sufficient condition for a matrix to be diagonalisable.

Theorem 4.18.3. An $n \times n$ matrix A is diagonalisable if and only if all its eigen values are regular.

Worked Examples.

1. Show that the matrix $A = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}$ is not diagonalisable.

The eigen values of A are 1, 1. The eigen vectors corresponding to the eigen value 1 are $c\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ where $c (\neq 0) \in \mathbb{R}$. The characteristic subspace corresponding to the eigen value 1 is of dimension 1 and therefore two linearly independent eigen vectors corresponding to the eigen value 1 cannot be found. Therefore a non-singular 2×2 matrix P having two linearly independent eigen vectors as its columns cannot be found and so A cannot be similar to a diagonal matrix.

[Note that the algebraic multiplicity of the eigen value 1 is not equal to its geometric multiplicity.]

2. Find a matrix P such that $P^{-1}AP$ is a diagonal matrix, where

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

The eigen values of A are $1, 2, -1$. The eigen vectors corresponding to the eigen values $1, 2, -1$ are respectively $c_1\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, c_2\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, c_3\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ where c_1, c_2, c_3 are arbitrary non-zero real numbers.

Since the eigen values of A are distinct, there are three linearly independent eigen vectors of A, one corresponding to each eigen value.

The eigen vectors $\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ are linearly independent.

Let $P = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Then P is a non-singular matrix.

$$AP = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1.3 & 2.1 & -1.1 \\ 1.2 & 2.3 & -1.0 \\ 1.1 & 2.1 & -1.1 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 0 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

So $P^{-1}AP = \text{D} = \text{diag}\ (1, 2, -1)$, where $P = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 3 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

3. Diagonalise the matrix $A = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}$.

The eigen values of A are $-2, -2, 4$.

The eigen vectors corresponding to the eigen value -2 are $c\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} +$

$d\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ where c and d are real numbers and $(c,d) \neq (0,0)$.

Two linearly independent eigen vectors corresponding to the eigen value -2 are $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

The eigen vectors corresponding to the eigen value 4 are $c\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$, where c is a non-zero real number.

Three linearly independent eigen vectors are $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

Let $P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$. Then P is a non-singular matrix.

$$AP = \begin{pmatrix} 1 & -3 & 3 \\ 3 & -5 & 3 \\ 6 & -6 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} -2.1 & -2.0 & 4.1 \\ -2.1 & -2.1 & 4.1 \\ -2.0 & -2.1 & 4.2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}\begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

So $P^{-1}AP = D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, a diagonal matrix.

We state without proof a very important property of a real symmetric matrix in the form of the following theorem.

Theorem 4.18.4. If λ be a multiple eigen value of a real $n \times n$ symmetric matrix A, then the algebraic multiplicity of λ is equal to its geometric multiplicity.

The thoerem states that the characteristic subspace of an r-fold eigen value λ of a real symmetric matrix is of dimension r and therefore it is always possible to select r independent eigen vectors (forming a basis of the subspace) corresponding to λ.

Therefore for an $n \times n$ real symmetric matrix, it is always possible to select a full set of n linearly independent eigen vectors.

Let P be the $n \times n$ matrix whose column vectors are these n linearly independent eigen vectors. Then $P^{-1}AP$ is a diagonal matrix and this proves that a real symmetric matrix is diagonalisable.

Worked Example (continued).

4. Diagonalise the symmetric matrix $A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$.

The eigen values of A are 8, 2, 2.

The eigen vectors corresponding to the eigen value 8 are $c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, where c is a non-zero real number.

The eigen vectors corresponding to the eigen value 2 are $c\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + d\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ where $c, d \in \mathbb{R}$ and $(c,d) \neq (0,0)$.

Two linearly independent eigen vectors corresponding to the eigen value 2 are $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

Three linearly independent eigen vectors of A are $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

Let $P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}$. Then P is a non-singular matrix.

$$AP = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix} = \begin{pmatrix} 8.1 & 2.1 & 2.0 \\ 8.1 & 2.0 & 2.1 \\ 8.1 & 2.-1 & 2.-1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

So $P^{-1}AP = D = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, a diagonal matrix.

4.19. Orthogonal diagonalisation of real matrices.

Definition. A square matrix A is said to be *orthogonally diagonalisable* if there exists an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix. The matrix P is said to diagonalise A orthogonally.

Theorem 4.19.1. A square matrix A is orthogonally diagonalisable if and only if A is symmetric.

Proof. Let A be orthogonally diagonalisable. Then there exists an orthogonal matrix P such that $P^{-1}AP$ is diagonal.

Let $P^{-1}AP = D$ where D is a diagonal matrix.

Then $A = PDP^{-1} = PDP^t$ since $P^{-1} = P^t$.

$$A^t = (PDP^t)^t \; = \; PDP^t, \text{ since } D = D^t$$
$$= \; A.$$

This proves that A is symmetric.

Conversely, let A be an $n \times n$ real symmetric matrix. Then by the Theorem 4.18.4, A has a full set of n linearly independent eigen vectors. By Gram-Schmidt process these n eigen vectors can be orthogonalised and then be multiplied by appropriate scalars to give a set of n orthonormal eigen vectors.

Let P be the $n \times n$ matrix whose column vectors are these n orthonormal eigen vectors. Then P is an orthogonal matrix and $P^{-1}AP$ is diagonal. Therefore A is orthogonally diagonalisable.

Note. We have seen previously that a real symmetric matrix A is always diagonalisable. This theorem asserts that there exists always an orthogonal matrix P which can diagonalise A.

Worked Examples.

1. Diagonalise the matrix $A = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}$ orthogonally.

The eigen values of A are $1, -2, 4$.

The eigen vectors corresponding to the eigen values 1, -2, 4 are respectively $c_1 \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}$, $c_2 \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ and $c_3 \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$, where c_1, c_2, c_3 are arbitrary non-zero real numbers. Taking one eigen vector corresponding to each eigen value we get three orthogonal eigen vectors

$\alpha = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \beta = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ and $\gamma = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}$. [Theorem 4.17.10.]

The corresponding orthonormal eigen vectors are $\frac{1}{3}\alpha, \frac{1}{3}\beta$ and $\frac{1}{3}\gamma$.

Let $P = \frac{1}{3}\begin{pmatrix} -2 & 1 & 2 \\ -1 & 2 & -2 \\ 2 & 2 & 1 \end{pmatrix}$. Then P is an orthogonal matrix.

$$AP = \frac{1}{3}\begin{pmatrix} 2 & -2 & 0 \\ -2 & 1 & -2 \\ 0 & -2 & 0 \end{pmatrix}\begin{pmatrix} -2 & 1 & 2 \\ -1 & 2 & -2 \\ 2 & 2 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2 & -2 & 8 \\ -1 & -4 & -8 \\ 2 & -4 & 4 \end{pmatrix}$$

$$= \frac{1}{3}\begin{pmatrix} -2 & 1 & 2 \\ -1 & 2 & -2 \\ 2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

So $P^{-1}AP = D = \text{diag}\ (1, -2, -4)$, where $P = \frac{1}{3}\begin{pmatrix} -2 & 1 & 2 \\ -1 & 2 & -2 \\ 2 & 2 & 1 \end{pmatrix}$.

2. Find an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix where $A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$.

The eigen values of A are 1, 1, 10.

The eigen vectors corresponding to the eigen value 1 are $c\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + d\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ where $c, d \in \mathbb{R}$ and $(c, d) \neq (0, 0)$.

Let $\alpha = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ and let us choose c, d such that $\beta = c\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + d\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$ is orthogonal to α. Then $2.(2c + 2d) + c = 0$, or, $5c + 4d = 0$.

Taking $c = 4$ and $d = -5$, we have $\beta = \begin{pmatrix} -2 \\ 5 \\ 4 \end{pmatrix}$.

The eigen vectors corresponding to the eigen value 10 are

$c\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$ where $c \in \mathbb{R}$ and $c \neq 0$.

Thus we obtain three mutually orthogonal eigen vectors of A as

$\alpha = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \beta = \begin{pmatrix} -2 \\ 5 \\ 4 \end{pmatrix}, \gamma = \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$. Normalising them we obtain orthonormal eigen vectors $\frac{1}{\sqrt{5}}\alpha, \frac{1}{3\sqrt{5}}\beta, \frac{1}{3}\gamma$.

Let $P = \frac{1}{3\sqrt{5}}\begin{pmatrix} 6 & -2 & \sqrt{5} \\ 0 & 5 & 2\sqrt{5} \\ 3 & 4 & -2\sqrt{5} \end{pmatrix}$.

Then P is an orthogonal matrix and $AP =$

$$\begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix} \frac{1}{3\sqrt{5}} \begin{pmatrix} 6 & -2 & \sqrt{5} \\ 0 & 5 & 2\sqrt{5} \\ 3 & 4 & -2\sqrt{5} \end{pmatrix} = \frac{1}{3\sqrt{5}} \begin{pmatrix} 6 & -2 & 10\sqrt{5} \\ 0 & 5 & 20\sqrt{5} \\ 3 & 4 & -20\sqrt{5} \end{pmatrix}$$

$$= \frac{1}{3\sqrt{5}} \begin{pmatrix} 6 & -2 & \sqrt{5} \\ 0 & 5 & 2\sqrt{5} \\ 3 & 4 & -2\sqrt{5} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}.$$

So $P^{-1}AP = \text{diag}\,(1, 1, 10)$, where $P = \frac{1}{3\sqrt{5}} \begin{pmatrix} 6 & -2 & \sqrt{5} \\ 0 & 5 & 2\sqrt{5} \\ 3 & 4 & -2\sqrt{5} \end{pmatrix}$.

Exercises 14

1. Diagonalise the following matrices.

(i) $\begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$, (ii) $\begin{pmatrix} 3 & 2 & 2 \\ 1 & 4 & 1 \\ -2 & -4 & -1 \end{pmatrix}$, (iii) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

2. Show that the following matrices are not diagonalisable.

(i) $\begin{pmatrix} 2 & 8 \\ 0 & 2 \end{pmatrix}$, (ii) $\begin{pmatrix} 3 & 1 & 1 \\ 4 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix}$, (iii) $\begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 2 & 1 & 2 \end{pmatrix}$.

3. Prove that any $n \times n$ matrix similar to a nilpotent matrix of index p is nilpotent of index p. Deduce that

(i) a nilpotent matrix is not diagonalisable;

(ii) a srictly triangular matrix is not diagonalisable.

[**Hint.** (ii) A strictly triangular matrix is nilpotent.]

4. If $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ find a matrix P such that $P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

5. Find a matrix P such that $P^{-1}AP$ is a diagonal matrix, where

(i) $A = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$.

Use the result to find A^2 and A^{-1} in each case.

6. Find an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix, where

$A = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 3 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$.

4.20. Real quadratic form.

An expression of the form $\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}x_ix_j$, where a_{ij} are real and $a_{ij} = a_{ji}$, is said to be a *real quadratic form* in n variables $x_1, x_2, \ldots, x_n$.

The matrix notation for the quadratic form is X^tAX where $X = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$ and $A = (a_{ij})_{n,n}$. A is a real symmetric matrix since $a_{ij} \in \mathbb{R}$ and $a_{ij} = a_{ji}$.

To every real quadratic form in n variables is associated an $n \times n$ *real symmetric matrix* which is said to be the *matrix of the quadratic form*.

Examples.

1. $5x_1^2 + 2x_1x_2 - x_2^2$ is a real quadratic form in two variables x_1, x_2.

The associated matrix is $\begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix}$.

2. $x_1x_2 - x_2x_3$ is a real quadratic form in three variables x_1, x_2, x_3.

The associated matrix is $\begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}$.

3. $x_1^2 - x_2^2 + 2x_3^2$ is a real quadratic form in three variables.

The associated matrix is the diagonal matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Definitions.

A real quadratic form $Q(x_1, x_2, \ldots, x_n)$ assumes the value 0 when $X = O$. But Q takes up different real values for different non-zero X.

A real quadratic form $Q = X^tAX$ is said to be

(i) *positive definite* if $Q > 0$ for all $X \neq O$;

(ii) *positive semi definite* if $Q \geq 0$ for all X and $Q = 0$ for some $X \neq O$;

(iii) *negative definite* if $Q < 0$ for all $X \neq O$;

(iv) *negative semi definite* if $Q \leq 0$ for all X and $Q = 0$ for some $X \neq O$;

(v) *indefinite* if $Q \geq 0$ for some $X \neq O$ and $Q \leq 0$ for some other $X \neq O$.

In corresponding cases, the associated real symmetric matrix A is said to be positive definite, positive semi-definite, negative definite, negative semi-definite and indefinite respectively.

In simple cases, positive definiteness etc. of a real quadratic form can be established by direct calculation of the form.

Examples (continued).

4. Let $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 4x_3^2 + 2x_1x_2 - 4x_2x_3 - 2x_3x_1$.

$Q = (x_1 + x_2 - x_3)^2 + (x_2 - x_3)^2 + 2x_3^2$.

$Q \geq 0$ for all (x_1, x_2, x_3) and $Q = 0$ only when $x_1 = x_2 = x_3 = 0$.

Therefore Q is positive definite. The associated matrix $\begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & -2 \\ -1 & -2 & 4 \end{pmatrix}$ is therefore positive definite.

5. Let $Q(x_1, x_2, x_3) = x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 - 4x_2x_3 - 2x_3x_1$.

$Q = (x_1 + x_2 - x_3)^2 + (x_2 - x_3)^2$.

$Q \geq 0$ for all (x_1, x_2, x_3). But Q may be zero for non-zero (x_1, x_2, x_3). For example, $Q(0, 1, 1) = 0$.

Therefore Q is positive semi-definite.

6. Let $Q(x_1, x_2, x_3) = x_1^2 + x_2^2 + 2x_1x_2 + 2x_2x_3 + 2x_3x_1$.

$Q = (x_1^2 + x_2 + x_3)^2 - x_3^2$.

Q is indefinite because $Q > 0$ for some (x_1, x_2, x_3), and $Q < 0$ for some (x_1, x_2, x_3). For example, $Q(1, 1, 0) > 0, Q(0, -1, 1) < 0$.

By the transformation $X = PY$, where P is a non-singular matrix the real quadratic form $Q = X^tAX$ transforms to $Y^t(P^tAP)Y$. This is a quadratic form in Y since P^tAP is a symmetric matrix.

Thus Q is transformed to a quadratic form Q'. The associated matrix of Q' being congruent to A, Q' is said to be congruent to Q.

For a real symmetric matrix A of rank $r (\leq n)$, there exists a non-singular matrix P such that P^tAP becomes a diagonal matrix

$$D = \begin{pmatrix} I_m & & \\ & -I_{r-m} & \\ & & O \end{pmatrix}$$ of rank r, where $0 \leq m \leq r$.

Therefore by a suitable transformation $X = PY$, where P is non-singular, the real quadratic form Q transforms to

$y_1^2 + y_2^2 + \cdots + y_m^2 - y_{m+1}^2 - y_{m+2}^2 - \cdots - y_r^2$, where $0 \leq m \leq r \leq n$.

This is said to be the *normal (diagonal)* form of Q.

Theorem 4.20.1. The integer m which is the number of positive elements in the normal (diagonal) form of a real quadratic form Q, is invariant.

Proof. Assuming otherwise, let Q be transformed to $y_1^2 + y_2^2 + \cdots + y_m^2 - y_{m+1}^2 - y_{m+2}^2 - \cdots - y_r^2$ by the transformation $X = UY$ and to $z_1^2 + z_2^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$ by the transformation $X = VZ$, where U and V are non-singular and $m \neq k$.

Without loss of generality, let us assume $m < k$.

$Y = U^{-1}X, Z = V^{-1}X$.

Let $U^{-1} = (u_{ij})$, $V^{-1} = (v_{ij})$ and let us consider $m + n - k$ equations in n unknowns

$$\begin{aligned} u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n &= 0 \\ \cdots \quad & \cdots \\ u_{m1}x_1 + u_{m2}x_2 + \cdots + u_{mn}x_n &= 0 \\ v_{k+11}x_1 + v_{k+12}x_2 + \cdots + v_{k+1n}x_n &= 0 \\ \cdots \quad & \cdots \\ v_{n1}x_1 + v_{n2}x_2 + \cdots + v_{nn}x_n &= 0. \end{aligned}$$

Since $m < k$, there exists a non-zero solution say $(x_1', x_2', \ldots, x_n')$, of the homogeneous system.

$$\begin{aligned} \text{When } X &= (x_1', x_2', \ldots, x_n') \text{ let } Y = (y_1', y_2', \ldots, y_n') \text{ and} \\ Z &= (z_1', z_2', \ldots, z_n'). \end{aligned}$$

The first m equations of the system give $y_1' = y_2' = \cdots = y_m' = 0$, and the last $n - k$ equations give $z_{k+1}' = z_{k+2}' = \cdots = z_n' = 0$.

$$\begin{aligned} \text{Then } Q(x_1', x_2', \ldots, x_n') &= -{y_{m+1}'}^2 - {y_{m+2}'}^2 - \cdots - {y_r'}^2 \\ &= {z_1'}^2 + {z_2'}^2 + \cdots + {z_k'}^2. \end{aligned}$$

This equality can hold if and only if $y_{m+1}' = y_{m+2}' = \cdots = y_r' = 0$ and $z_1' = z_2' = \cdots = z_k' = 0$.

Therefore $Y = O$, $Z = O$.

Since V and U are non-singular matrices, $(x_1', x_2', \ldots, x_n') = (0, 0, \ldots, 0)$ and this is a contradiction.

Therefore our assumption that $m \neq k$ is wrong. Thus m is invariant of the quadratic form Q and the theorem is done.

Corollary. If a real $n \times n$ symmetric matrix A of rank r be congruent to the diagonal matrix $D = \begin{pmatrix} I_m & & \\ & -I_{r-m} & \\ & & O \end{pmatrix}$, m is invariant.

Since r remains invariant under congruence, the signature of the matrix $= m - (r - m) = 2m - r$ is also invariant.

Definitions.

The *rank of a real quadratic form* is defined to be the rank of the associated real symmetric matrix.

The *signature of a real quadratic form* is defined to be the signature of the associated real symmetric matrix.

Theorem 4.20.2. By the transformation $X = PY$, where P is non-singular, the character of a real quadratic form X^tAX regarding positive definiteness etc. remains invariant.

Proof. The quadratic form $Q = X^tAX$ is transformed by the substitution $X = PY$ to the form $Q' = Y^t(P^tAP)Y$.

Since P is non-singular, $Y = P^{-1}X$.
$X \neq O \Rightarrow Y \neq O$ and $X = O \Rightarrow Y = O$.

(i) Let Q be positive definite.
Then $Q > 0$ for all $X \neq O$ and therefore $Q > 0$ for all $PY \neq O$.

This implies $Q' > 0$ for all $Y \neq O$, proving that Q' is positive definite.

(ii) Let Q be positive semi-definite.
Then $Q \geq 0$ for all X and $Q = 0$ for some $X \neq O$.
Therefore $Q \geq 0$ for all PY and $Q = 0$ for some $PY \neq O$.

This implies $Q' \geq 0$ for all Y and $Q' = 0$ for some $Y \neq O$, proving that Q' is positive semi-definite.

The proof of the remaining cases left to the reader.

Theorem 4.20.3. A real quadratic form $Q(x_1, x_2, \ldots, x_n)$ of rank r and index m is

(i) positive definite, if $r = n, m = r$;

(ii) positive semi-definite, if $r < n, m = r$;

(iii) negative definite, if $r = n, m = 0$;

(iv) negative semi-definite, if $r < n, m = 0$;

(v) indefinite, if $r \leq n, 0 < m < r$.

Proof. By a suitable non-singular transformation $X = PY$, Q is transformed to the normal form $D(y_1, y_2, \ldots, y_n)$, where $D = y_1^2 + y_2^2 + \cdots + y_m^2 - y_{m+1}^2 - \cdots - y_r^2$ and $0 \leq m \leq r \leq n$.

(i) If $r = n$, $m = r$ then $D = y_1^2 + y_2^2 + \cdots + y_n^2$.
Hence D is positive definite and therefore Q is so.

(ii) If $r < n$, $m = r$ then $D = y_1^2 + y_2^2 + \cdots + y_r^2, r < n$.
Hence D is positive semi-definite and therefore Q is so.

(iii) If $r = n$, $m = 0$ then $D = -y_1^2 - y_2^2 - \cdots - y_n^2$.
Hence D is negative definite and therefore Q is so.

(iv) If $r < n$, $m = 0$ then $D = -y_1^2 - y_2^2 - \cdots - y_r^2, r < n$.
Hence D is negative semi-definite and therefore Q is so.

(v) If $r \leq n$, $0 < m < r$ then $D = y_1^2 + y_2^2 + \cdots + y_m^2 - y_{m+1}^2 - \cdots - y_r^2, r \leq n$.
Hence D is indefinite and therefore Q is so.

Corollary. Q is positive definite if rank of Q = signature of $Q = n$.

Theorem 4.20.4. A real symmetric matrix is positive definite if and only if all its eigen values are positive.

Proof. Let A be a real symmetric matrix of order n. Then all its eigen values are real. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigen values of A.

Since A is a real symmetric matrix, there exists an orthogonal matrix P such that $P^{-1}AP(= P^tAP)$ is a diagonl matrix. But A and $P^{-1}AP$ have the same eigen values.

$$P^{-1}AP = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

Let A be positive definite. Then A is congruent to I_n. Also P being non-singular, P^tAP is congruent to A.

Consequently, P^tAP is congruent to I_n. Therefore $\lambda_i > 0$ for $i = 1, 2, \ldots, n$.

Conversely, let $\lambda_i > 0$ for all i. Then diag $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ is positive definite. But A is congruent to $P^tAP = \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Therefore A is positive definite.

This completes the proof.

Corollary. If A be a positive definite real symmetric matrix then A is non-singular and det $A > 0$.

This follows from the fact that det A = det $P^{-1}AP = \lambda_1\lambda_2 \ldots \lambda_n$.

Theorem 4.20.5. A real symmetric matrix is negative definite if and only if all its eigen values are negative.

Similar proof.

Worked Examples.

1. Reduce the quadratic form $5x^2+y^2+10z^2-4yz-10zx$ to the normal form and show that it is positive definite.

The associated symmetric matrix is $A=\begin{pmatrix}5 & 0 & -5\\ 0 & 1 & -2\\ -5 & -2 & 10\end{pmatrix}$.

Let us apply congruence operations on A to reduce it to the normal form

$$A \xrightarrow{R_3+R_1} \begin{pmatrix}5 & 0 & -5\\ 0 & 1 & -2\\ 0 & -2 & 5\end{pmatrix} \xrightarrow{C_3+C_1} \begin{pmatrix}5 & 0 & 0\\ 0 & 1 & -2\\ 0 & -2 & 5\end{pmatrix}$$

$$\xrightarrow{R_3+2R_2} \begin{pmatrix}5 & 0 & 0\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix} \xrightarrow{C_3+2C_2} \begin{pmatrix}5 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

$$\xrightarrow{\frac{1}{\sqrt{5}}R_1} \begin{pmatrix}\sqrt{5} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \xrightarrow{\frac{1}{\sqrt{5}}C_1} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

The normal form is $x^2+y^2+z^2$.

The rank of the quadratic form is 3 and its signature is 3.

The quadratic form is positive definite.

2. Show that the quadratic form $x^2+2y^2+3z^2-2xy+4yz$ is indefinite.

The associated symmetric matrix is $A=\begin{pmatrix}1 & -1 & 0\\ -1 & 2 & 2\\ 0 & 2 & 3\end{pmatrix}$.

Let us apply congruence operations on A.

$$A \xrightarrow{R_2+R_1} \begin{pmatrix}1 & -1 & 0\\ 0 & 1 & 2\\ 0 & 2 & 3\end{pmatrix} \xrightarrow{C_2+C_1} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 2\\ 0 & 2 & 3\end{pmatrix}$$

$$\xrightarrow{R_3-2R_2} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 2\\ 0 & 0 & -1\end{pmatrix} \xrightarrow{C_3-2C_2} \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix} = D.$$

The normal form is $x^2+y^2-z^2$. The form is indefinite.

3. Obtain a non-singular transformation that will reduce the quadratic form $x^2+2y^2+3z^2-2xy+4yz$ to the normal form.

The form is X^tAX where $A=\begin{pmatrix}1 & -1 & 0\\ -1 & 2 & 2\\ 0 & 2 & 3\end{pmatrix}$, $X=\begin{pmatrix}x\\ y\\ z\end{pmatrix}$.

Let $X=PX'$, where $X'=\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}$ and P is non-singular, transforms the form into the normal form X'^tDX'. Then $D\ (=P^tAP)$ is a diagonal matrix.

By the previous example,
$E_{32}(-2)E_{21}(1)A\{E_{21}(1)\}^t\{E_{32}(-2)\}^t=D$.

Therefore $P^t = E_{32}(-2)E_{21}(1)$

$$= \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -2 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ -2 & -2 & 1\end{pmatrix}.$$

$P=\begin{pmatrix}1 & 1 & -2\\ 0 & 1 & -2\\ 0 & 0 & 1\end{pmatrix}$ and the transformation $X=PX'$ becomes

$$\begin{aligned} x &= x'+y'-2z'\\ y &= y'-2z'\\ z &= z'. \end{aligned}$$

4. B is a real $n\times n$ matrix. Show that the symmetric matrix B^tB is either positive definite or positive semi-definite and it is positive definite or positive semi-definite according as B is non-singular or singular.

The real quadratic form corresponding to the matrix B^tB is $X^t(B^tB)X$, where X is an $n\times 1$ real matrix.

$X^t(B^tB)X=(BX)^t(BX)=Y^tY$, where $Y=BX$.

Let $X=\begin{pmatrix}x_1\\ x_2\\ \cdot\\ \cdot\\ x_n\end{pmatrix}$ and $Y=\begin{pmatrix}y_1\\ y_2\\ \cdot\\ \cdot\\ y_n\end{pmatrix}$. Then $Y^tY=y_1^2+y_2^2+\cdots+y_n^2$.

The quadratic form Y^tY assumes real values greater than or equal to 0. So the quadratic form $X^t(B^tB)X$ is either positive definite or positive semi-definite and therefore the matrix B^tB is either positive definite or positive semi-definite.

Since Y is a real $n\times 1$ matrix, $Y^tY=0$ occurs only when $Y=O$, i.e., when $BX=O$.

If B be non-singular, $BX = 0$ occurs only when $X = O$.

If B be singular, $BX = O$ holds for $X = O$ and also for some $X \neq O$.

Therefore the quadratic form $X^t(B^tB)X$ is positive definite if B be non-singular and positive semi-definite if B be singular. Hence the matrix B^tB is positive definite or positive semi-definite according as B is non-singular or singular.

5. Show that $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$ is positive definite and express A as the matrix P^tP for some non-singular matrix P.

Let us apply congruence operations on A to reduce it to the normal form.

$$A \xrightarrow[R_3-R_1]{R_2-R_1} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \xrightarrow[C_3-C_1]{C_2-C_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

$$\xrightarrow{R_3-R_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_3-C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The normal form under congruence is $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

A is positive definite.

$E_{32}(-1)E_{31}(-1)E_{21}(-1)A\{E_{21}(-1)\}^t\{E_{31}(-1)\}^t\{E_{32}(-1)\}^t = I_3$.

Let $B = \{E_{32}(-1)E_{31}(-1)E_{21}(-1)\}^t$. Then $B^tAB = I_3$.

B is the product of elementary matrices and therefore it is non-singular.

$B^tAB = I_3 \Rightarrow A = \{B^t\}^{-1}I_3(B)^{-1} = \{B^{-1}\}^tB^{-1} = P^tP$, where $P = B^{-1}$ and P is non-singular.

$$P = \{B^t\}^{-1} = [\{E_{21}(-1)\}^t\{E_{31}(-1)\}^t\{E_{32}(-1)\}^t]^{-1}$$
$$= \{E_{12}(-1)E_{13}(-1)E_{23}(-1)\}^{-1}$$
$$= E_{23}(1)E_{13}(1)E_{12}(1), \text{ since } \{E_{ij}(c)\}^{-1} = E_{ij}(-c)$$
$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \text{ Therefore } A = P^tP \text{ where } P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercises 15

1. A real symmetric matrix A is positive definite. Prove that
 (i) $\det A > 0$,
 (ii) A^{-1} is positive definite,
 (iii) $adj A$ is positive definite.

[Hint. (iii) If λ be an eigen value of A then $\det(adj A - \frac{\det A}{\lambda} I_n) = 0$.]

2. If A be a real non-singular matrix, show that the eigen values of the matrix A^tA are all real and positive.

[Hint. $X^t(A^tA)X = (AX)^t(AX) > 0$ for all $X \neq O$.]

3. Reduce the following quadratic forms to their normal forms. Find the rank and signature of each.
 (i) $2x^2 + 5y^2 + 10z^2 + 4xy + 12yz + 6zx$,
 (ii) $x^2 + 2y^2 + 6xz + 4yz$,
 (iii) $2x^2 + 3y^2 + 4z^2 - 4xy + 4yz$,
 (iv) $xy + yz + zx$.

4. Obtain a non-singular transformation that will reduce the quadratic form into the normal form.
 (i) $x_1^2 + 2x_1x_2 + 2x_2x_3$,
 (ii) $x^2 + y^2 + z^2 - 2xy - 2yz - 2zx$.

5. Show that the following quadratic forms are positive definite.
 (i) $x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 + 4x_2x_3 + 2x_1x_3$,
 (ii) $x_1^2 + 2x_2^2 + 6x_3^2 - 2x_1x_3 + 4x_2x_3$,
 (iii) $5x^2 + y^2 + 5z^2 + 4xy - 8xz - 4yz$.

6. Show that the following forms are positive semi-definite.
 (i) $x^2 + 2y^2 + 2z^2 + 2xy + 2xz$,
 (ii) $x^2 + 5y^2 + 2z^2 - 4xy - 6yz + 2zx$,
 (iii) $2x^2 + 2y^2 + 5z^2 - 4xy - 2xz + 2yz$.

7. Show that the following quadratic forms are indefinite.

(i) $x^2 + 2y^2 + z^2 + 4xy$,

(ii) $x^2 + y^2 + 2xy + 2yz$,

(iii) $x_1^2 + x_2^2 + x_3^2 + 3x_2x_3$.

8. Find λ for which the quadratic form $f(x, y, z)$ is positive definite.

(i) $f(x,y,z) = x^2 + \lambda(y^2 + z^2) + 2xy$,

(ii) $f(x,y,z) = x^2 + \lambda(y^2 + z^2) + 2yz$.

9. Show that the real quadratic form $ax^2+2hxy+by^2$ $(a \neq 0)$ in two variables x, y is positive definite if and only if $a > 0$ and $\begin{vmatrix} a & h \\ h & b \end{vmatrix} > 0$.

10. Without finding the characteristic equation, prove that every eigen value of the matrix $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$ is a positive real number.

11. Examine if the matrix A is positive definite, where

(i) $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$,

(iii) $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, (iv) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

12. Prove that a real symmetric matrix A is positive definite if and only if $A = BB^t$ for some non-singular matrix B. Show that A is positive definite and find the matrix B such that $A = BB^t$, where

(i) $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 11 \end{pmatrix}$, (ii) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$.

[**Hint.** If A is positive definite then $P^tAP = I$ for some non-singular matrix P and so $A = (P^{-1})^tP^{-1}$. Conversely, $A = BB^t \Rightarrow B^{-1}A(B^t)^{-1} = I \Rightarrow P^tAP = I$, where $P = (B^t)^{-1}$.]

13. A and B are $n \times n$ real symmetric matrices and A is positive definite. Prove that there exists a non-singular matrix P such that $PAP^t = I_n$ and $PBP^t = D$, where D is a diagonal matrix.

[**Hint.** For some non-singular matrix Q, $QAQ^t = I_n$. Since B is symmetric, QBQ^t is symmetric and there exists an orthogonal matrix S such that $S(QBQ^t)S^{-1}(= S(QBQ^t)S^t) = D$. Take $P = SQ$.]

4.21. Linear mappings.

Let V and W be vector spaces over the same field F. A mapping $T : V \to W$ is said to be a *linear mapping* (or a *linear transformation*) if it satisfies the following conditions –

1. $T(\alpha + \beta) = T(\alpha) + T(\beta)$ for all α, β in V
2. $T(c\alpha) = cT(\alpha)$ for all c in F and all α in V.

These two conditions can be replaced by the single condition–

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

Note 1. A linear mapping $T : V \to W$ is also a *homomorphism* of V to W.

2. Generally, a linear mapping T is a transformation from one vector space V to another vector space W, both over the same field of scalars. But the co-domain space W may be the space V itself. In this case T is said to be a *linear mapping on* V.

There is another important case when the co-domain space is F, regarded as a vector space over itself. In this case $T : V \to F$ is said to be a *linear functional*.

Examples.

1. **The identity mapping.** The mapping $T : V \to V$ defined by $T(\alpha) = \alpha$ for all α in V, is a linear mapping. This is called the *identity mapping* on V and is denoted by I_V.

2. **The zero mapping.** The mapping $T : V \to W$ defined by $T(\alpha) = \theta'$ for each α in V, θ' being the null vector in W, is a linear mapping. This is called the *zero mapping* and is denoted by 0.

3. The mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x_1, x_2, x_3) = (x_1, x_2, 0)$, $(x_1, x_2, x_3) \in \mathbb{R}^3$ is linear.

Let $\alpha = (x_1, x_2, x_3)$, $\beta = (y_1, y_2, y_3) \in \mathbb{R}^3$.

Then $\alpha + \beta = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$.

$T(\alpha + \beta) = (x_1 + y_1, x_2 + y_2, 0) = (x_1, x_2, 0) + (y_1, y_2, 0) = T(\alpha) + T(\beta)$.

For $c \in \mathbb{R}$, $c\alpha = (cx_1, cx_2, cx_3)$. $T(c\alpha) = (cx_1, cx_2, 0) = c(x_1, x_2, 0) = cT(\alpha)$.

Therefore T is a linear mapping.

4. Let P be the vector space of all real polynomials. The mapping $D : P \to P$ defined by $Dp(x) = \frac{d}{dx}p(x), p(x) \in P$ is a linear mapping.

5. Let V be the vector space of all real valued functions continuous on the closed interval $[a, b]$ and let $T : V \to \mathbb{R}$ be defined by $T(f) = \int_a^b f(t)dt, f \in V$. Then T is a linear functional.

6. The mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x_1, x_2, x_3) = (x_1 + 1, x_2 + 1, x_3 + 1), (x_1, x_2, x_3) \in \mathbb{R}^3$ is not a linear mapping.

Let $\alpha = (x_1, x_2, x_3)$, $\beta = (y_1, y_2, y_3) \in \mathbb{R}^3$.
Then $\alpha + \beta = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$.

$$\begin{aligned} T(\alpha + \beta) &= (x_1 + y_1 + 1, x_2 + y_2 + 1, x_3 + y_3 + 1) \\ &= (x_1, x_2, x_3) + (y_1, y_2, y_3) + (1, 1, 1) \\ &\neq T(\alpha) + T(\beta). \end{aligned}$$

This shows that T is not a linear mapping.

Theorem 4.21.1. Let V and W be vector spaces over a field F and $T : V \to W$ be a linear mapping. Then

(i) $T(\theta) = \theta'$, where θ, θ' are null elements in V and W respectively;

(ii) $T(-\alpha) = -T(\alpha)$ for all $\alpha \in V$.

Proof. (i) In $V, \theta + \theta = \theta$.
Since T is linear, $T(\theta) + T(\theta) = T(\theta)$ in W.

This implies $-T(\theta) + [T(\theta) + T(\theta)] = -T(\theta) + T(\theta)$
$\Rightarrow [-T(\theta) + T(\theta)] + T(\theta) = \theta'$, since θ' is the null vector in W
$\Rightarrow \theta' + T(\theta) = \theta'$
$\Rightarrow T(\theta) = \theta'$.

(ii) Proof left as an exercise.

Kernel of a linear mapping.

Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping. The set of all vectors $\alpha \in V$ such that $T(\alpha) = \theta', \theta'$ being the null vector in W, is said to be the *kernel* of T and is denoted by $Ker\, T$.

$$Ker\, T = \{\alpha \in V : T(\alpha) = \theta'\}.$$

Theorem 4.21.2. Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping. Then $Ker\, T$ is a subspace of V.

Proof. $Ker\, T = \{\alpha \in V : T(\alpha) = \theta'\}$.
Since $T(\theta) = \theta', \theta \in Ker\, T$. Therefore $Ker\, T$ is non-empty.

Case 1. $Ker\, T = \{\theta\}$. Then $Ker\, T$ is a subspace of V.

Case 2. $Ker\, T \neq \{\theta\}$. Let $\alpha \in Ker\, T$. Then $T(\alpha) = \theta'$.

Let $c \in F$. Then $T(c\alpha) = cT(\alpha)$, since T is linear
$= c\theta' = \theta'$.

This implies $c\alpha \in Ker\, T$.

Let $\alpha, \beta \in Ker\, T$. Then $T(\alpha) = \theta'$, $T(\beta) = \theta'$.
$T(\alpha + \beta) = T(\alpha) + T(\beta)$, since T is linear
$= \theta' + \theta' = \theta'$.

This implies $\alpha + \beta \in Ker\, T$.

Thus $\alpha, \beta \in Ker\, T \Rightarrow \alpha + \beta \in Ker\, T$ and $\alpha \in Ker\, T \Rightarrow c\alpha \in Ker\, T$ for all $c \in F$. This proves that $Ker\, T$ is a subspace of V.
This completes the proof.

Note. $Ker\, T$ is also called the *null space* of T and is denoted by $N(T)$.

Theorem 4.21.3. Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping. Then T is injective if and only if $Ker\, T = \{\theta\}$.

Proof. Let T be injective. Since $T(\theta) = \theta'$ in W, θ is a pre-image of θ' and since T is injective, θ is the only pre-image of θ'. So $Ker\, T = \{\theta\}$.

Conversely, let $Ker\, T = \{\theta\}$ and α, β be two elements of V such that $T(\alpha) = T(\beta)$ in W.

$\theta' = T(\alpha) - T(\beta)$
$= T(\alpha - \beta)$, since T is linear.

This implies $\alpha - \beta \in Ker\, T$ and since $Ker\, T = \{\theta\}, \alpha = \beta$.
Thus $T(\alpha) = T(\beta) \Rightarrow \alpha = \beta$ and therefore T is injective.
This completes the proof.

Theorem 4.21.4. Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping such that $Ker\, T = \{\theta\}$. Then the images of a linearly independent set of vectors $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ in V are linearly independent in W.

Proof. To prove linear independence of the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_r)$ in W, let us consider the relation

$c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_rT(\alpha_r) = \theta'$, where $c_1, c_2, \ldots, c_r \in F$.

This implies $T(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r) = \theta'$, since T is linear.
$\Rightarrow c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r = \theta$, since $Ker\, T = \{\theta\}$
$\Rightarrow c_1 = c_2 = \cdots = c_r = 0$, since $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ is linearly independent. This proves linear independence of the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_r)$ in W.

Corollary. If $T : V \to V$ be a linear mapping on V such that $Ker\ T = \{\theta\}$, then a basis of V is mapped onto another basis of V.

Worked Examples.

1. A mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x_1, x_2, x_3) = (x_1 + x_2 + x_3, 2x_1 + x_2 + 2x_3, x_1 + 2x_2 + x_3)$, $(x_1, x_2, x_3) \in \mathbb{R}^3$. Show that T is a linear mapping. Find $Ker\ T$ and the dimension of $Ker\ T$.

Let $\alpha = (a_1, a_2, a_3) \in \mathbb{R}^3$, $\beta = (b_1, b_2, b_3) \in \mathbb{R}^3$.

Then $T(\alpha) = (a_1 + a_2 + a_3, 2a_1 + a_2 + 2a_3, a_1 + 2a_2 + a_3)$,
$T(\beta) = (b_1 + b_2 + b_3, 2b_1 + b_2 + 2b_3, b_1 + 2b_1 + b_3)$.
$\alpha + \beta = (a_1 + b_1, a_2 + b_2, a_3 + b_3)$.

$$\begin{aligned} T(\alpha+\beta) &= ((a_1+b_1)+(a_2+b_2)+(a_3+b_3), 2(a_1+b_1)+(a_2+b_2)\\ &\quad +2(a_3+b_3), (a_1+b_1)+2(a_2+b_2)+(a_3+b_3))\\ &= ((a_1+a_2+a_3)+(b_1+b_2+b_3), (2a_1+a_2+2a_3)\\ &\quad +(2b_1+b_2+2b_3), (a_1+2a_2+a_3)+(b_1+2b_2+b_3))\\ &= T(\alpha)+T(\beta). \end{aligned}$$

Let $c \in \mathbb{R}$. Then $c\alpha = (ca_1, ca_2, ca_3)$.

$$\begin{aligned} T(c\alpha) &= (ca_1 + ca_2 + ca_3, 2ca_1 + ca_2 + 2ca_3, ca_1 + 2ca_2 + ca_3)\\ &= c(a_1 + a_2 + a_3, 2a_1 + a_2 + 2a_3, a_1 + 2a_2 + a_3)\\ &= cT(\alpha). \end{aligned}$$

Thus $T(\alpha + \beta) = T(\alpha) + T(\beta)$ for all $\alpha, \beta \in \mathbb{R}^3$
and $T(c\alpha) = cT(\alpha)$ for all $c \in \mathbb{R}$ and $\alpha \in \mathbb{R}^3$.

Hence T is a linear mapping.

$Ker\ T = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : T(x_1, x_2, x_3) = (0, 0, 0)\}$.
Let $(x_1, x_2, x_3) \in Ker\ T$.

Then $x_1 + x_2 + x_3 = 0, 2x_1 + x_2 + 2x_3 = 0, x_1 + 2x_2 + x_3 = 0$.

From the first two equations we have $\frac{x_1}{1} = \frac{x_2}{0} = \frac{x_3}{-1} = k$, say.
Then $x_1 = k, x_2 = 0, x_3 = -k$. The last equation is satisfied.
Therefore $(x_1, x_2, x_3) = k(1, 0, -1)$, $k \in \mathbb{R}$.

Let $\alpha = (1, 0, -1)$. Then $Ker\ T = L\{\alpha\}$ and $\dim Ker\ T = 1$.

2. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^4$ is defined by $T(x_1, x_2, x_3) = (x_2 + x_3, x_3 + x_1, x_1 + x_2, x_1 + x_2 + x_3)$, $(x_1, x_2, x_3) \in \mathbb{R}^3$. Find $Ker\ T$. Verify that the set $\{T(\epsilon_1), T(\epsilon_2), T(\epsilon_3)\}$ is linearly independent in $\mathbb{R}^4$, where $\epsilon_1 = (1, 0, 0)$, $\epsilon_2 = (0, 1, 0)$, $\epsilon_3 = (0, 0, 1)$.

$Ker\ T = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : T(x_1, x_2, x_3) = (0, 0, 0, 0)\}$.
Let $(x_1, x_2, x_3) \in ker\ T$.

Then $x_2 + x_3 = 0, x_3 + x_1 = 0, x_1 + x_2 = 0, x_1 + x_2 + x_3 = 0$.
The solution is $x_1 = x_2 = x_3 = 0$. Therefore $Ker\ T = \{\theta\}$.
$T(\epsilon_1) = (0, 1, 1, 1)$, $T(\epsilon_2) = (1, 0, 1, 1)$, $T(\epsilon_3) = (1, 1, 0, 1)$.

To examine linear independence of the set $\{T(\epsilon_1), T(\epsilon_2), T(\epsilon_3)\}$, let $c_1T(\epsilon_1) + c_2T(\epsilon_2) + c_3T(\epsilon_3) = \theta$, $c_i \in \mathbb{R}$.
Then $c_1(0, 1, 1, 1) + c_2(1, 0, 1, 1) + c_3(1, 1, 0, 1) = (0, 0, 0, 0)$.
Therefore $c_2 + c_3 = 0, c_1 + c_3 = 0, c_1 + c_2 = 0, c_1 + c_2 + c_3 = 0$.
The solution is $c_1 = c_2 = c_3 = 0$.

This proves that $\{T(\epsilon_1), T(\epsilon_2), T(\epsilon_3)\}$ is linearly independent in $\mathbb{R}^4$.

Image of a linear mapping.

Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping. The images of the elements of V under the mapping T form a subset of W. This subset is said to be the *image* of T and is denoted by $Im\ T$.

$$Im\ T = \{T(\alpha) : \alpha \in V\}.$$

Theorem 4.21.5. Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping. Then $Im\,T$ is a subspace of W.

Proof. Let θ, θ' be the null elements of V and W respectively.
Since $T(\theta) = \theta'$, $\theta' \in Im\ T$. Therefore $Im\ T$ is not empty.

Case 1. $Im\ T = \{\theta'\}$. Then $Im\ T$ is a subspace of W.

Case 2. $Im\ T \neq \{\theta'\}$. Let $\xi \in Im\ T$. Then there exists an element α in V such that $T(\alpha) = \xi$.
This implies $T(c\alpha) = cT(\alpha) = c\xi$ for all $c \in F$.
Therefore $c\xi \in Im\ T$ for all $c \in F$.

Let $\xi \in Im\ T, \eta \in Im\ T$. Then there exist elements α, β in V such that $T(\alpha) = \xi, T(\beta) = \eta$.
$T(\alpha + \beta) = T(\alpha) + T(\beta) = \xi + \eta$. This implies $\xi + \eta \in Im\ T$.
Thus $\xi, \eta \in Im\ T \Rightarrow \xi + \eta \in Im\ T$ and $\xi \in Im\ T, c \in F \Rightarrow c\xi \in Im\ T$. This proves that $Im\ T$ is a subspace of W.

Note. $Im\ T$ is also called the *range* of T and is denoted by $R(T)$.

Theorem 4.21.6. Let V and W be vector spaces over a field F. Let $T : V \to W$ be a linear mapping and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. Then the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ generate $Im\ T$.

Proof. Let $\xi \in Im\ T$. Then there exists an element α in V such that $T(\alpha) = \xi$. Let $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$ for some scalars $c_i \in F$.

$$\text{Then } \xi = T(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n)$$
$$= c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_nT(\alpha_n), \text{ since } T \text{ is linear.}$$

Since each $T(\alpha_i) \in Im\ T$, it follows that $Im\ T$ is generated by the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$.

Note. The vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ are not necessarily distinct. If $Ker\ T = \{\theta\}$ then these vectors are distinct and by Theorem 4.21.4, they are linearly independent and therefore they form a basis of $Im\ T$.

Worked Examples (continued).

3. Find $Im\ T$ and the dimension of $Im\ T$, where T is the linear mapping in Example 1.

If $\{\alpha_1, \alpha_2, \alpha_3\}$ be a basis of the domain space $\mathbb{R}^3$, $Im\ T$ is the linear span of the vectors $T(\alpha_1), T(\alpha_2), T(\alpha_3)$.

$\{\epsilon_1 = (1,0,0), \epsilon_2 = (0,1,0), \epsilon_3 = (0,0,1)\}$ is a basis of $\mathbb{R}^3$.

$T(\epsilon_1) = (1,2,1), T(\epsilon_2) = (1,1,2), T(\epsilon_3) = (1,2,1)$.

Since $T(\epsilon_1) = T(\epsilon_3)$, $Im\ T$ is the linear span of the vectors (1, 2, 1) and (1, 1, 2). Hence $Im\ T = L\{(1,2,1),(1,1,2)\}$.

Since the vectors (1, 2, 1), (1, 1, 2) are linearly independent, the dimension of $Im\ T$ is 2.

Alternative method.

Let ξ be a arbitrary vector in $Im\ T$.

$$\text{Then } \xi = (x_1 + x_2 + x_3, 2x_1 + x_2 + 2x_3, x_1 + 2x_2 + x_3)$$
$$= x_1(1,2,1) + x_2(1,1,2) + x_3(1,2,1).$$

Thus ξ is a linear combination of the vectors (1, 2, 1), (1, 1, 2).

Hence $Im\ T = L\{(1,2,1),(1,1,2)\}$.

Since the set $\{(1,2,1),(1,1,2)\}$ is linearly independent, the dimension of $Im\ T$ is 2.

4. Find $Im\ T$ and the dimension of $Im\ T$, where T is the linear mapping in Example 2.

$Im\ T$ is the linear span of the vectors $T(\alpha_1), T(\alpha_2), T(\alpha_3)$ where $\{\alpha_1, \alpha_2, \alpha_3\}$ is a basis of $\mathbb{R}^3$.

$\{\epsilon_1 = (1,0,0), \epsilon_2 = (0,1,0), \epsilon_3 = (0,0,1)\}$ is a basis of $\mathbb{R}^3$.

$T(\epsilon_1) = (0,1,1,1), T(\epsilon_2) = (1,0,1,1), T(\epsilon_3) = (1,1,0,1)$.

Therefore $Im\ T = L\{(0,1,1,1),(1,0,1,1),(1,1,0,1)\}$.

The set $\{(0,1,1,1),(1,0,1,1),(1,1,0,1)\}$ is linearly independent. (Example 2) Therefore the dimension of $Im\ T$ is 3.

Nullity and Rank of a linear mapping.

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

Then $Ker\ T$ is a subspace of V. The dimension of $Ker\ T$ is called the *nullity* of T. $Im\ T$ is a subspace of W. The dimension of $Im\ T$ is called the *rank* of T.

If V be a finite dimensional vector space then $Ker\ T$, being a subspace of V, is finite dimensional. The number of vectors in a basis of V is finite. As $Im\ T$ is generated by the images of the vectors in a basis of V, $Im\ T$ also is finite dimensional.

Theorem 4.21.7. Let V and W be vector spaces over a field F and V is finite dimensional. If $T: V \to W$ be a linear mapping then

$\dim Ker\ T + \dim Im\ T = \dim V$.

In other words, the nullity of T+rank of $T = \dim V$.

Proof. **Case 1.** Let $Ker\ T = V$. Then $Im\ T$ consists of θ' only, where θ' is the null element in W. Therefore $\dim Im\ T = 0$. Thus $\dim Ker\ T + \dim Im\ T = \dim V + 0 = \dim V$.

Case 2. Let $Ker\ T = \{\theta\}$ and let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. Then $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a basis of $Im\ T$.

Therefore $\dim Ker\ T = 0$, $\dim V = n$, $\dim Im\ T = n$ and the theorem holds good.

Case 3. Let $Ker\ T$ be a non-trivial proper subspace of V. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ be a basis of $Ker\ T$. This basis of $Ker\ T$ can be extended to a basis of V and let the extended basis of V be $\{\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$.

$Im\ T$ is generated by the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$, by Theorem 4.21.6.

As $T(\alpha_1) = T(\alpha_2) = \cdots = T(\alpha_k) = \theta'$, $Im T$ is generated by the vectors $T(\alpha_{k+1}), T(\alpha_{k+2}), \ldots, T(\alpha_n)$.

Let us consider the relation $c_{k+1}T(\alpha_{k+1}) + c_{k+2}T(\alpha_{k+2}) + \cdots + c_nT(\alpha_n) = \theta'$ for some scalars c_i in F.

Since T is linear, this implies $T(c_{k+1}\alpha_{k+1} + c_{k+2}\alpha_{k+2} + \cdots + c_n\alpha_n) = \theta'$.

This shows that $c_{k+1}\alpha_{k+1} + c_{k+2}\alpha_{k+2} + \cdots + c_n\alpha_n$ is in $Ker\ T$.

Since $\{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ is an assumed basis for $Ker\ T$, $c_{k+1}\alpha_{k+1} + c_{k+2}\alpha_{k+2} + \cdots + c_n\alpha_n = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_k\alpha_k$ for some scalars $b_i \in F$.

Therefore $b_1\alpha_1 + b_2\alpha_2 + \cdots + b_k\alpha_k - c_{k+1}\alpha_{k+1} - c_{k+2}\alpha_{k+2} - \cdots - c_n\alpha_n = \theta$.

The linear independence of $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ requires $b_1 = b_2 = \cdots = b_k = c_{k+1} = \cdots = c_n = 0$.

The fact that $c_{k+1} = c_{k+2} = \cdots = c_n = 0$ proves that the set $\{T(\alpha_{k+1}), T(\alpha_{k+2}), \cdots, T(\alpha_n)\}$ is linearly independent and so this is a basis for $Im\ T$.

So dim $Im\ T = n - k$ and dim $Ker\ T+$ dim $Im\ T = k + (n-k) = n$.

Thus dim $Ker\ T +$ dim $Im\ T =$ dim V and the theorem is done.

Theorem 4.21.8. Let V and W be finite dimensional vector spaces of the same dimension over a field F and $T : V \to W$ be a linear mapping. Then T is one-to-one $\Leftrightarrow$ T is onto.

Proof. Let T be one-to-one . Then $Ker\ T = \{\theta\}$ and dim $Ker\ T = 0$.

The relation dim $Ker\ T+$ dim $Im\ T =$ dim V gives dim $Im\ T =$ dim V and therefore dim $Im\ T =$ dim W.

But $Im\ T$ is a subspace of W and as dim $Im\ T =$ dim W, $Im\ T = W$. Hence T is onto.

Conversely, let T be onto. Then $Im\ T = W$.

The relation dim $Ker\ T+$ dim $Im\ T =$ dim V gives dim $Ker\ T+$ dim $W =$ dim V.

As dim $V =$ dim W, dim $Ker\ T = 0$.

So $Ker\ T = \{\theta\}$ and this implies T is one-to-one.

Corollary. If T be a linear mapping on a finite dimensional vector space V, then T is one-to-one $\Leftrightarrow$ T is onto.

Linear mapping with prescribed images.

Theorem 4.21.9. Let V and W be vector spaces over a field F. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V and $\beta_1, \beta_2, \ldots, \beta_n$ be arbitrarily chosen elements (not necessarily distinct) in W. Then there exists one and only one linear mapping $T : V \to W$ such that $T(\alpha_i) = \beta_i$ for $i = 1, 2, \ldots, n$.

Proof. Let α be an arbitrary element of V. Since $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ is a basis, there exist unique scalars $r_1, r_2, \ldots, r_n$ in F such that $\alpha = r_1\alpha_1 + r_2\alpha_2 + \cdots + r_n\alpha_n$.

Let us define a mapping $T : V \to W$ by $T(\alpha) = r_1\beta_1 + r_2\beta_2 + \cdots + r_n\beta_n$ for $\alpha(= r_1\alpha_1 + r_2\alpha_2 + \cdots + r_n\alpha_n) \in V$.

The mapping T is such that $T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_2, \ldots, T(\alpha_n) = \beta_n$.

First we prove that T is linear.

Let $\xi, \eta \in V$ and
$$\begin{aligned} \xi &= s_1\alpha_1 + s_2\alpha_2 + \cdots + s_n\alpha_n, \\ \eta &= t_1\alpha_1 + t_2\alpha_2 + \cdots + t_n\alpha_n, \end{aligned}$$
where s_i, t_i are unique scalars in F determined by the basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Then $\xi + \eta = (s_1 + t_1)\alpha_1 + (s_2 + t_2)\alpha_2 + \cdots + (s_n + t_n)\alpha_n$
and $c\xi = (cs_1)\alpha_1 + (cs_2)\alpha_2 + \cdots + (cs_n)\alpha_n, c \in F$.

By definition,
$$\begin{aligned} T(\xi + \eta) &= (s_1 + t_1)\beta_1 + \cdots + (s_n + t_n)\beta_n; \\ T(c\xi) &= (cs_1)\beta_1 + (cs_2)\beta_2 + \cdots + (cs_n)\beta_n. \end{aligned}$$
$$\begin{aligned} T(\xi) + T(\eta) &= (s_1\beta_1 + \cdots + s_n\beta_n) + (t_1\beta_1 + \cdots + t_n\beta_n) \\ &= (s_1 + t_1)\beta_1 + \cdots + (s_n + t_n)\beta_n \\ &= T(\xi + \eta); \end{aligned}$$
$$\begin{aligned} \text{and } cT(\xi) &= c(s_1\beta_1 + s_2\beta_2 + \cdots + s_n\beta_n) \\ &= (cs_1)\beta_1 + (cs_2)\beta_2 + \cdots + (cs_n)\beta_n = T(c\xi). \end{aligned}$$
Therefore T is linear.

To prove that T is unique, let us assume that there exists another linear mapping $f : V \to W$ such that $f(\alpha_1) = \beta_1, f(\alpha_2) = \beta_2, \ldots, f(\alpha_n) = \beta_n$.

$$\begin{aligned} f(\alpha) &= f(r_1\alpha_1 + r_2\alpha_2 + \cdots + r_n\alpha_n) \\ &= r_1 f(\alpha_1) + r_2 f(\alpha_2) + \cdots + r_n f(\alpha_n), \text{ since } f \text{ is linear} \\ &= r_1\beta_1 + r_2\beta_2 + \cdots + r_n\beta_n \\ &= T(\alpha). \end{aligned}$$

Thus $f(\alpha) = T(\alpha)$ for all α in V and this implies $f = T$.
This completes the proof.

Worked Examples (continued).

5. Prove that the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (x + y, y + z, z + x), (x, y, z) \in \mathbb{R}^3$ is one-to one and onto.

Let $(x, y, z) \in Ker\ T$. Then $(x + y, y + z, z + x) = (0, 0, 0)$.
Therefore $x + y = 0,\ y + z = 0,\ z + x = 0$.
The solution is $x = 0, y = 0, z = 0$.

$Ker\ T = \{\theta\}$ and hence T is one-to-one.

$\{(1,0,0), (0,1,0), (0,0,1)\}$ is a basis of $\mathbb{R}^3$. $Im\ T$ is the linear span of the vectors $T(1,0,0), T(0,1,0), T(0,0,1)$.

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

The set of vectors $\{(1,0,1), (1,1,0), (0,1,1)\}$ being linearly independent, $Im\ T = \mathbb{R}^3$ and therefore T is an onto mapping.

6. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ which maps the basis vectors (1,0,0), (0,1,0), (0,0,1) of $\mathbb{R}^3$ to the vectors (1,1), (2,3), (3,2) respectively.

(i) Find $T(1,1,0), T(6,0,-1)$. (ii) Find $Ker\ T$ and $Im\ T$.

(iii) Prove that T is not one-to-one, but onto.

Let $\xi = (x, y, z)$ be an arbitrary vector in $\mathbb{R}^3$.
$\xi = x(1,0,0) + y(0,1,0) + z(0,0,1)$.

Since T is linear,

$$\begin{aligned} T(\xi) &= xT(1,0,0) + yT(0,1,0) + zT(0,0,1) \\ &= x(1,1) + y(2,3) + z(3,2) \\ &= (x + 2y + 3z, x + 3y + 2z). \end{aligned}$$

So T is defined by $T(x, y, z) = (x + 2y + 3z, x + 3y + 2z), (x, y, z) \in \mathbb{R}^3$.

(i) $T(1,1,0) = (3,4)$, $T(6,0,-1) = (3,4)$.

(ii) Let $(x, y, z) \in Ker\ T$. Then $T(x, y, z) = (0,0)$.
This gives $x + 2y + 3z = 0$, $x + 3y + 2z = 0$.

The solution is $\frac{x}{-5} = \frac{y}{1} = \frac{z}{1} = k$. Therefore $(x, y, z) = k(-5,1,1)$, where $k \in \mathbb{R}$.

Consequently, $Ker\ T = L\{\alpha\}$ where $\alpha = (-5,1,1)$.

$Im\ T$ is the linear span of the vectors $T(\alpha_1), T(\alpha_2), T(\alpha_3)$ where $\{\alpha_1, \alpha_2, \alpha_3\}$ is any basis of $\mathbb{R}^3$.

Since $\{(1,0,0),(0,1,0),(0,0,1)\}$ is a basis of $\mathbb{R}^3$, $Im\ T = L\{(1,1),(2,3),(3,2)\}$.

(iii) Since $Ker\ T \neq \{\theta\}$, T is not one-to-one.

$Im\ T = L\{(1,1),(2,3),(3,2)\}$. These vectors are linearly dependent in $\mathbb{R}^2$. The subset $\{(1,1),(2,3)\}$ is linearly independent in $\mathbb{R}^2$. Therefore $\dim Im\ T = 2$.

Since $Im\ T$ is a subspace of $\mathbb{R}^2$ and $\dim Im\ T = 2, Im\ T = \mathbb{R}^2$. Therefore T is onto.

7. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ which maps the basis vectors $(0,1,1),(1,0,1),(1,1,0)$ of $\mathbb{R}^3$ to (1, 1, 1), (1, 1, 1), (1, 1, 1) respectively. Verify that $\dim Ker\ T + \dim Im\ T = 3$.

Let $\xi = (x, y, z)$ be an arbitrary vector of the domain space $\mathbb{R}^3$.

Let $\xi = c_1(0,1,1) + c_2(1,0,1) + c_3(1,1,0)$, where c_i are unique scalars in $\mathbb{R}$. Then $c_2 + c_3 = x$, $c_3 + c_1 = y$, $c_1 + c_2 = z$.

Solving, we have $c_1 = \frac{y+z-x}{2}$, $c_2 = \frac{z+x-y}{2}$, $c_3 = \frac{x+y-z}{2}$.

Since T is linear,

$$\begin{aligned} T(\xi) &= c_1T(0,1,1) + c_2T(1,0,1) + c_3T(1,1,0) \\ &= c_1(1,1,1) + c_2(1,1,1) + c_3(1,1,1) \\ &= (c_1 + c_2 + c_3, c_1 + c_2 + c_3, c_1 + c_2 + c_3) \\ &= (\tfrac{y+z+x}{2}, \tfrac{y+z+x}{2}, \tfrac{y+z+x}{2}). \end{aligned}$$

Hence T is defined by $T(x, y, z) = (\frac{x+y+z}{2}, \frac{x+y+z}{2}, \frac{x+y+z}{2}), (x, y, z) \in \mathbb{R}^3$.

Let $(x, y, z) \in Ker\ T$. Then $T(x, y, z) = (0,0,0)$.

This gives $x + y + z = 0$. Let $y = c, z = d$, where $c, d \in \mathbb{R}$. Then $x = -c - d$.

$(x, y, z) = (-c - d, c, d) = c(-1,1,0) + d(-1,0,1)$, where $c, d \in \mathbb{R}$.

Hence $Ker\ T = L\{(-1,1,0),(-1,0,1)\}$ and since $(-1,1,0)$ and $(-1,0,1)$ are linearly independent, $\dim Ker\ T = 2$.

$Im\ T$ is the linear span of the vectors $T(\alpha), T(\beta), T(\gamma)$, where $\{\alpha, \beta, \gamma\}$ is any basis of the domain space $\mathbb{R}^3$.

Since (0,1,1), (1,0,1), (1,1,0) is a basis of $\mathbb{R}^3$, $Im\ T = L\{(1,1,1)\}$.

Hence $\dim Im\ T = 1$ and $\dim Ker\ T + \dim Im\ T = 2 + 1 = 3$.

8. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ which maps the basis vectors $(0,1,1),(1,0,1),(1,1,0)$ of $\mathbb{R}^3$ to the vectors (2, 0, 0), (0, 2, 0), (0, 0, 2) respectively. Find $Ker\ T$ and $Im\ T$. Verify that $\dim Ker\ T + \dim Im\ T = 3$.

Let $\xi = (x, y, z)$ be an arbitrary vector in $\mathbb{R}^3$.

Let $\xi = c_1(0,1,1) + c_2(1,0,1) + c_3(1,1,0)$, where c_i are unique scalars in $\mathbb{R}$. Then $c_2 + c_3 = x$, $c_3 + c_1 = y$, $c_1 + c_2 = z$.

Solving, we have $c_1 = \frac{y+z-x}{2}$, $c_2 = \frac{z+x-y}{2}$, $c_3 = \frac{x+y-z}{2}$.

Since T is linear,

$$\begin{aligned} T(\xi) &= c_1T(0,1,1) + c_2T(1,0,1) + c_3T(1,1,0) \\ &= \tfrac{y+z-x}{2}(2,0,0) + \tfrac{z+x-y}{2}(0,2,0) + \tfrac{x+y-z}{2}(0,0,2) \\ &= (y + z - x, z + x - y, x + y - z). \end{aligned}$$

Therefore $T(x, y, z) = (y + z - x, z + x - y, x + y - z)$, $(x, y, z) \in \mathbb{R}^3$.

Let $(x, y, z) \in Ker\ T$. Then $y+z-x = 0$, $z+x-y = 0$, $x+y-z = 0$.

This gives $x = y = z = 0$. Hence $Ker\ T = \{\theta\}$ and $\dim Ker\ T = 0$.

$Im\ T$ is the linear span of the vectors $T(\alpha_1), T(\alpha_2), T(\alpha_3)$, where $\{\alpha_1, \alpha_2, \alpha_3\}$ is any basis of the domain space $\mathbb{R}^3$. Since $\{(0,1,1),(1,0,1),(1,1,0)\}$ is a basis of $\mathbb{R}^3$, $Im\ T = L\{(2,0,0),(0,2,0),(0,0,2)\}$.

Since the set of the vectors $\{(2,0,0),(0,2,0),(0,2,0)\}$ is linearly independent, $\dim Im\ T = 3$.

Hence $\dim Ker\ T + \dim Im\ T = 0 + 3 = 3$.

9. Find a linear operator T on $\mathbb{R}^3$ such that $Ker\ T$ is the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ of $\mathbb{R}^3$.

Let $\xi = (a, b, c)$ be a vector in the subspace U. Then $a + b + c = 0$.
$\xi = (a, b, -a-b) = a(1, 0, -1) + b(0, 1, -1)$, $a, b \in \mathbb{R}$. This shows that $\xi \in L\{(1, 0, -1), (0, 1, -1)\}$. Because the vectors $(1, 0, -1)$ and $(0, 1, -1)$ are linearly independent, $\{(1, 0, -1), (0, 1, -1)\}$ is a basis of U.

Since $U = ker\ T$, $T(1, 0, -1) = (0, 0, 0), T(0, 1, -1) = (0, 0, 0)$.

The basis $\{(1, 0, -1), (0, 1, -1)\}$ of U can be extended to a basis $\{(1, 0, -1), (0, 1, -1), (1, 0, 0)\}$ of $\mathbb{R}^3$.

Let T be the linear operator on $\mathbb{R}^3$ such that $T(1, 0, -1) = (0, 0, 0), T(0, 1, -1) = (0, 0, 0), T(1, 0, 0) = (1, 0, 0)$, then $ker\ T = U$.

Let $(x, y, z) \in \mathbb{R}^3$ and $(x, y, z) = c_1(1, 0, -1) + c_2(0, 1, -1) + c_3(1, 0, 0)$. Then $c_1 = -y - z, c_2 = y, c_3 = x + y + z$.

$$\begin{aligned} T(x, y, z) &= (-y - z)T(1, 0, -1) + yT(0, 1, -1) + (x + y + z)T(1, 0, 0) \\ &= (-y - z)(0, 0, 0) + y(0, 0, 0) + (x + y + z)(1, 0, 0) \\ &= (x + y + z, 0, 0),\ (x, y, z) \in \mathbb{R}^3. \end{aligned}$$

Note. As the image of the basis vector $(1, 0, 0)$ can be chosen arbitrarily (other than $(0, 0, 0)$), T is not unique.

10. Find a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $Im\ T$ is the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x + y - z = 0\}$ of $\mathbb{R}^3$.

Let $\xi = (a, b, c)$ be a vector in the subspace U. Then $a + b - c = 0$.
$\xi = (a, b, a + b) = a(1, 0, 1) + b(0, 1, 1)$, $a, b \in \mathbb{R}$. This shows that $\xi \in L\{(1, 0, 1), (0, 1, 1)\}$. Because the vectors $(1, 0, 1)$ and $(0, 1, 1)$ are linearly independent, $\{(1, 0, 1), (0, 1, 1)\}$ is a basis of U.

$Im\ T$ is generated by the images of the vectors of a basis. Let us take the standard basis $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ of $\mathbb{R}^3$ and let $T(1, 0, 0) = (1, 0, 1), T(0, 1, 0) = (0, 1, 1), T(0, 0, 1) = (0, 1, 1)$.
Then $Im\ T = L\{(1, 0, 1), (0, 1, 1)\} = U$.

Let $(x, y, z) \in \mathbb{R}^3$.

$$\begin{aligned} T(x, y, z) &= xT(1, 0, 0) + yT(0, 1, 0) + zT(0, 0, 1), \text{ since } T \text{ is linear} \\ &= x(1, 0, 1) + y(0, 1, 1) + z(0, 1, 1) \\ &= (x, y + z, x + y + z),\ (x, y, z) \in \mathbb{R}^3. \end{aligned}$$

Note. As the image of the basis vector $(0, 0, 1)$ can be chosen arbitrarily (as a scalar multiple of $(0, 1, 1)$, or as a scalar multiple of $(1, 0, 1)$), T is not unique.

Exercises 16

1. Examine whether T is a linear mapping. If T is linear, find $Ker\ T$ and $Im\ T$.
   (i) $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(x, y) = (x + y, x - y)$, $(x, y) \in \mathbb{R}^2$;
   (ii) $T : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $T(x, y) = (x + 2y, 2x + y, x + y)$, $(x, y) \in \mathbb{R}^2$;
   (iii) $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (yz, zx, xy)$, $(x, y, z) \in \mathbb{R}^3$;
   (iv) $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(x, y, z) = (x+2y+3z, 3x+2y+z, x+y+z)$, $(x, y, z) \in \mathbb{R}^3$;
   (v) $T : \mathbb{R}^3 \to \mathbb{R}^4$ defined by $T(x, y, z) = (-x + y + z, x - y + z, x + y - z, x + y + z)$, $(x, y, z) \in \mathbb{R}^3$;
   (vi) $T : \mathbb{R}^3 \to \mathbb{R}$ defined by $T(x, y, z) = x + y + z$, $(x, y, z) \in \mathbb{R}^3$;
   (vii) $T : \mathbb{R}_{2\times 2} \to \mathbb{R}_{2\times 2}$ defined by $T(A) = \frac{1}{2}(A + A^t)$, $A \in \mathbb{R}_{2\times 2}$.

2. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ that maps the basis vectors $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ of $\mathbb{R}^3$ to the vectors $(0, 1, 0)$, $(0, 0, 1)$, $(1, 0, 0)$ respectively. Find $Ker\ T$ and $Im\ T$. Verify that dim $Ker\ T+$ dim $Im\ T = 3$.

3. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ that maps the basis vectors $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ of $\mathbb{R}^3$ to the vectors $(2, 1, 1)$, $(1, 2, 1)$, $(1, 1, 2)$ respectively. Find $Ker\ T$ and $Im\ T$. Verify that dim $Ker\ T+$ dim $Im\ T = 3$.

4. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ that maps the basis vectors $(2, 1, 1)$, $(1, 2, 1)$, $(1, 1, 2)$ of $\mathbb{R}^3$ to the vectors $(1, 1, 1)$, $(1, 1, 1)$, $(1, 1, 1)$ respectively. Find $Ker\ T$ and $Im\ T$. Verify that dim $Ker\ T+$ dim $Im\ T = 3$.

5. Determine the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^4$ that maps the basis vectors $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ of $\mathbb{R}^3$ to the vectors $(0, 1, 1, 1)$, $(1, 0, 1, 1)$, $(1, 1, 0, 1)$ respectively. Find $Ker\ T$ and $Im\ T$. Verify that dim $Ker\ T+$ dim $Im\ T = 3$.

6. Prove that the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ that maps the basis vectors $(1, 2, 2)$, $(2, 1, 2)$, $(2, 2, 1)$ of $\mathbb{R}^3$ to the vectors $(0, 1, 1)$, $(1, 0, 1)$, $(1, 1, 0)$ respectively is one-to-one and onto.

7. Find a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $ker\ T$ is the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x - y - z = 0\}$ of $\mathbb{R}^3$.

8. Find a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $Im\ T$ is the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ of $\mathbb{R}^3$.

9. Find a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that $Im\ T$ is the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x - y = 0, y - z = 0\}$ of $\mathbb{R}^3$.

10. Find a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ such that the subspace $U = \{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$ of $\mathbb{R}^3$ is mapped to itself, T not being the identity mapping.

Composition of linear mappings.

Let V, W and U be vector spaces over a field F and let $T : V \to W, S : W \to U$ be linear mappings. The composite mapping $S \circ T : V \to U$ is defined by $S \circ T(\alpha) = S\{T(\alpha)\}, \alpha \in V$.

The composite $S \circ T$ is generally denoted by ST and it is also said to be the *product mapping* ST.

Theorem 4.21.10. Let V, W and U be vector spaces over a field F and $T : V \to W$, $S : W \to U$ be linear mappings. Then the composite mapping $ST : V \to W$ is linear.

Proof. Let α, β in V and $c \in F$.

$$\begin{aligned} ST(\alpha+\beta) &= S[T(\alpha+\beta)] \\ &= S(T\alpha + T\beta), \text{ since } T \text{ is linear} \\ &= S(T\alpha) + S(T\beta), \text{ since } S \text{ is linear} \\ &= ST(\alpha) + ST(\beta). \end{aligned}$$

$$\begin{aligned} ST(c\alpha) &= S[T(c\alpha)] \\ &= S[cT(\alpha)], \text{ since } T \text{ is linear} \\ &= cS[T(\alpha)], \text{ since } S \text{ is linear} \\ &= cST(\alpha). \end{aligned}$$

This proves that ST is linear.

We recall here that for two non-empty sets A and B, a mapping $f : A \to B$ is said to be invertible if there exists a mapping $g : B \to A$ such that $g \circ f = i_A$ and $f \circ g = i_B$. f is invertible if and only if f is both one-to-one and onto.

Definition. Let V and W be vector spaces over a field F. A linear mapping $T : V \to W$ is said to be *invertible* if there exists a mapping $S : W \to V$ such that $ST = I_V$ and $TS = I_W$.

In this case, S is said to be *an inverse* of T.

Theorem 4.21.11. Let V and W be vector spaces over a field F. If a linear mapping $T : V \to W$ be invertible then T has a unique inverse.

Proof. If possible, let there be two inverses $S_1 : W \to V$ and $S_2 : W \to V$ of T.

Then $S_1T = S_2T = I_V$ and $TS_1 = TS_2 = I_W$.

But $S_1(TS_2) = (S_1T)S_2$ and this implies $S_1I_W = I_VS_2$.

That is, $S_1 = S_2$. This proves that the inverse of T is unique.

Note. The unique inverse of the mapping T is denoted by T^{-1}.

Theorem 4.21.12. Let V and W be vector spaces over a field F. A linear mapping $T : V \to W$ is invertible if and only if T is one-to-one and onto.

Proof. Let $T : V \to W$ be invertible. Then there exists a mapping $S : W \to V$ such that $ST = I_V$ and $TS = I_W$.

To prove that T is one-to-one, let $\alpha, \beta \in V$.

$$\begin{aligned} T(\alpha) = T(\beta) &\Rightarrow ST(\alpha) = ST(\beta) \\ &\Rightarrow \alpha = \beta, \text{ since } ST = I_V. \end{aligned}$$

Therefore T is one-to-one.

To prove that T is onto, let $\gamma \in W$.

Since $TS = I_W$, we have $TS(\gamma) = \gamma$.

That is, $T\{S(\gamma)\} = \gamma$ and this shows that $S(\gamma)$ is a pre-image of γ under T. So T is onto.

Therefore T is both one-to-one and onto.

Conversely, let $T : V \to W$ be both one-to-one and onto.

Let $\alpha \in V$ and $T(\alpha) = \gamma$. Since T is one-to-one, α is the unique pre-image of γ under T. Since T is onto, each γ in W has a pre-image in V. Let us define a mapping $S : W \to V$ by $S(\gamma) = \alpha$ (the pre-image of γ under T), $\gamma \in W$.

Then $ST(\alpha) = S(\gamma) = \alpha$ for all $\alpha \in V$ and $TS(\gamma) = T(\alpha) = \gamma$ for all $\gamma \in W$.

Therefore $ST = I_V$ and $TS = I_W$ and this proves that T is invertible. This completes the proof.

Theorem 4.21.13. Let V and W be vector spaces over a field F. If a linear mapping $T : V \to W$ be invertible, then the inverse mapping $T^{-1} : W \to V$ is linear.

Proof. Let $\alpha', \beta' \in W$ and $T^{-1}(\alpha') = \alpha, T^{-1}(\beta') = \beta$.

Then $\alpha, \beta \in V$ and $T(\alpha) = \alpha'$ and $T(\beta) = \beta'$.

Since T is linear, $T(\alpha+\beta) = T(\alpha) + T(\beta) = \alpha' + \beta'$.

Therefore $T^{-1}(\alpha'+\beta') = \alpha + \beta = T^{-1}(\alpha') + T^{-1}(\beta')$.

$$\begin{aligned} \text{Since } T \text{ is linear, } T(c\alpha) &= cT(\alpha), c \in F \\ &= c\alpha'. \end{aligned}$$

Therefore $T^{-1}(c\alpha') = c\alpha = cT^{-1}(\alpha')$ for all $c \in F$.

This proves that T^{-1} is linear.

Note 1. The linear mapping $T^{-1} : W \to V$ has the property that $T^{-1}T = I_V$, $TT^{-1} = I_W$.

2. If $T : V \to V$ be an invertible linear mapping on V then the linear mapping $T^{-1} : V \to V$ has the property that $T^{-1}T = TT^{-1} = I_V$.

Definition. A linear mapping $T : V \to W$ is said to be *non-singular* if T be invertible.

Isomorphism. Let V and W be vector spaces over a field F. A linear mapping $T : V \to W$ is said to be an *isomorphism* if T is both one-to-one and onto.

Since T is both one-to-one and onto, T is invertible and $T^{-1} : W \to V$ is also a linear mapping which is both one-to-one and onto.

Thus the existence of an isomorphism $T : V \to W$ implies the existence of another isomorphism $T^{-1} : W \to V$. In this case the vector spaces V and W are said to be *isomorphic*.

Theorem 4.21.14. Two finite dimensional vector spaces V and W over a field F are isomorphic if and only if $\dim V = \dim W$.

Proof. Let V and W be isomorphic. Then there exists a linear mapping $T : V \to W$ such that T is both one-to-one and onto.

Since T is onto, $Im\ T = W$; and since T is one-to-one, $Ker\ T = \{\theta\}$. Therefore the relation $\dim Ker\ T + \dim Im\ T = \dim V$ gives $\dim W = \dim V$.

Conversely, let $\dim V = \dim W = n$. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V and $\{\beta_1, \beta_2, \ldots, \beta_n\}$ be a basis of W. By the Theorem 4.21.9, there exists a unique linear mapping $T : V \to W$ such that $T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_2, \ldots, T(\alpha_n) = \beta_n$.

Let $\alpha \in Ker\ T$ and $\alpha = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$ for some $c_i \in F$.

Then $T(\alpha) = c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_nT(\alpha_n)$

Since $T(\alpha) = \theta'$, we have $\theta' = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$.

This implies $c_1 = 0, c_2 = 0, \ldots, c_n = 0$, since $\{\beta_1, \beta_2, \ldots, \beta_n\}$ is a basis. Therefore $\alpha = \theta$ and so $KerT = \{\theta\}$ and this implies T is one-to-one.

The relation $\dim Ker\ T + \dim Im\ T = \dim V$ gives $\dim Im\ T = \dim V$, since $\dim Ker\ T = 0$. That is, $\dim Im\ T = \dim W$.

But $Im\ T \subset W$. Therefore $W = Im\ T$ and this implies T is onto.

T being both one-to-one and onto, V and W are isomorphic.
This completes the proof.

Theorem 4.21.15. Let V and W be finite dimensional vector spaces of same dimension over a field F and $T : V \to W$ be a linear mapping. Then T is an isomorphism if and only if T maps a basis of V to a basis of W.

Proof. Let $\dim V = \dim W = n$. Let T be an isomorphism and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V.

Since T is linear and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis of V, the vectors $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ generate $Im\ T$.

Because T is an isomorphism, T is one-to-one and onto.

Since T is onto, $Im\ T = W$. Since T is one-to-one, the set of vectors $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is linearly indepedent.

Therefore $Im\ T$ (i.e., W) is generated by the linearly independent set of vectors $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$. Since $\dim W = n$, $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a basis of W.

Conversely, let $\dim V = \dim W = n$ and a linear mapping $T : V \to W$ be such that T maps a basis of V to a basis of W.

Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V and T maps this basis to the basis $\{\beta_1, \beta_2, \ldots, \beta_n\}$ of W such that $T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_2, \ldots, T(\alpha_n) = \beta_n$.

Let $\eta \in W$ and $\eta = c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$ for some $c_i \in F$.

Then $\eta = c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_nT(\alpha_n)$
$= T(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n)$.

As $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n \in V$, this shows that T is onto.

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

Then $\theta = T(\alpha) = a_1T(\alpha_1) + a_2T(\alpha_2) + \cdots + a_nT(\alpha_n)$.

This implies $a_1 = 0, a_2 = 0, \ldots, a_n = 0$, since $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a linearly independent set in W.

Therefore $\alpha = \theta$, i.e., $Ker\ T = \{\theta\}$. This implies T is one-to-one.

Since T is one-to-one and onto, T is an isomorphism.

This completes the proof.

Isomorphisms from V to F^n.

Theorem 4.21.16. Let V be a vector space of dimension n over a field F. Then V is isomorphic of F^n.

Proof. An isomorphism between V and F^n can be established in many ways.

Let $(\beta_1, \beta_2, ..., \beta_n)$ be an ordered basis of V. Then any vector ξ of V can be expressed as

$\xi = c_1\beta_1 + c_2\beta_2 + ... + c_n\beta_n$ where $c_1, c_2, ..., c_n$ are unique scalars in F.

Let us define a mapping $\phi : V \to F^n$ by

$$\phi(\xi) = \begin{pmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_n \end{pmatrix}, \text{ where } \xi = (c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n) \in V.$$

Let $\alpha = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n \in V$,

$\beta = b_1\beta_1 + b_2\beta_2 + \cdots + b_n\beta_n \in V$.

Then $\alpha + \beta = (a_1 + b_1)\beta_1 + (a_2 + b_2)\beta_2 + \cdots + (a_n + b_n)\beta_n \in V$.

$$\phi(\alpha) = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_n \end{pmatrix}, \ \phi(\beta) = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix} \text{ and}$$

$$\phi(\alpha+\beta) = \begin{pmatrix} a_1 + b_1 \\ a_1 + b_2 \\ \cdot \\ \cdot \\ a_n + b_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix} = \phi(\alpha) + \phi(\beta) \text{ (i)}$$

Let $p \in F$. Then $p\alpha \in V$ and $p\alpha = (pa_1)\beta_1 + (pa_2)\beta_2 + \cdots + (pa_n)\beta_n$.

$$\phi(p\alpha) = \begin{pmatrix} pa_1 \\ pa_2 \\ \cdot \\ \cdot \\ pa_n \end{pmatrix} = p\begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_n \end{pmatrix} = p\phi(\alpha) \text{ (ii)}$$

From (i) and (ii) ϕ is a homomorphism.

To prove that ϕ is one-to-one, let $\alpha, \beta \in V$ be such that

$$\phi(\alpha) = \phi(\beta) \quad \text{where } \alpha = a_1\beta_1 + a_2\beta_2 + \cdots + a_n\beta_n,$$
$$\beta = b_1\beta_1 + b_2\beta_2 + \cdots + b_n\beta_n.$$

$$\phi(\alpha) = \phi(\beta) \Rightarrow \begin{pmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}$$

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

$\Rightarrow \alpha = \beta$.

So ϕ is one-to-one.

To prove that ϕ is onto, let $\begin{pmatrix} r_1 \\ r_2 \\ \cdot \\ \cdot \\ r_n \end{pmatrix}$ be an element in F^n.

Then $r_1\beta_1 + r_2\beta_2 + \cdots + r_n\beta_n \in V$

and $\phi(r_1\beta_1 + r_2\beta_2 + \cdots + r_n\beta_n) = \begin{pmatrix} r_1 \\ r_2 \\ \cdot \\ \cdot \\ r_n \end{pmatrix}$.

So ϕ is onto.

Since ϕ is both one-to-one and onto, ϕ is an isomorphism.

Since ϕ is isomorphism, V is isomorphic to F^n.

This completes the proof.

Note 1. The isomorphism ϕ has been established on the choice of the ordered basis $(\beta_1, \beta_2, \ldots, \beta_n)$ of V. Different isomorphisms can be established on the choice of different ordered bases.

2. As ϕ is an isomorphism, V and F^n have the same structure as vector spaces except for the names of their elements. Therefore F^n serves as a prototype of a vector space V over F of dimension n.

3. A real vector space V of dimension n and the vector space $\mathbb{R}^n$ of n-tuples are isomorphic and therefore they have the same structure as vector spaces.

Worked Examples.

1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(x, y) = (x, 0)$, $(x, y) \in \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $S(x, y) = (0, y)$, $(x, y) \in \mathbb{R}^2$.

Describe the mappings ST and TS.

The mapping $ST : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $ST(x, y) = (0, 0)$, $(x, y) \in \mathbb{R}^2$.

The mapping $TS : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $TS(x, y) = (0, 0)$, $(x, y) \in \mathbb{R}^2$.

Both the linear mappings ST and TS on $\mathbb{R}^2$ are zero mappings although the mappings S and T are each different from 0.

2. Let V be a finite dimensional vector space over a field F and $S : V \to V$, $T : V \to V$ are linear mappings such that $ST = I$, I being the identity mapping on V. Prove that $TS = I$ too.

Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V. We prove that $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ are linearly independent.

Let us consider the relation $c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_nT(\alpha_n) = \theta$, where $c_i \in F$.

Then $S(c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_nT(\alpha_n)) = S(\theta) = \theta$.

Therefore $c_1ST(\alpha_1) + c_2ST(\alpha_2) + \cdots + c_nST(\alpha_n)) = \theta$, since S is linear.

Because $ST = I$, this implies $c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n = \theta$.

As $\alpha_1, \alpha_2, \ldots, \alpha_n$ are linearly indepedent, we have $c_1 = 0, c_2 = 0, \ldots, c_n = 0$ and this proves that $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ are linearly independent.

Since $\dim V = n$ and $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ are linearly independent in V, $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a basis of V.

Let $\beta \in V$. Then $\beta = b_1T(\alpha_1) + b_2T(\alpha_2) + \cdots + b_nT(\alpha_n)$ for some $b_i \in F$.

$TS(\beta) = b_1TST(\alpha_1) + b_2TST(\alpha_2) + \cdots + b_nTST(\alpha_n)$, since TS is linear.

or, $TS(\beta) = b_1T(\alpha_1) + b_2T(\alpha_2) + \cdots + b_nT(\alpha_n)$, since $ST = I$
$= \beta$.

This holds for all $\beta \in V$. Therefore $TS = I$.

3. Let V be a finite dimensional vector space over a field F and $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V. A linear mapping $T : V \to V$ is such that $T(\alpha_1) = \alpha_2, T(\alpha_2) = \alpha_3, \ldots, T(\alpha_{n-1}) = \alpha_n, T(\alpha_n) = \alpha_1$. Prove that $T^n = I$, I being the identity mapping on V.

$T^n(\alpha_1) = T^{n-1}(\alpha_2) = T^{n-2}(\alpha_3) = \cdots = T(\alpha_n) = \alpha_1$;
$T^n(\alpha_2) = T^{n-1}(\alpha_3) = T^{n-2}(\alpha_4) = \cdots = T^2(\alpha_n) = T(\alpha_1) = \alpha_2$;
... ...
$T^n(\alpha_n) = T^{n-1}(\alpha_1) = T^{n-2}(\alpha_2) = \cdots = T(\alpha_{n-1}) = \alpha_n$.

Let Let $\beta \in V$. Then $\beta = b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n$ for some $b_i \in F$.

$T^n(\beta) = b_1T^n(\alpha_1) + b_2T^n(\alpha_2) + \ldots + b_nT^n(\alpha_n)$, since T^n is linear
$= b_1\alpha_1 + b_2\alpha_2 + \ldots + b_n\alpha_n = \beta$.

This holds for all $\beta \in V$. Therefore $T^n = I$, I being the identity mapping on V.

4. A linear mapping $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ maps the vectors (0, 1,1), (1,0,1) and (1,1,0) to the vectors $(1, 1, -1), (1, -1, 1)$ and $(1, 0, 0)$ respectively. Show that ϕ is not an isomorphism.

The set of vectors $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ is a linearly independent set in $\mathbb{R}^3$. Therefore it is a basis of $\mathbb{R}^3$. The images of this basis vectors are the vectors $(1, 1, -1), (1, -1, 1), (1, 0, 0)$ respectively. The set

$\{(1, 1, -1), (1, -1, 1), (1, 0, 0)\}$ is a linearly dependent set in $\mathbb{R}^3$ and therefore it is not a basis of $\mathbb{R}^3$.

The linear mapping ϕ maps a basis of $\mathbb{R}^3$ to a set which is not a basis of $\mathbb{R}^3$. So ϕ is not an isomorphism.

Another method.

The set of vectors $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ is a basis of $\mathbb{R}^3$.

As $Im\ \phi$ is the linear span of the images of any basis vectors, $Im\ \phi = L\{(1, 1, -1), (1, -1, 1), (1, 0, 0)\}$. The set $\{(1, 1, -1), (1, -1, 1), (1, 0, 0)\}$ is a linearly dependent set in $\mathbb{R}^3$ and therefore the dimension of $Im\ \phi$ is less than 3. Consequently, ϕ is not onto and therefore ϕ is not an isomorphism.

5. A linear mapping $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ maps the vectors (1, 2, 3), (2, 3, 1) and (3, 1, 2) to the vectors (0,1,1), (1,0, 1) and (1,1,0) respectively. Show that ϕ is an isomorphism.

The set of vectors $\{(1, 2, 3), (2, 3, 1), (3, 1, 2)\}$ is a linearly independent set in $\mathbb{R}^3$, a vector space of dimension 3. Therefore it is a basis of $\mathbb{R}^3$.

The set of vectors $\{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ is a linearly independent set in $\mathbb{R}^3$, a vector space of dimension 3. Therefore it is a basis of $\mathbb{R}^3$.

The domain space and the co-domain space of the linear mapping ϕ are of the same dimension and ϕ maps a basis of the domain space to a basis of the co-domain space. So ϕ is an isomorphism.

6. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x, y, z) = (x - y, x + 2y, y + 3z)$, $(x, y, z) \in \mathbb{R}^3$. Show that T is non-singular and determine T^{-1}.

T is a linear mapping. Let us find $Ker\ T$.

Let $(a, b, c) \in Ker\ T$. Then $(a - b, a + 2b, b + 3c) = (0, 0, 0)$.

Therefore $a - b = 0, a + 2b = 0, b + 3c = 0$. This gives $a = b = c = 0$.

$Ker\ T = \{\theta\}$ and therefore T is one-to-one.

Here $V = \mathbb{R}^3, W = \mathbb{R}^3$ and therefore $\dim V = \dim W$. Since $T : V \to W$ is one-to-one, T is onto. Since T is one-to-one and onto, it is non-singular.

Let $T^{-1}(x, y, z) = (a, b, c)$.

Then $(x, y, z) = T(a, b, c) = (a - b, a + 2b, b + 3c)$.

Therefore $a - b = x,\ a + 2b = y,\ b + 3c = z$.

This gives $a = \frac{1}{3}(2x + y), b = \frac{1}{3}(-x + y), c = \frac{1}{9}(x - y + 3z)$.

Therefore $T^{-1} : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T^{-1}(x, y, z) = (\frac{2}{3}x + \frac{1}{3}y, -\frac{1}{3}x + \frac{1}{3}y, \frac{1}{9}x - \frac{1}{9}y + \frac{1}{3}z), (x, y, z) \in \mathbb{R}^3$.

Exercises 17

1. Let S and T be linear mappings of $\mathbb{R}^2$ to $\mathbb{R}^2$ defined by $S(x,y) = (x+y, y)$, $(x,y) \in \mathbb{R}^2$ and $T(x,y) = (x, x+y)$, $(x,y) \in \mathbb{R}^2$.

(i) Determine TS and ST. (ii) Prove that $(TS - ST)^2 = (ST - TS)^2 = I_{\mathbb{R}^2}$.

2. Let S and T be linear mappings of $\mathbb{R}^3$ to $\mathbb{R}^3$ defined by $S(x,y,z) = (z,y,x)$, $(x,y,z) \in \mathbb{R}^3$ and $T(x,y,z) = (x+y+z, y+z, z)$, $(x,y,z) \in \mathbb{R}^3$.

(i) Determine TS and ST.

(ii) Prove that S and T are invertible. Verify that $(ST)^{-1} = T^{-1}S^{-1}$.

3. Let T be a linear operator on a vector space V over a field F. Prove that

(i) Ker $T \subset$ Ker T^2, (ii) Im $T^2 \subset$ Im T.

4. Let V be a finite dimensional vector space over a field F and T is a linear operator on V such that $Ker\ T = Im\ T$. Prove that

(i) dim V is even; (ii) $T^2 = O$, O being the zero operator on V;

(iii) if dim V=2, $T^2 = O$ $(T \neq O)$ implies $Ker\ T = Im\ T$.

(iv) Give an example of a linear operator T on a vector space V such that $Ker\ T = Im\ T$.

(v) Give an example of a non-zero linear operator T on a vector space V such that $T^2 = O$ but $Ker\ T \neq Im\ T$.

5. Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ and $S : \mathbb{R}^2 \to \mathbb{R}^3$ be linear mappings. Prove that the mapping ST is neither one-to-one nor onto.

[**Hint.** S is not a surjection. T is not an injection.]

6. Let V be a vector space over a field F with a basis $(\alpha_1, \alpha_2, \alpha_3)$ and T is a linear operator on V such that $T(\alpha_1) = \alpha_2, T(\alpha_2) = \alpha_3, T(\alpha_3) = \theta$. Prove that $T^3 = O$, O being the zero operator on V.

7. Let U, V, W be finite dimensional vactor spaces over a field F. Let $T : V \to W$ be a linear mapping and $S : W \to U$ be an isomorphism. Prove that
(i) dim ker T= dim ker ST, (ii) dim Im T= dim Im ST.

8. A linear mapping $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ maps the vectors $(0, 1, 1)$, $(1, 0, 1)$ and $(1, 1, 0)$ to $(2, 1, 1)$, $(1, 2, 1)$ and $(1, 1, 2)$ respectively. Show that ϕ is an isomorphism.

9. A linear mapping $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ maps the vectors $(2, 1, 1)$, $(1, 2, 1)$ and $(1, 1, 2)$ to $(1, 1, -1)$, $(1, -1, 1)$ and $(1, 0, 0)$ respectively. Show that ϕ is not an isomorphism.

10. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ maps the basis vectors α, β, γ to $\alpha+\beta, \beta+\gamma, \gamma$ respectively. Show that T is an isomorphism.

Matrix representation of a linear mapping.

Let V and W be finite dimensional vector space over a field F with dim $V = n$ and dim $W = m$. Let $T : V \to W$ be a linear mapping. T is completely determined by its action on a given basis of V. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V and $(\beta_1, \beta_2, \ldots, \beta_m)$ be an ordered basis of W.

T is completely determined by the images $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$.

Each $T(\alpha_i)$ in W is a linear combination of the vectors $\beta_1, \beta_2, \ldots, \beta_m$.

$$
\begin{aligned}
\text{Let } T(\alpha_1) &= a_{11}\beta_1 + a_{21}\beta_2 + \cdots + a_{m1}\beta_m \\
T(\alpha_2) &= a_{12}\beta_1 + a_{22}\beta_2 + \cdots + a_{m2}\beta_m \\
&\cdots \quad \cdots \\
T(\alpha_n) &= a_{1n}\beta_1 + a_{2n}\beta_2 + \cdots + a_{mn}\beta_m,
\end{aligned}
$$

where a_{ij} are unique scalars in F determined by the ordered basis $(\beta_1, \beta_2, \ldots, \beta_m)$.

Let $\xi = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n$ be an arbitrary vector of V and let $T(\xi) = y_1\beta_1 + y_2\beta_2 + \cdots + y_m\beta_m$, $x_i, y_i \in F$.

$$
\begin{aligned}
T(\xi) &= T(x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n) \\
&= x_1T(\alpha_1) + x_2T(\alpha_2) + \cdots + x_nT(\alpha_n), \text{ since } T \text{ is linear} \\
&= x_1(a_{11}\beta_1 + a_{21}\beta_2 + \cdots + a_{m1}\beta_m) + x_2(a_{12}\beta_1 + a_{22}\beta_2 + \\
&\quad \cdots + a_{m2}\beta_m) + \cdots + x_n(a_{1n}\beta_1 + a_{2n}\beta_2 + \cdots + a_{mn}\beta_m).
\end{aligned}
$$

Since $\{\beta_1, \beta_2, ..., \beta_m\}$ is linearly independent,

$$
\begin{aligned}
y_1 &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\
y_2 &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\
&\cdots \quad \cdots \\
y_m &= a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n
\end{aligned}
$$

$$
\text{or, } \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}
$$

$$\text{or, } Y = AX, \quad \ldots \quad \text{(i)}$$

where $A = (a_{ij})_{m,n}$, X is the co-ordinate vector of an arbitrary element ξ in V relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and Y is the co-ordinate vector of $T(\xi)$ in W relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_m)$.

(i) is the *matrix representation* of the linear mapping T relative to the chosen ordered bases of V and W.

If $\xi = \alpha_j$, then $x_1 = 0, x_2 = 0, \ldots, x_{j-1} = 0, x_j = 1, x_{j+1} = 0, \ldots, x_n = 0$; and $y_1 = a_{1j}, y_2 = a_{2j}, \ldots, y_m = a_{mj}$.

Therefore the co-ordinate vector of $T(\alpha_j)$ relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_m)$ is given by the jth column of A.

The matrix $A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$ is said to be the *matrix associated with* the linear mapping T *relative to* the chosen ordered bases of V and W. A is also called the *matrix of* T relative to the chosen ordered bases.

Worked Examples.

1. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $T(x_1, x_2, x_3) = (3x_1 - 2x_2 + x_3, x_1 - 3x_2 - 2x_3), (x_1, x_2, x_3) \in \mathbb{R}^3$. Find the matrix of T relative to the ordered bases

(i) $((1,0,0),(0,1,0),(0,0,1))$ of $\mathbb{R}^3$ and $((1,0),(0,1))$ of $\mathbb{R}^2$;

(ii) $((0,1,0),(1,0,0),(0,0,1))$ of $\mathbb{R}^3$ and $((0,1),(1,0))$ of $\mathbb{R}^2$;

(iii) $((0,1,1),(1,0,1),(1,1,0))$ of $\mathbb{R}^3$ and $((1,0),(0,1))$ of $\mathbb{R}^2$.

(i)
$$\begin{aligned} T(1,0,0) &= (3,1) = 3(1,0) + 1(0,1); \\ T(0,1,0) &= (-2,-3) = -2(1,0) - 3(0,1); \\ T(0,0,1) &= (1,-2) = 1(1,0) - 2(0,1). \end{aligned}$$

Therefore the matrix of $T = \begin{pmatrix} 3 & -2 & 1 \\ 1 & -3 & -2 \end{pmatrix}$.

(ii)
$$\begin{aligned} T(0,1,0) &= (-2,-3) = -3(0,1) - 2(1,0); \\ T(1,0,0) &= (3,1) = 1(0,1) + 3(1,0); \\ T(0,0,1) &= (1,-2) = -2(0,1) + 1(1,0). \end{aligned}$$

Therefore the matrix of $T = \begin{pmatrix} -3 & 1 & -2 \\ -2 & 3 & 1 \end{pmatrix}$.

(iii)
$$\begin{aligned} T(0,1,1) &= (-1,-5) = -1(1,0) - 5(0,1); \\ T(1,0,1) &= (4,-1) = 4(1,0) - 1(0,1); \\ T(1,1,0) &= (1,-2) = 1(1,0) - 2(0,1). \end{aligned}$$

Therefore the matrix of $T = \begin{pmatrix} -1 & 4 & 1 \\ -5 & -1 & -2 \end{pmatrix}$.

2. Let $(\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2)$ be ordered bases of the real vector spaces V and W respectivly. A linear mapping $T : V \to W$ maps the basis vectors as $T(\alpha_1) = \beta_1 + \beta_2$, $T(\alpha_2) = 3\beta_1 - \beta_2$, $T(\alpha_3) = \beta_1 + 3\beta_2$.

Find the matrix of T relative to the ordered bases

(i) $(\alpha_1, \alpha_2, \alpha_3)$ of V and (β_2, β_1) of W;

(ii) $(\alpha_1 + \alpha_2, \alpha_2, \alpha_3)$ of V and $(\beta_1, \beta_1 + \beta_2)$ of W.

(i) $T(\alpha_1) = \beta_2 + \beta_1$; $T(\alpha_2) = -\beta_2 + 3\beta_1$; $T(\alpha_3) = 3\beta_2 + \beta_1$.

Therefore the matrix of $T = \begin{pmatrix} 1 & -1 & 3 \\ 1 & 3 & 1 \end{pmatrix}$.

(ii) $T(\alpha_1 + \alpha_2) = T(\alpha_1) + T(\alpha_2) = 4\beta_1 = 4\beta_1 + 0(\beta_1 + \beta_2)$;
$T(\alpha_2) = 3\beta_1 - \beta_2 = 4\beta_1 - (\beta_1 + \beta_2)$;
$T(\alpha_3) = \beta_1 + 3\beta_2 = -2\beta_1 + 3(\beta_1 + \beta_2)$.

Therefore the matrix of $T = \begin{pmatrix} 4 & 4 & -2 \\ 0 & -1 & 3 \end{pmatrix}$.

3. The matrix of a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ relative to the ordered bases $((0,1,1),(1,0,1),(1,1,0))$ of $\mathbb{R}^3$ and $((1,0),(1,1))$ of $\mathbb{R}^2$ is $\begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & 0 \end{pmatrix}$. Find T. Also find the matrix of T relative to the ordered bases $((1,1,0),(1,0,1),(0,1,1))$ of $\mathbb{R}^3$ and $((1,1),(0,1))$ of $\mathbb{R}^2$.

$T(0,1,1) = 1(1,0) + 2(1,1) = (3,2)$;
$T(1,0,1) = 2(1,0) + 1(1,1) = (3,1)$;
$T(1,1,0) = 4(1,0) + 0(1,1) = (4,0)$.

Let $(x,y,z) \in \mathbb{R}^3$ and let $(x,y,z) = c_1(0,1,1) + c_2(1,0,1) + c_3(1,1,0)$, where c_i are scalars in $\mathbb{R}$.

Then $c_2 + c_3 = x, c_1 + c_3 = y, c_1 + c_2 = z$.

Therefore $c_1 = \frac{1}{2}(y + z - x); c_2 = \frac{1}{2}(z + x - y), c_3 = \frac{1}{2}(x + y - z)$.

$$\begin{aligned} T(x,y,z) &= c_1T(0,1,1) + c_2T(1,0,1) + c_3T(1,1,0) \\ &= c_1(3,2) + c_2(3,1) + c_3(4,0) \\ &= (3c_1 + 3c_2 + 4c_3, 2c_1 + c_2) \\ &= (2x + 2y + z, \tfrac{1}{2}(-x + y + 3z)). \end{aligned}$$

T is given by $T(x,y,z) = (2x + 2y + z, \frac{1}{2}(-x + y + 3z)), (x,y,z) \in \mathbb{R}^3$.

$T(1,1,0) = (4,0)$; $T(1,0,1) = (3,1)$; $T(0,1,1) = (3,2)$.

Let $(4,0) = c_1(1,1) + c_2(0,1)$. Then $c_1 = 4, c_1 + c_2 = 0$. This gives $c_1 = 4, c_2 = -4$.

Let $(3,1) = c_1(1,1) + c_2(0,1)$. Then $c_1 = 3, c_1 + c_2 = 1$. This gives $c_1 = 3, c_2 = -2$.

Let $(3,2) = c_1(1,1) + c_2(0,1)$. Then $c_1 = 3, c_1 + c_2 = 2$. This gives $c_1 = 3; c_2 = -1$.

Therefore the matrix of $T = \begin{pmatrix} 4 & 3 & 3 \\ -4 & -2 & -1 \end{pmatrix}$.

Theorem 4.21.17. Let V and W be vector spaces of finite dimension over a field F and $T : V \to W$ be a linear mapping. Then

rank of T = rank of the matrix $m(T)$,

where $m(T)$ is the matrix of T relative to *any* chosen pair of ordered bases of V and W respectively.

Proof. Let $\dim V = n$, $\dim W = m$. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\beta_1, \beta_2, \ldots, \beta_m)$ be some ordered bases of V and W respectively. Let the matrix of T relative to these chosen bases be

$$A = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & & \cdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix}, \text{ where } c_{ij} \in F.$$

Then $T(\alpha_j) = c_{ij}\beta_1 + c_{2j}\beta_2 + \cdots + c_{mj}\beta_m$ for $j = 1, 2, \ldots, n$.

Let us consider the isomorphism $\phi : W \to F^m$ defined by

$$\phi(c_1\beta_1 + c_2\beta_2 + \cdots + c_m\beta_m) = \begin{pmatrix} c_1 \\ c_2 \\ \cdots \\ c_m \end{pmatrix}.$$

Since $\phi T : V \to F^m$ is a linear mapping and $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a basis of V, $\phi T(\alpha_1), \phi T(\alpha_2), \ldots, \phi T(\alpha_n)$ generate $Im\ \phi T$.

But $\phi T(\alpha_j)$ is the j-th column vector of the matrix A. Therefore $Im\ \phi T$ is the column space of the matrix A.

Hence rank of $\phi T = \dim Im\ \phi T$= column rank of A = rank of A... (i)

We have $\dim Ker\ T + \dim Im\ T = \dim V$; and
$\dim Ker\ \phi T + \dim Im\ \phi T = \dim V$ ... (ii)

We now prove that $\dim Ker\ T = \dim Ker\ \phi\ T$.

Let $\alpha \in Ker\ T$. Then $T(\alpha) = \theta \in W$. Therefore $\phi T(\alpha) = \theta \in F^m$ and this implies $\alpha \in Ker\ \phi T$. Hence $Ker\ T \subset Ker\ \phi T$ .... (iii)

Let $\beta \in Ker\ \phi T$. Then $\phi T(\beta) = \theta \in F^m$ and this implies $T(\beta) = \theta \in W$, since ϕ is an isomorphism. Therefore $\beta \in Ker\ T$. Hence $Ker\ \phi T \subset Ker\ T$ ... (iv)

From (iii) and (iv) $Ker\ T = Ker\ \phi T$ and therefore $\dim Ker\ T = \dim Ker\ \phi T$. From (ii) we have $\dim Im\ T = \dim Im\ \phi T$, i.e., rank of T = rank of ϕT ... (v)

From (i) and (v) we have rank of T = rank of A, i.e., rank of T = rank of $m(T)$.

This completes the proof.

Matrix of the composite mapping.

Theorem 4.21.18. Let $T : V \to U$ and $S : U \to W$ be linear mappings where V, U and W are finite dimensional vector spaces over a field F. Then relative to a choice of ordered bases

$$m(ST) = m(S) \cdot m(T),$$

where $m(T)$ is the matrix of T relative to the chosen bases.

Proof. Let $\dim V = n$, $\dim U = p$, $\dim W = m$. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V, $(\beta_1, \beta_2, \ldots, \beta_p)$ be an ordered basis of U, $(\gamma_1, \gamma_2, \ldots, \gamma_m)$ be an ordered basis of W.

Relative to the bases, let the matrices of T, S and ST be

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & & \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}, \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ & \cdots & \cdots & \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{bmatrix}, \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ & & \cdots & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}$$

respectively.

$$\begin{aligned} \text{Then } T(\alpha_j) &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{pj}\beta_p \text{ for } j = 1, 2, \ldots, n \\ S(\beta_j) &= b_{1j}\gamma_1 + b_{2j}\gamma_2 + \cdots + b_{mj}\gamma_m \text{ for } j = 1, 2, \ldots, p \\ ST(\alpha_j) &= c_{1j}\gamma_1 + c_{2j}\gamma_2 + \cdots + c_{mj}\gamma_m \text{ for } j = 1, 2, \ldots, n. \\ \text{But } ST(\alpha_j) &= S[T(\alpha_j)], \text{ since} S \text{ and } T \text{ are linear} \\ &= S[a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{pj}\beta_p] \\ &= a_{1j}S(\beta_1) + a_{2j}S(\beta_2) + \cdots + a_{pj}S(\beta_p) \\ &= a_{1j}[b_{11}\gamma_1 + \cdots + b_{m1}\gamma_m] + a_{2j}[b_{12}\gamma_1 + \cdots \\ &\quad + b_{m2}\gamma_m] + \cdots + a_{pj}[b_{1p}\gamma_1 + \cdots + b_{mp}\gamma_m] \\ &= \sum_{k=1}^{p} b_{1k}a_{kj}\gamma_1 + \sum_{k=1}^{p} b_{2k}a_{kj}\gamma_2 + \cdots + \sum_{k=1}^{p} b_{mk}a_{kj}\gamma_m \end{aligned}$$

Therefore $c_{1j} = \sum_{k=1}^{p} b_{1k}a_{kj}$, $c_{2j} = \sum_{k=1}^{p} b_{2k}a_{kj}, \ldots, c_{mj} = \sum_{k=1}^{p} b_{mk}a_{kj}$.

So we have

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \cdots & \cdots & & \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ & \cdots & \cdots & \\ b_{m1} & b_{m2} & \cdots & b_{mp} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix}.$$

That is, $m(ST) = m(S) \cdot m(T)$. This completes the proof.

Theorem 4.21.19. Let V and W be vector spaces of finite dimensions over a field F and let $T : V \to W$ be a linear mapping. Then T is invertible (non-singular) if and only if the matrix of T relative to any chosen pair of ordered bases of V and W is non-singular.

Proof. Let $T : V \to W$ be invertible. Then T is one-to-one and onto. Since T is one-to-one, dim $Ker\ T = 0$. Since T is onto, $Im\ T = W$.

dim $Ker\ T +$ dim $Im\ T =$ dim V gives dim $V =$ dim W.

Let dim $V =$ dim $W = n$.

Let A be the matrix of T relative to the pair of bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V and $(\beta_1, \beta_2, \ldots, \beta_n)$ of W. Then A is an $n \times n$ matrix.

Rank of $A=$ rank of $T =$dim $Im\ T = n$. Therefore A being an $n \times n$ matrix of rank n, is non-singular.

Conversely, let the matrix $m(T)$ relative to any chosen pair of bases of V and W be non-singular.

Then $m(T)$ is a square matrix. Let the order of $m(T)$ be n. Then the rank of $m(T)$ is n.

Since the order of $m(T)$ is n, dim $V =$ dim $W = n$.

Also the rank of $T = n$, since rank of $T =$ rank of $m(T)$.

Therefore $Im\ T = W$ and this implies T is onto.

dim $Ker\ T +$ dim $Im\ T =$ dim V gives dim $Ker\ T = 0$.

Hence T is one-to-one. T being both one-to-one and onto, T is invertible. This completes the proof.

Matrix of the inverse mapping.

Theorem 4.21.20. Let V and W be finite dimensional vector spaces of the same dimension over a field F. Let $T : V \to W$ be an invertible mapping. If $m(T)$ be the matrix of T relative to any chosen pair of ordered bases of V and W then the matrix of the inverse mapping T' : $W \to V$ relative to the same ordered bases, is given by

$$m(T') = [m(T)]^{-1}.$$

Proof. Let dim V=dim $W = n$. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\beta_1, \beta_2, \ldots, \beta_n)$ be the ordered bases of V and W respectively. Relative to the chosen bases, let the matrices of T and T' be $A = (a_{ij})_{nn}$, $B = (b_{ij})_{nn}$ respectively.

$$\begin{aligned} \text{Then } T(\alpha_j) &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n \text{ for } j = 1, 2, \ldots, n \\ T'(\beta_j) &= b_{1j}\alpha_1 + b_{2j}\alpha_2 + \cdots + b_{nj}\alpha_n \text{ for } j = 1, 2, \ldots, n. \end{aligned}$$

Since the mapping $T' : W \to V$ is the inverse of T, $T'T = I_V$ and $TT' = I_W$.

$$\begin{aligned} T'T(\alpha_j) &= T'[a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n] \\ &= a_{1j}T'(\beta_1) + a_{2j}T'(\beta_2) + \cdots + a_{nj}T'(\beta_n) \text{ as } T' \text{ is linear,} \\ &= a_{1j}[b_{11}\alpha_1 + b_{21}\alpha_2 + \cdots + b_{n1}\alpha_n] + a_{2j}[b_{12}\alpha_1 + b_{22}\alpha_2 + \cdots + b_{n2}\alpha_n] + \cdots + a_{nj}[b_{1n}\alpha_1 + b_{2n}\alpha_2 + \cdots + b_{nn}\alpha_n] \\ &= (b_{11}a_{1j} + b_{12}a_{2j} + \cdots + \cdots +)\alpha_1 + (b_{21}a_{1j} + b_{22}a_{2j} + \cdots + b_{2n}a_{nj})\alpha_2 + \cdots + (b_{n1}a_{1j} + b_{n2}a_{2j} + \cdots + b_{nn}a_{nj})\alpha_n \text{ for } j = 1, 2, \ldots, n. \end{aligned}$$

But $T'T = I_V$ implies $T'T(\alpha_j) = \alpha_j$. Therefore

$$\begin{aligned} b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{in}a_{nj} &= 1 \text{ if } i = j \\ &= 0 \text{ if } i \neq j. \end{aligned}$$

This holds for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$.

It follows that $BA = I_n$. ... ... (i)

$TT' = I_W$ implies $TT'(\beta_j) = \beta_j$. Therefore

$$\begin{aligned} a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} &= 1 \text{ if } i = j \\ &= 0 \text{ if } i \neq j. \end{aligned}$$

This holds for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$.

It follows that $AB = I_n$. ... ... (ii)

From (i) and (ii) we have $B = A^{-1}$.

Therefore $m(T') = [m(T)]^{-1}$, where $m(T)$ and $m(T')$ are the matrices of T and T' respectively relative to any chosen ordered bases, one of V and one of W.

This completes the proof.

Another proof.

Let dim $V =$ dim $W = n$. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\beta_1, \beta_2, \ldots, \beta_n)$ be ordered bases of V and W respectively.

Relative to the chosen bases, let the matrices of T and T' be A and B respectively.

Then the matrix of $T'T : V \to V$ is BA and the matrix of $TT' : W \to W$ is AB relative to the chosen bases of V and W.

Since the mapping $T' : W \to V$ is the inverse of T, $T'T = I_V$ and $TT' = I_W$.

Since $T'T = I_V$, where I_V is the identity mapping on V and the matrix of the identity mapping is I_n, it follows that $AB = I_n$. Since $TT' = I_W$, it follows that $BA = I_n$.

$AB = BA = I_n$ implies B is the inverse of A and therefore $m(T') = [m(T)]^{-1}$. This completes the proof.

Worked Example (continued).

4. Let $(\alpha_1, \alpha_2, \alpha_3), (\beta_1, \beta_2, \beta_3)$ be ordered bases of the real vector spaces V and W respectivly. A linear mapping $T : V \to W$ maps the basis vectors as $T(\alpha_1) = \beta_1, T(\alpha_2) = \beta_1 + \beta_2, T(\alpha_3) = \beta_1 + \beta_2 + \beta_3$.

Find the matrix of T relative to the ordered bases $(\alpha_1, \alpha_2, \alpha_3)$ of V and $(\beta_1, \beta_2, \beta_3)$ of W. Show that T is non-singular.

Find the matrix of T^{-1} relative to the same chosen ordered bases.

$T(\alpha_1) = 1\beta_1 + 0\beta_2 + 0\beta_3$;
$T(\alpha_2) = 1\beta_1 + 1\beta_2 + 0\beta_3$;
$T(\alpha_3) = 1\beta_1 + 1\beta_2 + 1\beta_3$.

Therefore $m(T)$ (the matrix of T) $= \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

$m(T)$ is non-singular and therefore T is non-singular.

T^{-1} exists and $T^{-1}(\beta_1) = \alpha_1$; $T^{-1}(\beta_1+\beta_2) = \alpha_2$; $T^{-1}(\beta_1+\beta_2+\beta_3) = \alpha_3$.

or, $T^{-1}(\beta_1) = \alpha_1$;
$T^{-1}(\beta_1) + T^{-1}(\beta_2) = \alpha_2$;
$T^{-1}(\beta_1) + T^{-1}(\beta_2) + T^{-1}(\beta_3) = \alpha_3$, since T^{-1} is linear.

$T^{-1}(\beta_1) = \alpha_1 = 1\alpha_1 + 0\alpha_2 + 0\alpha_3$;
$T^{-1}(\beta_2) = \alpha_2 - \alpha_1 = -1\alpha_1 + 1\alpha_2 + 0\alpha_3$;
$T^{-1}(\beta_3) = \alpha_3 - \alpha_2 = 0\alpha_1 - 1\alpha_2 + 1\alpha_3$.

Therefore $m(T^{-1})$ (the matrix of T^{-1}) $= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

Note. $m(T).m(T^{-1}) = m(T^{-1}).m(T) = I_3$ and therefore $m(T^{-1}) = [m(T)]^{-1}$.

4.22. Linear space of linear mappings.

Algebraic operations on the set of all linear mappings.

Let V and W be vector spaces over a field F. Let $T : V \to W$, $S : V \to W$ be two linear mappings. We define the sum $T + S$ and the scalar multiple cT, c being a scalar in F.

$T + S : V \to W$ is defined by $(T+S)(\alpha) = T(\alpha) + S(\alpha)$ for all $\alpha \in V$

$cT : V \to W$ is defined by $(cT)(\alpha) = c(T(\alpha))$ for all $\alpha \in V$.

To prove that $T + S$ is a linear mapping, let $\alpha, \beta \in V$.

Then $(T+S)(\alpha) = T(\alpha) + S(\alpha)$; $(T+S)(\beta) = T(\beta) + S(\beta)$.

$$\begin{aligned} (T+S)(\alpha+\beta) &= T(\alpha+\beta) + S(\alpha+\beta) \\ &= [T(\alpha) + T(\beta)] + [S(\alpha) + S(\beta)], \text{ since } T, S \text{ are linear} \\ &= [T(\alpha) + S(\alpha)] + [T(\beta) + S(\beta)] \\ &= (T+S)(\alpha) + (T+S)(\beta). \end{aligned}$$

For a scalar p, $(T+S)(p\alpha) = T(p\alpha) + S(p\alpha)$
$= pT(\alpha) + pS(\alpha)$, since T, S are linear
$= p[T(\alpha) + S(\alpha)]$
$= p(T+S)(\alpha)$.

Hence $T + S$ is linear.

$$\begin{aligned} \text{Again } (cT)(\alpha+\beta) &= c[T(\alpha+\beta)] \\ &= c[T(\alpha) + T(\beta)], \text{ since } T \text{ is linear} \\ &= cT(\alpha) + cT(\beta) \\ &= (cT)(\alpha) + (cT)(\beta). \end{aligned}$$

$$\begin{aligned} \text{For a scalar } p, (cT)(p\alpha) &= c[T(p\alpha)] \\ &= c[pT(\alpha)], \text{ since } T \text{ is linear} \\ &= cpT(\alpha) \\ &= p(cT)(\alpha). \end{aligned}$$

Hence cT is linear.

Let L be the set of all linear mappings with domain V and co-domain W, both having the same scalar field F. Then L is a linear space over the field F, if addition and scalar multiplication be defined as follows –

for $T \in L, S \in L$, $(T+S)(\alpha) = T(\alpha) + S(\alpha)$ for all $\alpha \in V$,

for $T \in L, c \in F$, $(cT)(\alpha) = cT(\alpha)$ for all $\alpha \in V$.

It is a matter of simple verification that

$T + S = S + T$ for all $T, S \in L$;
$T + (S + U) = (T + S) + U$ for all $T, S, U \in L$;
$T + O = T$ for all $T \in L$, O being the zero mapping;
for each $T \in L$, there is a mapping $-T$ such that $T + (-T) = O$,
$-T : V \to W$ is defined by $(-T)(\alpha) = -[T(\alpha)]$ for all $\alpha \in V$.

$c(dT) = (cd)T$ for all $c, d \in F$, all $T \in L$;
$c(T + S) = cT + cS$ for all $c \in F$, all $T, S \in L$;
$(c + d)T = cT + dT$ for all $c, d \in F$, all $T \in L$;
$1T = T$, 1 being the identity element in F.

Therefore L is a linear space.

The linear space of all linear mappings with domain V and co-domain W is denoted by $L(V, W)$.

Since a linear mapping $T : V \to W$ is also a homomorphism of V into W, the linear space $L(V, W)$ is also denoted by $\text{Hom}(V, W)$.

Two particular cases of the linear space $L(V, W)$ are of profound interest. The first one is $L(V, V)$ and the second one is $L(V, F)$, where F is a vector space over F itself.

The linear space $L(V, F)$ is said to be the *dual space* of V.

Isomorphism between linear mappings and matrices.

Theorem 4.22.1. Let V and W be finite dimensional vector spaces over a field F with $\dim V = n$ and $\dim W = m$. Then $L(V, W)$ and $M_{m,n}$ are isomorphic, where $L(V, W)$ is the linear space over F of all linear mappings of V to W and $M_{m,n}$ is the vector space of all $m \times n$ matrices over F.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\beta_1, \beta_2, \ldots, \beta_m)$ be ordered bases of V and W respectively.

Let us define a mapping $m : L(V, W) \to M_{m,n}$ by $m(T) = (a_{ij})_{m,n}$ for $T \in L(V, W)$, $(a_{ij})_{m,n}$ being the matrix of T relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V and $(\beta_1, \beta_2, \ldots, \beta_m)$ of W.

Let $T \in L(V, W)$, $S \in L(V, W)$. Then $T + S \in L(V, W)$.

Let $m(T) = (a_{ij})_{m,n}$, $m(S) = (b_{ij})_{m,n}$, $m(T + S) = (c_{ij})_{m,n}$.

$$\begin{aligned} \text{Then } T(\alpha_j) &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m \text{ for } j = 1, 2, \ldots, n; \\ S(\alpha_j) &= b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{mj}\beta_m \text{ for } j = 1, 2, \ldots, n; \\ (T + S)(\alpha_j) &= c_{1j}\beta_1 + c_{2j}\beta_2 + \cdots + c_{mj}\beta_m \text{ for } j = 1, 2, \ldots, n. \end{aligned}$$

Since T and S are linear, $(T + S)(\alpha_j) = T(\alpha_j) + S(\alpha_j)$.

Therefore $c_{1j}\beta_1 + c_{2j}\beta_2 + \cdots + c_{mj}\beta_m = (a_{1j} + b_{1j})\beta_1 + (a_{2j} + b_{2j})\beta_2 + \cdots + (a_{mj} + b_{mj})\beta_m$.

Since the set $\{\beta_1, \beta_2, \ldots, \beta_m\}$ is linearly independent, we have $c_{ij} = a_{ij} + b_{ij}$ for $i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$.

It follows that $m(T + S) = m(T) + m(S)$ ... (i)

Let p be a scalar in F. Then $pT \in L(V, W)$.

Let $m(pT) = (d_{ij})_{m,n}$.

$$\begin{aligned} \text{Then } (pT)(\alpha_j) &= d_{1j}\beta_1 + d_{2j}\beta_2 + \cdots + d_{mj}\beta_m \text{ for } j = 1, 2, \ldots, n. \\ \text{Again } (pT)(\alpha_j) &= pT(\alpha_j), \text{ by the definition of } pT \\ &= p[a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m] \\ &= (pa_{1j})\beta_1 + (pa_{2j})\beta_2 + \cdots + (pa_{mj})\beta_m. \end{aligned}$$

Therefore $d_{ij} = pa_{ij}$ for $i = 1, 2, \ldots, m$, since $\{\beta_1, \beta_2, \ldots, \beta_m\}$ is linearly independent.

Consequently, $d_{ij} = pa_{ij}$ for $i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$.

It follows that $m(pT) = pm(T)$ ... (ii)

From (i) and (ii) m is a homomorphism.

To prove that m is an isomorphism, let $m(T) = m(S)$ for some $T, S \in L(V, W)$.

Let $m(T) = (a_{ij})_{m,n}, m(S) = (b_{ij})_{m,n}$

$$\begin{aligned} \text{Then } T(\alpha_j) &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m \text{ for } j = 1, 2, \ldots, n; \\ S(\alpha_j) &= b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{mj}\beta_m \text{ for } j = 1, 2, \ldots, n. \end{aligned}$$

$m(T) = m(S) \Rightarrow a_{ij} = b_{ij}$ for all i, j.

Therefore $T(\alpha_j) = S(\alpha_j)$ for $j = 1, 2, \ldots, n$.

Let γ be an arbitrary vector in V.

Since $T(\alpha_j) = S(\alpha_j)$ for all basis vectors $\alpha_j, T(\gamma) = S(\gamma)$ for all $\gamma \in V$ and this implies $T = S$.

Therefore $m(T) = m(S) \Rightarrow T = S$, proving that m is one-to-one.

To prove that m is onto, let $(a_{ij})_{m,n} \in M_{m,n}$. Then there exists a unique linear mapping $T : V \to W$ whose matrix is (a_{ij}), because if we prescribe the jth column of (a_{ij}) as the co-ordinates of $T(\alpha_j)$ relative to $(\beta_1, \beta_2, \ldots, \beta_m)$, i.e., $T(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m$ then T is determined uniquely with (a_{ij}) as the associated matrix.

Thus m is an isomorphism and therefore the linear spaces $L(V, W)$ and $M_{m,n}$ are isomorphic.

This completes the proof.

Dimension of the vector space $L(V, W)$.

Theorem 4.22.2. Let V and W be finite dimensional vector spaces over a field F with $\dim V = n$ and $\dim W = m$. Then the dimension of the linear space $L(V, W)$ is mn.

Proof. Since the linear spaces $L(V, W)$ and $M_{m,n}$ are isomorphic, the dimension of $L(V, W)$ = the dimension of the linear space $M_{m,n}$.

Let E_{ij} be the $m \times n$ matrix over F in which the ijth element is 1 and every element except the ijth element is zero. Then the set of mn elements $\{E_{11}, E_{12}, \ldots, E_{1n}, E_{21}, E_{22}, \ldots, E_{2n}, \ldots, E_{m1}, E_{m2}, \ldots, E_{mn}\}$ is linearly independent. Also each matrix $(a_{ij})_{m,n} \in M_{m,n}$ is a linear combination of the matrices in the set $\{E_{ij}\}$.

Therefore the set $\{E_{ij}\}$ is a basis of $M_{m,n}$ and the dimension of $M_{m,n}$ is mn.

So the dimension of $L(V, W) = mn$.

Worked Examples.

1. Let V, W, U be a finite dimensional vector spaces over a field F and $T_1 : V \to W$, $T_2 : W \to U$ are linear mappings. Prove that

$$\text{rank of } T_2T_1 \leq \min \{\text{rank of } T_2, \text{ rank of } T_1\}.$$

$T_2T_1 : V \to U$ is a linear mapping. We prove that $Im\ T_2T_1 \subset Im\ T_2$ and $Ker\ T_1 \subset Ker\ T_2T_1$.

Let $\gamma \in Im\ T_2T_1$. Then $T_2T_1(\alpha) = \gamma$ for some $\alpha \in V$. $T_2[T_1(\alpha)] = \gamma \Rightarrow \gamma \in Im\ T_2$. Thus $\gamma \in Im\ T_2T_1 \Rightarrow \gamma \in Im\ T_2$. So $Im\ T_2T_1 \subset Im\ T_2$.

Let $\alpha \in Ker\ T_1$. Then $T_1(\alpha) = \theta \in W$. Therefore $T_2T_1(\alpha) = \theta \in U$. This shows that $\alpha \in Ker\ T_2T_1$. Thus $\alpha \in Ker\ T_1 \Rightarrow \alpha \in Ker\ T_2T_1$. So $Ker\ T_1 \subset Ker\ T_2T_1$.

$Im\ T_2T_1 \subset Im\ T_2$ implies $\dim Im\ T_2T_1 \leq \dim Im\ T_2$, i.e., rank of $T_2T_1 \leq$ rank of T_2 ... (i)

$Ker\ T_1 \subset Ker\ T_2T_1$ implies $\dim Ker\ T_1 \leq \dim Ker\ T_2T_1$.

Because $\dim Ker\ T_1 + \dim Im\ T_1 = \dim Ker\ T_2T_1 + \dim Im\ T_2T_1 = \dim V$, $\dim Ker\ T_1 \leq \dim Ker\ T_2T_1$ implies $\dim Im\ T_2T_1 \leq \dim Im\ T_1$, i.e., rank of $T_2T_1 \leq$ rank of T_1 ... (ii)

From (i) and (ii) rank of $T_2T_1 \leq \min \{\text{rank of } T_2, \text{ rank of } T_1\}$.

2. Let V and W be finite dimensional vector spaces over a field F and $T_1 : V \to W$, $T_2 : V \to W$ are linear mappings. Prove that

$$\text{rank of } (T_1 + T_2) \leq \text{rank of } T_1 + \text{rank of } T_2.$$

Let $\alpha \in Im\ (T_1 + T_2)$. Then $\alpha = (T_1 + T_2)\beta$ for some $\beta \in V$
$$= T_1(\beta) + T_2(\beta).$$

But $T_1(\beta) \in Im\ T_1$, $T_2(\beta) \in Im\ T_2$.

Therefore $\alpha = T_1(\beta) + T_2(\beta) \in (Im\ T_1 + Im\ T_2)$, the linear sum of the subspaces $Im\ T_1$ and $Im\ T_2$.

Hence $Im\ (T_1 + T_2) \subset Im\ T_1 + Im\ T_2$ and Therefore
$\dim Im\ (T_1 + T_2) \leq \dim\ (Im\ T_1 + Im\ T_2)$ ... (i)

$Im\ T_1, Im\ T_2$ being subspaces of a finite dimensional vector space W, $\dim\ (Im\ T_1 + Im_2) = \dim Im\ T_1 + \dim Im\ T_2 - \dim\ [Im\ T_1 \cap Im\ T_2]$.

Therefore $\dim\ (Im\ T_1 + Im\ T_2) \leq \dim Im\ T_1 + \dim Im\ T_2$ ... (ii)

From (i) and (ii) $\dim Im\ (T_1 + T_2) \leq \dim Im\ T_1 + \dim Im\ T_2$. That is, rank of $(T_1 + T_2) \leq$ rank of $T_1 +$ rank of T_2.

Corollary. If $T_1 \in L(V, W)$ and $T_2 \in L(V, W)$, where V and W are finite dimensional vector spaces over a field F and $\dim V = n$ then

$$\text{nullity of } (T_1 + T_2) \geq \text{nullity of } T_1 + \text{nullity of } T_2 - n.$$

We have seen that the matrix $m(T)$ of a linear map $T \in L(V, W)$ depends on the choice of a pair of ordered bases of V and W.

How different matrices of a linear mapping $T \in L(V, W)$ relative to different pairs of ordered bases of V and W are related to one another is discussed in the following theorem.

Theorem 4.22.3. Let V and W be finite dimensional vector spaces over a field F and $\dim V = n$, $\dim W = m$. Let $T : V \to W$ be a linear mapping. Let A be the matrix of T relative to a pair of ordered bases of V and W and C be the matrix of T relative to a different pair of ordered bases of V and W. Then there exist non-singular matrices P and Q such that $C = P^{-1}AQ$.

Proof. Let A be the matrix of T relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V and $(\beta_1, \beta_2, \ldots, \beta_m)$ of W.

Let C be the matrix of T relative to the ordered bases $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ of V and $(\delta_1, \delta_2, \ldots, \delta_m)$ of W.

Let $A = (a_{ij})_{m,n}$, $C = (c_{ij})_{m,n}$.

Then $T(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m$ for $j = 1, 2, \ldots, n$
$T(\gamma_j) = c_{1j}\delta_1 + c_{2j}\delta_2 + \cdots + c_{mj}\delta_m$ for $j = 1, 2, \ldots, n$.

Let $T_1 : V \to V$ be such that $T_1(\alpha_i) = \gamma_i, i = 1, 2, \ldots, n$ and Q be the matrix of T_1 relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V. Since T_1 maps a basis of V onto another basis, T_1 is non-singular and therefore Q is non-singular.

Let $Q = (q_{ij})_{n,n}$. Then $\gamma_j = T_1(\alpha_j) = q_{1j}\alpha_1 + q_{2j}\alpha_2 + \cdots + q_{nj}\alpha_n$ for $j = 1, 2, \ldots, n$.

Let $T_2 : W \to W$ be such that $T_2(\beta_i) = \delta_i, i = 1, 2, \ldots, m$ and P be the matrix of T_2 relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_m)$ of W. Since T_2 maps a basis of W onto another basis, T_2 is non-singular and therefore P is non-singular.

Let $P = (p_{ij})_{m,m}$. Then $\delta_j = T_2(\beta_j) = p_{1j}\beta_1 + p_{2j}\beta_2 + \cdots + p_{mj}\beta_m$ for $j = 1, 2, \ldots, m$.

$$\begin{aligned} T(\gamma_j) &= c_{1j}\delta_1 + c_{2j}\delta_2 + \cdots + c_{mj}\delta_m \\ &= c_{1j}[p_{11}\beta_1 + p_{21}\beta_2 + \cdots + p_{m1}\beta_m] + c_{2j}[p_{12}\beta_1 + \\ &\quad + \cdots + p_{m2}\beta_m] + \cdots + c_{mj}[p_{1m}\beta_1 + \cdots + p_{mm}\beta_m] \\ &= (\sum_{k=1}^{m} p_{1k}c_{kj})\beta_1 + \cdots + (\sum_{k=1}^{m} p_{mk}c_{kj})\beta_m, \end{aligned}$$

and $T(\gamma_j) = T[q_{1j}\alpha_1 + q_{2j}\alpha_2 + \cdots + q_{nj}\alpha_n]$
$= q_{1j}T(\alpha_1) + q_{2j}T(\alpha_2) + \cdots + q_{nj}T(\alpha_n)$
$= q_{1j}[a_{11}\beta_1 + a_{21}\beta_2 + \cdots + a_{m1}\beta_m] + q_{2j}[a_{12}\beta_1 + \cdots + a_{m2}\beta_m] + \cdots + q_{nj}[a_{1n}\beta_1 + \cdots + a_{mn}\beta_m]$
$= (\sum_{k=1}^{n} a_{1k}q_{kj})\beta_1 + \cdots + (\sum_{k=1}^{n} a_{mk}q_{kj})\beta_m.$

Since $(\beta_1, \beta_2, \ldots, \beta_m)$ is a basis, we have
$\sum_{k=1}^{m} p_{ik}c_{kj} = \sum_{k=1}^{n} a_{ik}q_{kj}$ for $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$.

This gives $PC = AQ$. Since P is non-singular, $C = P^{-1}AQ$.

This completes the proof.

Note. The matrix C is equivalent to A. Thus a linear mapping $T \in L(V, W)$ has different matrices relative to different pairs of ordered bases of V and W but all such matrices are equivalent matrices.

We now come to the converse problem.

Let V and W be finite dimensional vector spaces over a field F with dim $V = n$ and dim $W = m$. Given two equivalent matrices A and C of order $m \times n$, does there exist a linear map $T \in L(V, W)$ such that A and C are matrices of T relative to two different pairs of ordered bases of V and W? This is discussed in the following theorem.

Theorem 4.22.4. Let V and W be finite dimensional vector spaces over a field F with dim $V = n$ and dim $W = m$ and A, C are $m \times n$ matrices over F such that $C = P^{-1}AQ$ for some non-singular matrices P and Q. Then there exists a linear map $T \in L(V, W)$ such that A and C are matrices of T relative to different pairs of ordered bases of V and W.

Proof. Let $A = (a_{ij})_{m,n}$; $C = (c_{ij})_{m,n}$; $P = (p_{ij})_{m,m}$; $Q = (q_{ij})_{n,n}$.

Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\beta_1, \beta_2, \ldots, \beta_m)$ be a pair of ordered bases of V and W respectively and let $T : V \to W$ be defined by $T(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n$, $j = 1, 2, \ldots, n$.

Then T is a uniquely determined linear mapping and the matrix of T relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $(\beta_1, \beta_2, \ldots, \beta_m)$, is A.

Let $T_1 : V \to V$ be defined by $T_1(\alpha_j) = q_{1j}\alpha_1 + q_{2j}\alpha_2 + \cdots + q_{nj}\alpha_n$.

Then T_1 is a uniquely determined mapping on V. Since Q is a non-singular matrix, $\{T_1(\alpha_1), T_1(\alpha_2), \ldots, T_1(\alpha_n)\}$ is a basis of V.

Let $T_1(\alpha_i) = \gamma_i$, $i = 1, 2, \ldots, n$.

Let $T_2 : W \to W$ be defined by $T_2(\beta_j) = p_{1j}\beta_1 + p_{2j}\beta_2 + \cdots + p_{mj}\beta_m$.

Then T_2 is a uniquely determined mapping on W. Since P is a non-singular matrix, $\{T_2(\beta_1), T_2(\beta_2), \ldots, T_2(\beta_m)\}$ is a basis of W.

Let $T_2(\beta_i) = \delta_i$, $i = 1, 2, \ldots, m$.

Let T' be the linear mapping $\in L(V, W)$ whose matrix relative to the ordered bases $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ of V and $(\delta_1, \delta_2, \ldots, \delta_m)$ of W be C.

Then $T'(\gamma_j) = c_{1j}\delta_1 + c_{2j}\delta_2 + \cdots + c_{mj}\delta_m$
$= c_{1j}[p_{11}\beta_1 + p_{21}\beta_2 + \cdots + p_{m1}\beta_m] + c_{2j}[p_{12}\beta_1 + \cdots + p_{m2}\beta_m] + \cdots + c_{mj}[p_{1m}\beta_1 + \cdots + p_{mm}\beta_m]$
$= (\sum_{k=1}^{m} p_{1k}c_{kj})\beta_1 + \cdots + (\sum_{k=1}^{m} p_{mk}c_{kj})\beta_m,$

and $T(\gamma_j) = T[q_{1j}\alpha_1 + q_{2j}\alpha_2 + \cdots + q_{nj}\alpha_n]$
$= q_{1j}T(\alpha_1) + q_{2j}T(\alpha_2) + \cdots + q_{nj}T(\alpha_n)$
$= q_{1j}[a_{11}\beta_1 + \cdots + a_{m1}\beta_m] + q_{2j}[a_{12}\beta_1 + \cdots + a_{m2}\beta_m] + \cdots + q_{nj}[a_{1n}\beta_1 + \cdots + a_{mn}\beta_m]$
$= (\sum_{k=1}^{n} a_{1k}q_{kj})\beta_1 + \cdots + (\sum_{k=1}^{n} a_{mk}q_{kj})\beta_m.$

$C = P^{-1}AQ \Rightarrow PC = AQ$
$\Rightarrow \sum_{k=1}^{m} p_{ik}c_{kj} = \sum_{k=1}^{n} a_{ik}q_{kj}$ for $j = 1, 2, \ldots, n$.

So $T'(\gamma_j) = T(\gamma_j)$ for $j = 1, 2, \ldots, n$ and this proves $T = T'$.

Thus A is the matrix of T relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V and $(\beta_1, \beta_2, \ldots, \beta_m)$ of W; C is the matrix of T relative to the ordered bases $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ of V and $(\delta_1, \delta_2, \ldots, \delta_m)$ of W.

P is the matrix of the non-singular mapping $T_2 \in L(W, W)$ that maps β_i to δ_i; Q is the matrix of the non-singular mapping $T_1 \in L(V, V)$ that maps α_i to γ_i.

Let us reverse the roles of matrices and linear mappings in $L(V, W)$. We have the following theorems.

Theorem 4.22.5. Let V and W be finite dimensional vector spaces over a field F with dim $V = n$ and dim $W = m$ and A be an $m \times n$ matrix over F. Relative to two different pairs of ordered bases of V and W, let A determines two linear maps T_1 and T_2 such that $m(T_1) = A, m(T_2) = A$. Then there exist non-singular linear maps $S \in L(V, V)$ and $R \in L(W, W)$ such that $T_1 = R^{-1}T_2S$.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\gamma_1, \gamma_2, \ldots, \gamma_n)$ be a pair of ordered bases of V and let $S : V \to V$ be defined by $S(\alpha_i) = \gamma_i$, $i = 1, 2, \ldots, n$.

Let $(\beta_1, \beta_2, \ldots, \beta_m)$, $(\delta_1, \delta_2, \ldots, \delta_m)$ be a pair of ordered bases of W and let $R : W \to W$ be defined by $R(\beta_i) = \delta_i$, $i = 1, 2, \ldots, m$.

Then S and R are non-singular. Let $A = (a_{ij})_{m,n}$.

Let $T_1(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m$, $j = 1, 2, \ldots, n$

and $T_2(\gamma_j) = a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m$, $j = 1, 2, \ldots, n$.

Then $T_2S(\alpha_j) = T_2(\gamma_j) = a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m$,

$$\begin{aligned} RT_1(\alpha_j) &= R(a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m) \\ &= a_{1j}R(\beta_1) + a_{2j}R(\beta_2) + \cdots + a_{mj}\beta_m \\ &= a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m. \end{aligned}$$

Since $T_2S(\alpha_j) = RT_1(\alpha_j)$ for $j = 1, 2, \ldots, n$, $T_2S = RT_1$ and therefore $T_1 = R^{-1}T_2S$.

This completes the proof.

The converse problem is discussed in the following theorem.

Theorem 4.22.6. Let V and W be finite dimensional vector spaces over a field F with $\dim V = n$ and $\dim W = m$ and let T_1, T_2 be two linear maps in $L(V, W)$ such that $T_1 = R^{-1}T_2S$, where R is a non-singular mapping in $L(W, W)$, S is a non-singular mapping in $L(V, V)$. Then $m(T_1) = m(T_2)$ relative to different pairs of ordered bases of V and W.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $(\beta_1, \beta_2, \ldots, \beta_m)$ be a pair of ordered bases of V and W respectively. Let $S(\alpha_i) = \gamma_i$ and $R(\beta_i) = \delta_i$. Since R and S are non-singular, $(S(\alpha_1), S(\alpha_2), \ldots, S(\alpha_n))$ is a basis of V and $(R(\beta_1), R(\beta_2), \ldots, R(\beta_n))$ is a basis of W.

Let $A = (a_{ij})_{m,n}$ be the matrix of T_1 relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V and $(\beta_1, \beta_2, \ldots, \beta_m)$ of W.

Then $T_1(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{mj}\beta_m$,

$$\begin{aligned} RT_1(\alpha_i) &= a_{1j}R(\beta_1) + a_{2j}R(\beta_2) + \cdots + a_{mj}R(\beta_m) \\ &= a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m. \end{aligned}$$

Since $RT_1 = T_2S$, we have

$T_2S(\alpha_j) = a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m$

or, $T_2(\gamma_j) = a_{1j}\delta_1 + a_{2j}\delta_2 + \cdots + a_{mj}\delta_m$.

This shows that A is the matrix of T_2 relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V and $(\beta_1, \beta_2, \ldots, \beta_m)$ of W.

This completes the proof.

Worked Example (continued).

3. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by $T(x, y, z) = (x + y, y + z)$, $(x, y, z) \in \mathbb{R}^3$.

Let the matrix of T relative to the ordered bases $((1,0,0), (0,1,0), (0,0,1))$ of $\mathbb{R}^3$ and $((1,0), (0,1))$ of $\mathbb{R}^2$ be A and the matrix of T relative to the ordered bases $((0,1,1), (1,0,1), (1,1,0))$ of $\mathbb{R}^3$ and $((1,2), (2,1))$ of $\mathbb{R}^2$ be B.

Let the linear mapping $S : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $S(1,0,0) = (0,1,1)$, $S(0,1,0) = (1,0,1)$, $S(0,0,1) = (1,1,0)$ and the linear mapping $S' : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $S'(1,0) = (1,2), S'(0,1) = (2,1)$.

Let the matrix of S relative to the ordered bases $((1,0,0), (0,1,0), (0,0,1))$ of $\mathbb{R}^3$ be Q and the matrix of S' relative to the ordered bases $((1,0), (0,1))$ of $\mathbb{R}^2$ be P.

Show that $B = P^{-1}AQ$.

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

Therefore $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

$T(0,1,1) = (1,2) = 1(1,2) + 0(2,1), T(1,0,1) = (1,1) = \frac{1}{3}(1,2) + \frac{1}{3}(2,1), T(1,1,0) = (2,1) = 0(1,2) + 1(2,1)$.

Therefore $B = \begin{pmatrix} 1 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix}$.

$S(1,0,0) = (0,1,1) = 0(1,0,0) + 1(0,1,0) + 1(0,0,1), T(0,1,0) = (1,0,1) = 1(1,0,0) + 0(0,1,0) + 1(0,0,1), T(0,0,1) = (1,1,0) = 1(1,0,0) + 1(0,1,0) + 0(0,0,1)$.

Therefore $Q = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

$S'(1,0) = (1,2) = 1(1,0) + 2(0,1), S'(0,1) = (2,1) = 2(1,0) + 1(0,1)$.

Therefore $P = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$.

$$AQ = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}.$$

$$PB = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}.$$

$AQ = PB$ and therefore $B = P^{-1}AQ$.

4.23. Linear operators.

So far we have discussed some properties of linear mappings from V to W, where V and W are vector spaces over a field F.

Now we shall consider the special case when $W = V$.

A linear mapping $T : V \to V$ is called a *linear operator on* V.

The set of all linear operators on a vector space V over a field F form, in its own right, a linear space over F, denoted by $L(V, V)$.

The important feature in this case is that we can define another binary composition, called multiplication, on this set.

Let T and S be two linear operators on V. Then the composite mappings $T \circ S$ and $S \circ T$ are both linear operators on V.

We define $S.T$ by $S \circ T$.

Then $S.T : V \to V$ is defined by $S.T(\alpha) = S[T(\alpha)]$ for all $\alpha \in V$.

Since the composition of linear mappings is associative, multiplication is associative.

Hence $(S.T).U = S.(T.U)$ for all $S, T, U \in L(V, V)$.

The mapping $I_V : V \to V$ defined by $I_V(\alpha) = \alpha$ for all $\alpha \in V$ is the identity operator.

Multiplication is related to addition by the laws–

(i) $T.(S + U) = (T.S) + (T.U)$; (ii) $(S + U).T = (S.T) + (U.T)$.

To prove (i), $[T.(S + U)]\alpha = T[(S + U)\alpha] = T[S\alpha + U\alpha]$
$= T(S\alpha) + T(U\alpha)$ [T is linear]
$= (TS + TU)\alpha$, for all $\alpha \in V$.

Therefore $T.(S + U) = (T.S) + (T.U)$.

Thus the linear space $L(V, V)$ is a *ring* under addition and multiplication. It is a non-commutative ring with unity, I_V being the unity in the ring.

Definition. A non-empty set A is said to be an *algebra* over a field F if A is a vector space over F and multiplication of vectors can be defined in such a way that A is also a ring in which scalar multiplication satisfies the condition–

$$c(\alpha.\beta) = (c\alpha).\beta = \alpha.(c\beta) \text{ for all } c \in F \text{ and all } \alpha, \beta \in A.$$

An algebra is an algebraic system which is a vector space as well as a ring. An algebra in which multiplication of vectors is commutative is called a *commutative algebra*.

Examples.

1. The set $\mathbb{C}$ of all complex numbers is a vector space over $\mathbb{R}$ and $\mathbb{C}$ is commutative ring. Scalar multiplication (multiplication by reals) satisfies the condition–

$$c(x.y) = (cx).y = x.(cy) \text{ for all } c \in \mathbb{R}, \text{ all } x, y \in \mathbb{C}.$$

It is a commutative algebra over $\mathbb{R}$.

2. The set $M_n(\mathbb{R})$ of all $n \times n$ real matrices is a vector space over $\mathbb{R}$ and this set forms a non-commutative ring under matrix addition and matrix multiplication. Scalar multiplication (multiplication by reals) satisfies the condition–

$$c(A.B) = (cA).B = A.(cB) \text{ for all } c \in \mathbb{R}, \text{ all } A, B \in M_n(\mathbb{R}).$$

It is a non-commutative algebra over $\mathbb{R}$.

3. The set $C[a, b]$ of all real valued continuous functions on the closed interval $[a, b]$ is a vector space over $\mathbb{R}$ and this set forms a commutative ring under addition and multiplication of functions. Scalar multiplication (multiplication by reals) satisfies the condition–

$$c(f.g) = (cf).g = f.(cg) \text{ for all } c \in \mathbb{R}, \text{ all } f, g \in C[a, b].$$

It is a commutative algebra over $\mathbb{R}$.

We have seen that the linear space $L(V, V)$ is a non-commutative ring under addition and multiplication of linear mappings.

If $T \in L(V, V)$ then $cT \in L(V, V)$ for all $c \in F$ and it is a matter of simple verification that $c(S.T) = (cS).T = S.(cT)$ for all $c \in F$, all $S, T \in L(V, V)$.

Theefore $L(V, V)$ is a non-commutative algebra over F.

Theorem 4.23.1. Let T be a linear operator on V, where V is a vector space of dimension n over a field F. Then the following statements are equivalent.

(i) T is non-singular;

(ii) T is one-to-one;

(iii) T is onto;

(iv) T maps a linearly independent set of V to another linearly independent set;

(v) T maps a basis of V to another basis.

Proof. We shall prove that (i) implies (ii), (ii) implies (iii), (iii) implies (iv), (iv) implies (v) and (v) implies (i).

Let (i) hold. Then T is invertible. Therefore T is one-to-one and onto. Hence (ii) holds.

Let (ii) hold. Then $\dim Ker\, T = 0$.
The relation $\dim Ker\, T + \dim Im\, T = \dim V$ gives $\dim Im\, T = n$.
But $Im\, T \subset V$ and $\dim V = n$.
Hence $Im\, T = V$ and this proves that T is onto. Hence (iii) holds.

Let (iii) hold. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a linearly independent set of V.

Since T is onto, $\dim Im\, T = \dim V$. Therefore $\dim Ker\, T = 0$ and consequently, $Ker\, T = \{\theta\}$.

Since $Ker\, T = \{\theta\}$, the images of a linearly independent set in V are linearly independent, by Theorem 4.21.4.

Therefore the set $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_r)\}$ is linearly independent in V. Hence (iv) holds.

Let (iv) hold. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. Then by (iv), $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_r)\}$ is linearly independent in V.

Since $\dim V = n$, the set $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ being a linearly independent set of n elements, is a basis of V. Hence (v) holds.

Let (v) hold. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V. Then $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is another basis of V.

Let $\xi \in Ker\, T$. Then $T(\xi) = \theta$. Let $\xi = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n$.
Since T is linear, $c_1T(\alpha_1) + c_2T(\alpha_2) + \cdots + c_nT(\alpha_n) = \theta$.

This implies $c_1 = c_2 = \cdots = c_n = 0$, since $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is a linearly independent set.
Therefore $\xi \in Ker\, T \Rightarrow \xi = \theta$ and consequently, $Ker\, T = \{\theta\}$.
Therefore T is one-to-one.

$\dim Ker\, T + \dim Im\, T = \dim V$ gives $\dim Im\, T = \dim V$.
But $Im\, T \subset V$. So $Im\, T = V$ and this implies that T is onto.

T being both one-to-one and onto, T is invertible and therefore non-singular. Hence (i) holds.

This completes the proof.

Matrix representation of a linear operator.

Let V be a vector space of dimension n over a field F and $T : V \to V$ is a linear operator on V. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V. T is completely determined by the images $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$.

Each $T(\alpha_i)$ is a linear combination of the vectors $\alpha_1, \alpha_2, \ldots, \alpha_n$.

$$\begin{aligned} \text{Let } T(\alpha_1) &= a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n \\ T(\alpha_2) &= a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n \\ \cdots & \quad \cdots \\ T(\alpha_n) &= a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n, \end{aligned}$$

where a_{ij} are unique scalars in F determined by the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Proceeding with similar arguments as in the case of linear mappings, the matrix representation of the linear operator T is

$$Y = AX$$

where X is the co-ordinate vector of any arbitrary element ξ in V relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and Y is the co-ordinate vector of $T(\xi)$ relative to the same basis.

Relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$, the co-ordinate vectors of $T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)$ are $\begin{pmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ a_{n1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ a_{n2} \end{pmatrix}, \ldots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ a_{nn} \end{pmatrix}$ respectively.

The matrix $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & & \cdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$ is said to be the matrix of T relative to the basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and is denoted by $m(T)$.

Worked Example.

1. The matrix of a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ with respect to the ordered basis $((0,1,1), (1,0,1), (1,1,0))$ of $\mathbb{R}^3$ is given by

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

Find T. Find the matrix of T relative to the ordered basis $((2,1,1), (1,2,1), (1,1,2))$ of $\mathbb{R}^3$.

$T(0,1,1) = 0(0,1,1) + 2(1,0,1) + 2(1,1,0) = (4,2,2)$;
$T(1,0,1) = 3(0,1,1) + 3(1,0,1) - (1,1,0) = (2,2,6)$;
$T(1,1,0) = 0(0,1,1) - 2(1,0,1) + 2(1,1,0) = (0,2,-2)$.

Let $(x,y,z) \in \mathbb{R}^3$ and let $(x,y,z) = c_1(0,1,1) + c_2(1,0,1) + c_3(1,1,0)$.

where $c_i \in \mathbb{R}$. Then $c_1 = \frac{-x+y+z}{2}$, $c_2 = \frac{x-y+z}{2}$, $c_3 = \frac{x+y-z}{2}$.

$$\begin{aligned} T(x,y,z) &= c_1T(0,1,1) + c_2T(1,0,1) + c_3T(1,1,0) \\ &= c_1(4,2,2) + c_2(2,2,6) + c_3(0,2,-2) \\ &= (-x+y+3z, x+y+z, x-3y+5z). \end{aligned}$$

Therefore $T(2,1,1) = (2,4,4), T(1,2,1) = (4,4,0), T(1,1,2) = (6,4,8)$.
Let $(2,4,4) = c_1(2,1,1) + c_2(1,2,1) + c_3(1,1,2)$.
Then $c_1 + c_2 + c_3 = \frac{5}{2}$ and therefore $c_1 = -\frac{1}{2}, c_2 = \frac{3}{2}, c_3 = \frac{3}{2}$.
Let $(4,4,0) = c_1(2,1,1) + c_2(1,2,1) + c_3(1,1,2)$.
Then $c_1 + c_2 + c_3 = 2$ and therefore $c_1 = 2, c_2 = 2, c_3 = -2$.
Let $(6,4,8) = c_1(2,1,1) + c_2(1,2,1) + c_3(1,1,2)$.
Then $c_1 + c_2 + c_3 = \frac{9}{2}$ and therefore $c_1 = \frac{3}{2}, c_2 = -\frac{1}{2}, c_3 = \frac{7}{2}$.

$$m(T) = \begin{pmatrix} -\frac{1}{2} & 2 & \frac{3}{2} \\ \frac{3}{2} & 2 & -\frac{1}{2} \\ \frac{3}{2} & -2 & \frac{7}{2} \end{pmatrix}.$$

As in the case of linear mappings of V to W where V and W are finite dimensional vector spaces over a field F, the following theorems relating to the linear operators on a vector space V, hold.

1. If $\dim V = n$, then the linear space $L(V,V)$ is isomorphic to the linear space $M_{n,n}$ of all $n \times n$ matrices over F.

2. If $\dim V = n$, the dimension of the linear space $L(V,V)$ in n^2.

Theorem 4.23.2. Let V be a vector space of dimension n over a field F. Then for a linear operator $T \in L(V,V)$,

rank of T = rank of $m(T)$,

where $m(T)$ is the matrix of T relative to any ordered basis of V.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V and let $A = (a_{ij})_{n,n}$ be the matrix of T relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Then $T(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$ for $j = 1, 2, \ldots, n$.

Let us consider the isomorphism $\phi : V \to F^n$ defined by

$$\phi(c_1\alpha_1 + c_2\alpha_2 + \cdots + c_n\alpha_n) = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}.$$

Since $\phi T : V \to F^n$ is a linear mapping and $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a basis of V, $\phi T(\alpha_1), \phi T(\alpha_2), \ldots, \phi T(\alpha_n)$ generate $Im\ \phi T$.

But $\phi T(\alpha_j)$ is the j-th column vector of the matrix A. Therefore $Im\ \phi T$ is the column space of the matrix A.

Hence rank of ϕT = $\dim Im\ \phi T$ = column rank of A = rank of A.

We have $\dim Ker\ T + \dim Im\ T = \dim V$; and ... (i)

$\dim Ker\ \phi T + \dim Im\ \phi T = \dim V$ ... (ii)

We now prove that $\dim Ker\ T = \dim Ker\ \phi\ T$.

Let $\alpha \in Ker\ T$. Then $T(\alpha) = \theta \in V$. Therefore $\phi T(\alpha) = \theta \in F^n$. This shows $\alpha \in Ker\ \phi T$. Hence $Ker\ T \subset Ker\ \phi T$ ... (iii)

Let $\beta \in Ker\ \phi T$. Then $\phi T(\beta) = \theta \in F^n$. This shows $T(\beta) = \theta \in V$, since ϕ is an isomorphism. Therefore $\beta \in Ker\ T$. Hence $Ker\ \phi T \subset Ker\ T$ ... (iv)

From (iii) and (iv) $Ker\ T = Ker\ \phi T$ and therefore $\dim Ker\ T = \dim Ker\ \phi T$. This together with (ii) gives $\dim Im\ T = \dim Im\ \phi T$, i.e., rank of T = rank of ϕT ..... (v)

From (i) and (v) we have rank of T = rank of A, i.e., rank of T = rank of $m(T)$. This completes the proof.

Theorem 4.23.3. Let V be a vector space of dimension n over a field F. Let $T \in L(V,V)$ and $m(T)$ be the matrix of T with respect to some ordered basis of V. Then

$$m(T_1T_2) = m(T_1).m(T_2) \text{ for all } T_1, T_2 \in L(V,V).$$

Proof. Relative to an ordered basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of V, let $m(T_1) = (a_{ij})_{n,n}$, $m(T_2) = (b_{ij})_{n,n}$; $a_{ij} \in F, b_{ij} \in F$.

Then $T_1(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$ for $j = 1, 2, \ldots, n$;
$T_2(\alpha_j) = b_{1j}\alpha_1 + b_{2j}\alpha_2 + \cdots + b_{nj}\alpha_n$ for $j = 1, 2, \ldots, n$.

$$\begin{aligned} T_1T_2(\alpha_j) &= T_1(T_2(\alpha_j)) \\ &= T_1(b_{1j}\alpha_1 + b_{2j}\alpha_2 + \cdots + b_{nj}\alpha_n) \\ &= b_{1j}T_1(\alpha_1) + \beta_{2j}T_1(\alpha_2) + \cdots + b_{nj}T_1(\alpha_n) \\ &= b_{1j}(a_{11}\alpha_1 + \cdots + a_{n1}\alpha_n) + b_{2j}(a_{12}\alpha_1 + \\ &\quad \cdots + a_{n2}\alpha_n) + \cdots + b_{nj}(a_{1n}\alpha_1 + \cdots + a_{nn}\alpha_n) \\ &= (\sum_{k=1}^{n} a_{1k}b_{kj})\alpha_1 + \cdots\cdots + (\sum_{k=1}^{n} a_{nk}b_{kj})\alpha_n. \end{aligned}$$

This shows that

$$\begin{aligned} m(T_1T_2) &= \begin{pmatrix} \Sigma a_{1k}b_{k1} & \Sigma a_{1k}b_{k2} & \cdots & \Sigma a_{1k}b_{kn} \\ \Sigma a_{2k}b_{k1} & \Sigma a_{2k}b_{k2} & \cdots & \Sigma a_{2k}b_{kn} \\ & & \cdots & \\ \Sigma a_{nk}b_{k1} & \Sigma a_{nk}b_{k2} & \cdots & \Sigma a_{nk}b_{kn} \end{pmatrix} \\ &= (a_{ij})_{n,n}.(b_{ij})_{n,n} = m(T_1).m(T_2). \end{aligned}$$

Theorem 4.23.4. Let V be a vector space of dimension n over a field F. Then a linear mapping $T \in L(V,V)$ is an isomorphism if and only if $m(T)$ is non-singular, where $m(T)$ is the matrix of T relative to any chosen ordered basis of V.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V and let $A = (a_{ij})_{n,n}$ be the matrix of T relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Let $T : V \to V$ be an isomorphism. Then T is one-to-one and onto.

Because T is onto, $Im\ T = V$. Rank of T = rank of A gives rank of A= dim $V = n$. Therefore A is non-singular.

Conversely, let $m(T)$ be non-singular. Then rank of $m(T)$ is n.

Rank of T = rank of $m(T)$ gives dim $Im\ T = n$. $Im\ T$ being a subspace of V, $Im\ T = V$. That is, T is onto.

dim $Ker\ T$+ dim $Im\ T$ = dim V gives dim $Ker\ T = 0$. That is, T is one-to-one.

Since T is one-to-one and onto, T is an isomorphism.

This completes the proof.

Matrix of the inverse operator.

Theorem 4.23.5. Let V be a finite dimensional vector space over a field F. Let $T \in (V,V)$ be an isomorphism and $m(T)$be the matrix of T relative to an ordered basis of V. Then the matrix of the inverse mapping $T' \in (V,V)$ relative to the same basis is given by

$$m(T') = [m(T)]^{-1}.$$

Proof. Let dim $V = n$. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an orderd basis of V. Let A, B be the matrices of T, T' respetively relative to the basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Then the matrix of TT' is AB and the matrix of $T'T$ is BA relative to the basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$.

Since $TT' = I_V$ and $T'T = I_V$, where I_V is the identity mapping on V and the matrix of the identity mapping is I_n it follows that $AB = I_n$ and $BA = I_n$. $AB = BA = I_n$ implies B is the inverse of A and therefore $m(T') = [m(T)]^{-1}$.

This completes the proof.

Dimension of the linear space $L(V,V)$.

The linear space $L(V,V)$ is isomorphic to the linear space $M_{n,n}$ of all $n \times n$ matrices over the field F. Therefore the dimension of the linear space $L(V,V)$ is n^2. [Theorem 4.22.2]

Worked Example (continued).

2\. Let $(\alpha_1, \alpha_2, \alpha_3)$ be an ordered basis of a real vector space V and a linear mapping $T : V \to V$ is defined by

$T(\alpha_1) = \alpha_1 + \alpha_2 + \alpha_3, T(\alpha_2) = \alpha_1 + \alpha_2, T(\alpha_3) = \alpha_1$.

Show that T is non-singular. Find the matrix of T^{-1} relative to the ordered basis $(\alpha_1, \alpha_2, \alpha_3)$.

Let $m(T)$ be the matrix of T relative to the ordered basis $(\alpha_1, \alpha_2, \alpha_3)$.

Then $m(T) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

$m(T)$ is non-singular and therefore T is non-singular.

$T^{-1}(\alpha_1 + \alpha_2 + \alpha_3) = \alpha_1, T^{-1}(\alpha_1 + \alpha_2) = \alpha_2, T^{-1}(\alpha_1) = \alpha_3$.

Therefore $T^{-1}(\alpha_1) + T^{-1}(\alpha_2) + T^{-1}(\alpha_3) = \alpha_1; T^{-1}(\alpha_1) + T^{-1}(\alpha_2)) = \alpha_2$ and $T^{-1}(\alpha_1) = \alpha_3$, since T^{-1} is linear.

We have $T^{-1}(\alpha_1) = \alpha_3$; $T^{-1}(\alpha_2)) = \alpha_2 - \alpha_3$; $T^{-1}(\alpha_3) = \alpha_1 - \alpha_2$.

Therefore $m(T^{-1}) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}$.

We have seen that the matrix $m(T)$ associated with a linear mapping $T \in L(V,V)$ depends on the choice of an ordered basis.

How different matrices of a linear mapping $T \in L(V,V)$ with respect to different ordered bases of V are related to one another, is discussed in the following theorem.

Theorem 4.23.6. Let V be a vector space of dimension n over a field F and let $T \in L(V,V)$. Let the matrix of T relative to an ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V be A and let the matrix of T relative to another ordered basis $(\beta_1, \beta_2, \ldots, \beta_n)$ of V be B. Then there exists a non-singular matrix P such that $B = P^{-1}AP$.

Proof. Let $A = (a_{ij})_{n,n}$; $a_{ij} \in F$; $B = (b_{ij})_{n,n}$, $b_{ij} \in F$.

$$\begin{aligned} \text{Then } T(\alpha_j) &= a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n \\ T(\beta_j) &= b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{nj}\beta_n. \end{aligned}$$

Let $S \in L(V,V)$ be such that S maps the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ to $(\beta_1, \beta_2, \ldots, \beta_n)$. Since S maps a basis onto another basis of V, S is non-singular. Let the matrix of S relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be $P = (p_{ij})_{n,n}$, $p_{ij} \in F$.

Then $\beta_j = S(\alpha_j) = p_{1j}\alpha_1 + p_{2j}\alpha_2 + \cdots + p_{nj}\alpha_n$, $j = 1, 2, \ldots, n$.

Since S is non-singular, P is a non-singular matrix, by Theorem 4.23.3.

$$\begin{aligned} \text{Now } T(\beta_j) &= T(p_{1j}\alpha_1 + p_{2j}\alpha_2 + \cdots + p_{nj}\alpha_n) \\ &= p_{1j}T(\alpha_1) + p_{2j}T(\alpha_2) + \cdots + p_{nj}T(\alpha_n) \\ &= p_{1j}[a_{11}\alpha_1 + \cdots + a_{n1}\alpha_n] + p_{2j}[a_{12}\alpha_1 + \\ &\quad \cdots + a_{n2}\alpha_n] + \cdots + p_{nj}[a_{1n}\alpha_1 + \cdots + a_{nn}\alpha_n] \\ &= (\sum_{k=1}^{n} a_{1k}p_{kj})\alpha_1 + \cdots + (\sum_{k=1}^{n} a_{nk}p_{kj})\alpha_n. \end{aligned}$$

$$\begin{aligned} \text{Also } T(\beta_j) &= b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{nj}\beta_n \\ &= b_{1j}[p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n] + b_{2j}[p_{12}\alpha_1 \\ &\quad + \cdots + p_{n2}\alpha_n] + \cdots + b_{nj}[p_{1n}\alpha_1 + \cdots + p_{nn}\alpha_n] \\ &= (\sum_{k=1}^{n} p_{1k}b_{kj})\alpha_1 + \cdots + (\sum_{k=1}^{n} p_{nk}b_{kj})\alpha_n. \end{aligned}$$

Since $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a basis, we have

$\sum_{k=1}^{n} a_{ik}p_{kj} = \sum_{k=1}^{n} p_{ik}b_{kj}$ for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, n$.

This gives $AP = PB$.

Since P is non-singular, $B = P^{-1}AP$.

Thus the matrix of T relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_n)$ is $P^{-1}AP$, where P is a non-singular matrix and P is the matrix of the linear mapping $S \in L(V, V)$ that maps the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ on to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_n)$.

Definition. An $n \times n$ matrix A over a field F is said to be *similar* to an $n \times n$ matrix B over F if there exists a non-singular matrix P over F such that $B = P^{-1}AP$.

We observe that a linear mapping $T \in L(V, V)$ has different matrices relative to different ordered bases of V but all such matrices are similar.

The relation of similarity is an equivalence relation on the set of all $n \times n$ matrices over F. The set is partitioned into classes of similar matrices.

The above theorem says that each linear operator $T \in L(V, V)$ corresponds to one such class of similar matrices and conversely, to each class of similar matrices there exists one linear operator in $L(V, V)$.

Since the class of similar matrices containing the identity matrix I_n is a class that contains I_n only (since $P^{-1}I_nP = I_n$ for any non-singular matrix P), the matrix of the identity operator on V is I_n relative to any chosen ordered basis of V.

Similar arguments establish that the matrix of the null operator on V is the null matrix O relative to any chosen ordered basis of V.

We now come to the converse problem. Given two $n \times n$ similar matrices A and B, does there exist a linear mapping $T \in L(V, V)$ such that A and B are matrices of T relative to two different ordered bases of V? This is discussed in the following Theorem.

Theorem 4.23.7. Let V be a vector space of dimension n over a field F and A, B are $n \times n$ matrices over F such that $B = P^{-1}AP$ for some non-singular $n \times n$ matrix P. Then there exists a linear mapping $T \in L(V, V)$ such that A and B are matrices of T relative to two different ordered bases of V.

Proof. Let $A = (a_{ij})_{n,n}$; $B = (b_{ij})_{n,n}$; $P = (p_{ij})_{n,n}$.

Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V and let $T : V \to V$ be defined by $T(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$, $j = 1, 2, \ldots, n$.

Then T is a uniquely determined linear mapping and the matrix of T relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is A.

Let $S : V \to V$ be defined by $S(\alpha_j) = p_{1j}\alpha_1 + p_{2j}\alpha_2 + \cdots + p_{nj}\alpha_n$.

Then S is a uniquely determined mapping on V. Since P is a non-singular matrix, $\{S(\alpha_1), S(\alpha_2), \ldots, S(\alpha_n)\}$ is a basis of V.

Let $S(\alpha_i) = \beta_i$.

Let T' be the linear mapping $\in L(V, V)$ whose matrix relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_n)$ is B.

$$\begin{aligned} \text{Then } T'(\beta_j) &= b_{1j}\beta_1 + b_{2j}\beta_2 + \cdots + b_{nj}\beta_n \\ &= b_{1j}[p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n] + b_{2j}[p_{12}\alpha_1 + \\ &\quad + \cdots + p_{n2}\alpha_n] + \cdots + b_{nj}[p_{1n}\alpha_1 + \cdots + p_{nn}\alpha_n] \\ &= (\sum_{k=1}^{n} p_{1k}b_{kj})\alpha_1 + \cdots + (\sum_{k=1}^{n} p_{nk}b_{kj})\alpha_n. \end{aligned}$$

$$\begin{aligned} \text{Again } T(\beta_j) &= T[p_{1j}\alpha_1 + p_{2j}\alpha_2 + \cdots + p_{nj}\alpha_n] \\ &= p_{1j}T(\alpha_1) + p_{2j}T(\alpha_2) + \cdots + p_{nj}T(\alpha_n) \\ &= p_{1j}[a_{11}\alpha_1 + \cdots + a_{n1}\alpha_n] + p_{2j}[a_{12}\alpha_1 + \cdots + \\ &\quad a_{n2}\alpha_n] + \cdots + p_{nj}[a_{1n}\alpha_1 + \cdots + a_{nn}\alpha_n] \\ &= (\sum_{k=1}^{n} a_{1k}p_{kj})\alpha_1 + \cdots + (\sum_{k=1}^{n} a_{nk}p_{kj})\alpha_n. \end{aligned}$$

$$\begin{aligned} B = P^{-1}AP &\Rightarrow PB = AP \\ &\Rightarrow \sum_{k=1}^{n} p_{ik}b_{kj} = \sum_{k=1}^{n} a_{ik}p_{kj} \text{ for } j = 1, 2, \ldots, n. \end{aligned}$$

So $T'(\beta_j) = T(\beta_j)$ for $j = 1, 2, \ldots, n$ and this proves $T = T'$.

Thus A and B are matrices of T relative to the ordered bases $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $(\beta_1, \beta_2, \ldots, \beta_n)$ respectively, where P is the matrix of the non-singular mapping S that maps α_i to β_i, $i = 1, 2, \ldots, n$.

Let us reverse the roles of matrices and linear mappings in $L(V, V)$.

Theorem 4.23.8. Let V be a vector space of dimension n over a field F and A be an $n \times n$ matrix over F. Relative to two different ordered bases of V let A determines two linear maps T_1 and T_2 in $L(V, V)$ such that $m(T_1) = A$ and $m(T_2) = A$. Then there exists a non-singular linear map $S \in L(V, V)$ such that $T_1 = S^{-1}T_2S$.

Proof. Let $A = (a_{ij})_{n,n}$ and $m(T_1) = A$ relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ and $m(T_2) = A$ relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_n)$. Then
$T_1(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$, $T_2(\beta_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n$.

Let $S : V \to V$ be defined by $S(\alpha_i) = \beta_i$, $i = 1, 2, \ldots, n$. Since S maps a basis to another basis S is non-singular.

$$T_2S(\alpha_j) = T_2(\beta_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n,$$
$$\begin{aligned} ST_1(\alpha_j) &= S(a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n \\ &= a_{1j}S(\alpha_1) + a_{2j}S(\alpha_2) + \cdots + a_{nj}S(\alpha_n) \\ &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n. \end{aligned}$$

$T_2S(\alpha_j) = ST_1(\alpha_j)$ for $j = 1, 2, \ldots, n$ and therefore $T_2S = ST_1$, i.e., $T_1 = S^{-1}T_2S$.
This completes the proof.

The converse problem is discussed in the following theorem.

Theorem 4.23.9. Let V be a vector space of dimension n over a field F. Let T_1, T_2 be two linear maps in $L(V, V)$ such that $T_1 = S^{-1}T_2S$, where S is a non-singular mapping in $L(V, V)$. Then $m(T_1) = m(T_2)$ relative to different ordered bases of V.

Proof. Let $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ be an ordered basis of V. Let $S(\alpha_i) = \beta_i$. Since S is non-singular, $(S(\alpha_1), S(\alpha_2), \ldots, S(\alpha_n))$ is a basis of V.

Let $A = (a_{ij})_{m,n}$ be the matrix of T_1 relative to the ordered basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V.

Then $T_1(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$,
$$\begin{aligned} ST_1(\alpha_j) &= a_{1j}S(\alpha_1) + a_{2j}S(\alpha_2) + \cdots + a_{nj}S(\alpha_n) \\ &= a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n. \end{aligned}$$

Since $ST_1 = T_2S$, we have $T_2S(\alpha_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n$
or, $T_2(\beta_j) = a_{1j}\beta_1 + a_{2j}\beta_2 + \cdots + a_{nj}\beta_n$.

This shows that A is the matrix of T_2 relative to the ordered basis $(\beta_1, \beta_2, \ldots, \beta_m)$ of V.
This completes the proof.

Exercises 18

1. Let $(\alpha_1, \alpha_2, \alpha_3)$, (β_1, β_2) be ordered bases of the real vector spaces V and W respectively and a linear mapping $T : V \to W$ maps the basis vectors as $T(\alpha_1) = \beta_1 + \beta_2$, $T(\alpha_2) = 2\beta_1 - \beta_2$, $T(\alpha_3) = \beta_1 + 3\beta_2$.

   Find the matrix of T relative to the ordered bases
   (i) $(\alpha_2, \alpha_3, \alpha_1)$ of V and (β_1, β_2) of W;
   (ii) $(\alpha_1 + \alpha_2, \alpha_2, \alpha_3)$ of V and $(\beta_1, \beta_1 + \beta_2)$ of W.

2. Let $(\alpha_1, \alpha_2, \alpha_3)$, $(\beta_1, \beta_2, \beta_3)$ be ordered bases of the real vector spaces V and W respectively. A linear mapping $T : V \to W$ maps the basis vectors as $T(\alpha_1) = \beta_1 + \beta_2$, $T(\alpha_2) = \beta_2 + \beta_3$, $T(\alpha_3) = \beta_3$.

   Find the matrix of T relative to the ordered bases $(\alpha_1, \alpha_2, \alpha_3)$ of V and $(\beta_1, \beta_2, \beta_3)$ of W. Deduce that T is invertible. Find the matrix of T^{-1} relative to the same chosen ordered bases.

3. Let $(\alpha_1, \alpha_2, \alpha_3)$ be an ordered basis of a real vector space V and a linear mapping $T : V \to V$ is defined by
   $T(\alpha_1) = \alpha_1 - \alpha_2 + \alpha_3$, $T(\alpha_1 + \alpha_2) = 2\alpha_1 + \alpha_2 - \alpha_3$, $T(\alpha_1 + \alpha_2 + \alpha_3) = \alpha_1 + \alpha_2$.
   Find the matrix of T relative to the ordered basis $(\alpha_1, \alpha_2, \alpha_3)$.

4. D and T are linear mappings on the real vector space P_4 defined by
   $Dp(x) = \frac{d}{dx}p(x), p(x) \in P_4$ and $Tp(x) = x\frac{d}{dx}p(x), p(x) \in P_4$.
   Relative to the basis $(1, x, x^2, x^3)$ of P_4 determine the matrix of each of the linear mappings (i) D, (ii) T, (iii) $TD - DT$, (iv) $T^2D^2 - D^2T^2$.

5. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 2x_1 + 3x_2 + x_3, 3x_1 + x_2 + 2x_3)$ $(x_1, x_2, x_3) \in \mathbb{R}^3$. Find the matrix of T relative to
   (i) the ordered basis $((-1, 1, 1), (1, -1, 1), (1, 1, -1))$ of $\mathbb{R}^3$;
   (ii) the ordered basis $((0, 1, 1), (1, 0, 1), (1, 1, 0))$ of $\mathbb{R}^3$.

6. Find the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ if $T(1, 0, 0) = (2, 3, 4)$, $T(0, 1, 0) = (1, 2, 3)$, $T(0, 0, 1) = (1, 1, 1)$.

   Find the matrix of T relative to the ordered basis $(\epsilon_1, \epsilon_2, \epsilon_3)$, where $\epsilon_1 = (1, 0, 0)$, $\epsilon_2 = (0, 1, 0)$, $\epsilon_3 = (0, 0, 1)$. Deduce that T is not invertible.

7. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(0, 1, 1) = (1, 0, 1)$, $T(1, 0, 1) = (2, 3, 4)$, $T(1, 1, 0) = (1, 2, 3)$.

   Find the matrix of T relative to the ordered basis $(\epsilon_1, \epsilon_2, \epsilon_3)$, where $\epsilon_1 = (1, 0, 0)$, $\epsilon_2 = (0, 1, 0)$, $\epsilon_3 = (0, 0, 1)$. Deduce that T is invertible.

   Verify that $m(T^{-1}) = [m(T)]^{-1}$, where $m(T)$ is the matrix of T and $m(T^{-1})$ is the matrix of T^{-1} relative to the ordered basis $(\epsilon_1, \epsilon_2, \epsilon_3)$.

8. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $T(x,y,z) = (2x+z, x+y+z, -3x-2z)$, $(x,y,z) \in \mathbb{R}^3$. Show that (i) T is as isomorphism, (ii) $T^{-1} = T$.

9. The matrix of a linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ relative to the ordered basis $((-1,1,1),(1,-1,1),(1,1,-1))$ of $\mathbb{R}^3$ is

$$\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 3 \\ 3 & 3 & 1 \end{pmatrix}.$$

Find the matrix of T relative to the ordered basis
(i) $((1,0,0),(0,1,0),(0,0,1))$ of $\mathbb{R}^3$; (ii) $((0,1,1),(1,0,1),(1,1,0))$ of $\mathbb{R}^3$.

10. Let V, W, U be a finite dimensional vector spaces over a field F and $T : V \to W$, $S : W \to U$ are linear mappings. Prove that

(i) rank of ST = rank of S, if T is surjective,

(ii) rank of ST = rank of T, if S is injective.

11. Let V be a finite dimensional vector space over a field F and T_1, T_2 are linear operators on V. Prove that

(i) rank of $T_2T_1 \leq \min\{\text{rank of } T_2, \text{rank of } T_1\}$,

(ii) rank of T_2T_1 = rank of T_1, if T_2 be invertible.

12. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by
$$T(x,y,z) = (y+z, x+z, x+y), (x,y,z) \in \mathbb{R}^3.$$

Let the matrix of T relative to the ordered bases $((1,0,0), (0,1,0), (0,0,1))$ of $\mathbb{R}^3$ be A and the matrix of T relative to the ordered bases $((2,1,1),(1,2,1),(1,1,2))$ of $\mathbb{R}^3$ be B.

Let the linear mapping $S : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by $S(1,0,0) = (2,1,1)$, $S(0,1,0) = (1,2,1)$, $S(0,0,1) = (1,1,2)$ and let the matrix of S relative to the ordered bases $((1,0,0), (0,1,0), (0,0,1))$ of $\mathbb{R}^3$ be P. Show that $B = P^{-1}AP$.

13. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by
$$T(x,y,z) = (x+y+z, x+z, x+y), (x,y,z) \in \mathbb{R}^3.$$

If the the matrix of T relative to the ordered bases $((1,0,0), (0,1,0), (0,0,1))$ of $\mathbb{R}^3$ be A and the matrix of T relative to the ordered bases $((0,1,1),(1,0,1),(1,1,0))$ of $\mathbb{R}^3$ be B, show that A and B are similar matrices.

14. A linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ is defined by
$$T(x,y,z) = (x+2y+z, 2x+y+2z, x+y+z), (x,y,z) \in \mathbb{R}^3.$$

If the the matrix of T relative to the ordered bases $((1,0,0), (0,1,0), (0,0,1))$ of $\mathbb{R}^3$ be A and the matrix of T relative to the ordered bases $((0,1,1),(1,0,1),(1,1,0))$ of $\mathbb{R}^3$ be B, show that A and B are similar matrices.

4.24. Linear functionals.

Let V be a vector space over a field F. A linear mapping $f : V \to F$ is said to be a *linear functional* (or a *linear form*). Therefore
(i) $f(\alpha+\beta) = f(\alpha) + f(\beta)$ for all $\alpha, \beta \in V$; and
(ii) $f(c\alpha) = cf(\alpha)$ for all $\alpha \in V$ and all $c \in F$.

Note. Because a linear functional f is a linear mapping, $f(\theta) = 0$, where θ is the null vector in V and 0 is the zero element in F.

Examples.

1. Let V be a vector space over a field F. The mapping $f : V \to F$ defined by $f(\alpha) = 0$ for all $\alpha \in V$ (0 being the zero element in F) is a linear functional. It is called the *zero functional*.

2. In the real vector space $\mathcal{C}[0,1]$ of all real functions continuous on $[0,1]$, let T be the functional defined by $T(f) = \int_0^1 f(x)dx$ for all $f \in \mathcal{C}[0,1]$.

T is a linear functional which assigns to each function in $\mathcal{C}[0,1]$ a real number which is the integral of f on $[0,1]$.

3. In the real vector space $\mathcal{P}$ of all real polynomials, let T be the functional defined by $T(f) = f(0)$ for all $f \in \mathcal{P}$.

T is a linear functional which assigns to each polynomial f in $\mathcal{P}$ the real number $f(0)$.

4. Let V be a vector space of dimension n over a field F. The mapping $f : V \to F$ defined by $f(\xi) = x_1+x_2+\cdots+x_n$ for all $\xi = (x_1, x_2, \ldots, x_n) \in V$, is a linear functional.

5. Let V be a Euclidean space of dimension n with standard inner product. Let $\gamma = (c_1, c_2, \ldots, c_n)$ be a fixed vector in V. The mapping $f : V \to \mathbb{R}$ defined by $f(\xi) = (\gamma, \xi) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ for all $\xi = (x_1, x_2, \ldots, x_n) \in V$, is a linear functional.

Theorem 4.24.1. Let V be a vector space of dimension n over a field F and $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis of V. For arbitrary n scalars $c_1, c_2, \ldots, c_n$ in F there exists a unique linear functional $f : V \to F$ such that $f(\alpha_1) = c_1$, $f(\alpha_2) = c_2$, ..., $f(\alpha_n) = c_n$.

Proof. Let ξ be an arbitrary vector in V. Since $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is a basis of V, there exist unique scalars $x_1, x_2, \ldots, x_n$ in F such that $\xi = x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n$.

Let us define a mapping $f : V \to F$ by $f(\xi) = c_1x_1 + c_2x_2 + \ldots + c_nx_n$ for $\xi (= x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n) \in V$.

The mapping f is such that $f(\alpha_1) = f(1,0,\ldots,0) = c_1$, $f(\alpha_2) = f(0,1,\ldots,0) = c_2, \ldots, f(\alpha_n) = f(0,0,\ldots,1) = c_n$.

Let $\eta = y_1\alpha_1 + y_2\alpha_2 + \ldots + y_n\alpha_n$ be a vector in V.

Then for any scalars c, d in F, $c\xi + d\eta = (cx_1 + dy_1)\alpha_1 + (cx_2 + dy_2)\alpha_2 + \ldots + (cx_n + dy_n)\alpha_n$.

$f(c\xi + d\eta)$
$= c_1(cx_1 + dy_1) + c_2(cx_2 + dy_2) + \ldots + c_n(cx_n + dy_n)$, by definition
$= c(c_1x_1 + c_2x_2 + \ldots + c_nx_n) + d(c_1y_1 + c_2y_2 + \ldots + c_ny_n)$
$= cf(\xi) + df(\eta)$.

This proves that f is a linear functional.

To prove that f is unique, let g be a linear functional on v such that $g(\alpha_1) = c_1, g(\alpha_2) = c_2, \ldots, g(\alpha_n) = c_n$.

Let $\rho = r_1\alpha_1 + r_2\alpha_2 + \ldots + r_n\alpha_n$ be an arbitrary vector in V. Then $g(\rho) = g(r_1\alpha_1 + r_2\alpha_2 + \ldots + r_n\alpha_n$
$= r_1g(\alpha_1) + r_2g(\alpha_2) + \cdots + r_ng(\alpha_n)$, since g is linear
$= r_1c_1 + r_2c_2 + \cdots + r_nc_n = f(\rho)$.

Hence $g(\rho) = f(\rho)$ for all $\rho \in V$. Therefore $g = f$.

This completes the proof.

The set of all linear functionals on a vector space V over a field F form a linear space over F and it is denoted by $L(V, F)$.

The linear space $L(V, F)$ is called the *dual space* of V and is denoted by V'.

Theorem 4.24.2. Let V be a vector space of dimension n over a field F. Then the dimension of the dual space is also n.

Proof. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be a basis of V.

For $i = 1, 2, \ldots, n$, let $f_i : V \to F$ be defined by $f_i(\alpha_j) = 1$ if $i = j$
$= 0$, if $i \neq j$.

The linear functional f_i maps the basis vector α_i to 1 and all other basis vectors to 0.

We show that the set $\{f_1, f_2, \ldots, f_n\}$ is a basis of the linear space V'.

The zero element in V' is the zero linear functional f_θ that maps each vector in V to the zero element in F.

To establish linear independence of the set, let us consider the relation $c_1f_1 + c_2f_2 + \cdots + c_nf_n = f_\theta$ for $c_i \in F$

Then $(c_1f_1 + c_2f_2 + \cdots + c_nf_n)(\alpha_i) = f_\theta(\alpha_i) = 0$ for $i = 1, 2, \ldots, n$.

But $(c_1f_1 + c_2f_2 + \cdots + c_nf_n)(\alpha_i) = c_i$ for $i = 1, 2, \ldots, n$, since $f_k(\alpha_i) = 1$, if $i = k$ and $= 0$, if $i \neq k$.

Therefore $c_1 = 0, c_2 = 0, \ldots, c_n = 0$ and this proves linear independence of the set.

To establish that the set $\{f_1, f_2, \ldots, f_n\}$ spans V', let f be any element of V' defined by $f(\alpha_1) = d_1, f(\alpha_2) = d_2, \ldots, f(\alpha_n) = d_n, d_i \in F$.

Let $\alpha \in V$ and $\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$ for scalars $a_i \in F$.

Then $f(\alpha) = f(a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n)$
$= a_1f(\alpha_1) + a_2f(\alpha_2) + \cdots + a_nf(\alpha_n)$, since f is linear
$= a_1d_1 + a_2d_2 + \cdots + a_nd_n$, since $f(\alpha_i) = d_i$
$= d_1f_1(\alpha) + d_2f_2(\alpha) + \cdots + d_nf_n(\alpha)$
$= (d_1f_1 + d_2f_2 + \cdots + d_nf_n)(\alpha)$.

This holds for all $\alpha \in V$. Therefore $f = d_1f_1 + d_2f_2 + \cdots + d_nf_n$.

This shows that the set $\{f_1, f_2, \ldots, f_n\}$ spans V'.

Hence the set $\{f_1, f_2, \ldots, f_n\}$ is a basis of the linear space V' and therefore the dimension of V' is n.

This completes the proof.

Definition. The basis $\{f_1, f_2, \ldots, f_n\}$ is said to be the *dual basis* of $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$.

Worked Example.

1. Let S be a subspace of a linear space V of finite dimension over a field F and $\alpha \in V - S$. Show that there exists a linear functional on V such that $f(\alpha) = 1$ and $f(\xi) = 0$ for all $\xi \in S$.

Let dim $V = n$ and dim $S = r (< n)$. Let $\{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ be a basis of S. Since $\alpha \in V - S$, the set $\{\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha\}$ is linearly independent. If $n = r + 1$, this is a basis of V. If $n > r + 1$, it can be extended to a basis of V; and let the extended basis be $\alpha_1, \alpha_2, \ldots, \alpha_r, \alpha, \beta_1, \beta_2, \ldots, \beta_s\}$, where $r + 1 + s = n$.

Let f be a linear functional on V such that $f(\alpha_1) = 0, f(\alpha_2) = 0, \ldots, f(\alpha_r) = 0, f(\alpha) = 1, f(\beta_1) = a_1, f(\beta_2) = a_2, \ldots, f(\beta_s) = a_s$, where $a_1, a_2, \ldots, a_s$ are arbitrary scalars in F.

Let $\xi \in S$ and $\xi = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_r\alpha_r$ for $c_i \in F$. Then $f(\xi) = c_1f(\alpha_1) + c_2f(\alpha_2) + \cdots + c_rf(\alpha_r) = 0$.

Hence $f(\xi) = 0$ for all $\xi \in S$ and $f(\alpha) = 1$.

4.25. Orthogonal mappings on Euclidean spaces.

Let V be a Euclidean space. A linear mapping $T : V \to V$ is said to be *orthogonal* on V if $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in V$.

An orthogonal mapping preserves inner products.

Example.

1. Let $\mathbb{R}^2$ be the Euclidean space with standard inner product. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\epsilon_1) = \epsilon_2, T(\epsilon_2) = -\epsilon_1$. Show that T is orthogonal.

Let $\alpha, \beta \in \mathbb{R}^2$ and $\alpha = a_1\epsilon_1 + a_2\epsilon_2, \beta = b_1\epsilon_1 + b_2\epsilon_2$.

$T(\alpha) = a_1T(\epsilon_1) + a_2T(\epsilon_2) = a_1\epsilon_2 - a_2\epsilon_1$;

$T(\beta) = b_1T(\epsilon_1) + b_2T(\epsilon_2) = b_1\epsilon_2 - b_2\epsilon_1$.

$(T(\alpha), T(\beta)) = (a_1\epsilon_2 - a_2\epsilon_1, b_1\epsilon_2 - b_2\epsilon_1) = a_1b_1 + a_2b_2$

and $(\alpha, \beta) = (a_1\epsilon_1 + a_2\epsilon_2, b_1\epsilon_1 + b_2\epsilon_2) = a_1b_1 + a_2b_2$.

So $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in \mathbb{R}^2$ and therefore T is orthogonal.

Theorem 4.25.1. Let V be a Euclidean space. If a linear mapping $T : V \to V$ is orthogonal on V then for all $\alpha, \beta \in V$,

(i) $(\alpha, \beta) = 0 \Rightarrow (T(\alpha), T(\beta)) = 0$;

(ii) $\| T(\alpha) \| = \| \alpha \|$;

(iii) $\| T(\alpha) - T(\beta) \| = \| \alpha - \beta \|$;

(iv) T is one-to-one.

(i) says that T preserves orthogonality; (ii) says that T preserves norms; (iii) says that T preserves distance between two vectors.

Proof. (i) Since T is orthogonal on V, $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in V$. Therefore $(\alpha, \beta) = 0 \Rightarrow (T(\alpha), T(\beta)) = 0$.

(ii) Since T is orthogonal on V, $(T(\alpha), T(\alpha)) = (\alpha, \alpha)$ for all $\alpha \in V$.
Therefore $\| T(\alpha) \| = \| \alpha \|$.

(iii) $T(\alpha) - T(\beta) = T(\alpha - \beta)$, since T is linear.
Therefore $\| T(\alpha) - T(\alpha) \| = \| T(\alpha - \beta) \| = \| \alpha - \beta \|$ by (ii).

(iv) Let $T(\alpha) = \theta$. Since T is orthogonal, $\| T(\alpha) \| = \| \alpha \|$.
$T(\alpha) = \theta \Rightarrow | \alpha | = \theta \Rightarrow \alpha = \theta$. Therefore T is one-to-one.

This completes the proof.

If V be a finite dimensional Euclidean space then an orthogonal mapping on V can be characterised by its action on any basis of V.

Theorem 4.25.2. Let $\{\alpha_1, \alpha_2, ..., \alpha_n\}$ be a basis of Euclidean space V of dimension n. Then a linear mapping $T : V \to V$ is orthogonal on V if and only if $(T(\alpha_i), T(\alpha_j)) = (\alpha_i, \alpha_j)$ for all i, j.

Proof. Let T be orthogonal. Then $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in V$. Therefore $(T(\alpha_i), T(\alpha_j)) = (\alpha_i, \alpha_j)$ for all i, j.

Conversely, let $(T(\alpha_i), T(\alpha_j)) = (\alpha_i, \alpha_j)$ for all i, j.

Let α, β be any two elements of V and $\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n$, $a_i \in \mathbb{R}$, $\beta = b_1\beta_1 + b_2\beta_2 + \cdots + b_n\beta_n$, $b_i \in \mathbb{R}$.

Then $T(\alpha) = a_1T(\alpha_1) + a_2T(\alpha_2) + \cdots + a_nT(\alpha_n)$;

$T(\beta) = b_1T(\alpha_1) + b_2T(\alpha_2) + \cdots + b_nT(\alpha_n)$.

$(T(\alpha), T(\beta)) = (a_1T(\alpha_1) + \cdots + a_nT(\alpha_n), b_1T(\alpha_1) + \cdots + b_nT(\alpha_n))$

$= \sum_{j=1}^{n} \sum_{i=1}^{n} a_ib_j(T(\alpha_i), T(\alpha_j))$ (i)

and $(\alpha, \beta) = (a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n, b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n)$

$= \sum_{j=1}^{n} \sum_{i=1}^{n} a_ib_j(\alpha_i, \alpha_j)$. (ii)

From (i) and (ii) $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in V$. Therefore T is orthogonal on V and this completes the proof.

Theorem 4.25.3. Let V be a finite dimensional Euclidean space. Then a linear mapping $T : V \to V$ is orthogonal if and only if T maps an orthonormal basis to another orthonormal basis.

Proof. Let $B = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthonormal basis of V.

Let $T : V \to V$ be an orthogonal mapping. Then $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in V$. So $(T(\alpha_i), T(\alpha_j)) = (\alpha_i, \alpha_j)$ for all i, j.

Therefore $(T(\alpha_i), T(\alpha_j)) = (\alpha_i, \alpha_j) = 1$ if $i = j$
$= 0$ if $i \neq j$.

This proves that $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is an orthonormal set in V and as it contains n vectors, it is a basis of V.

Conversely, let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthonormal basis of V and $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is also orthonormal.

Let $\alpha = a_1\alpha_1 + a_1\alpha_2 + \cdots + a_n\alpha_n, a_i \in \mathbb{R}$ and

$\beta = b_1\alpha_1 + b_2\alpha_2 + \cdots + b_n\alpha_n, b_i \in \mathbb{R}$ belong to V.

$(\alpha, \beta) = a_1b_1 + a_1b_2 + \cdots + a_nb_n$, since $(\alpha_i, \alpha_j) = 1$ if $i = j$
$= 0$ if $i \neq j$.

$$T(\alpha) = a_1T(\alpha_1) + a_2T(\alpha_2) + \cdots + a_nT(\alpha_n);$$
$$T(\beta) = b_1T(\alpha_1) + b_2T(\alpha_2) + \cdots + b_nT(\alpha_n).$$

$(T(\alpha), T(\beta)) = a_1b_1 + a_2b_2 + \cdots + a_nb_n$,

since $(T(\alpha_i), T(\alpha_j)) = 1$ if $i = j$
$= 0$ if $i \neq j$.

Therefore $(T(\alpha), T(\beta)) = (\alpha, \beta)$ for all $\alpha, \beta \in V$ and this proves that T is an orthogonal mapping.

Matrix of an orthogonal transformation.

Theorem 4.25.4. Let V be a Euclidean space of dimension n. Let A be the matrix of a linear mapping $T : V \to V$ relative to an orthonormal basis. Then T is orthogonal on V if and only if A is a real orthogonal matrix.

Proof. Let $A = (a_{ij})_{n,n}$ be the matrix of T relative to the ordered orthonormal basis $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of V.

Then $T(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$ for $j = 1, 2, \ldots, n$

$(T(\alpha_i), T(\alpha_j)) = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj}$, since $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is orthonormal.

Let T be orthogonal on V. Then $(T(\alpha_i), T(\alpha_j)) = (\alpha_i, \alpha_j)$.

Since $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is an orthonormal set $(\alpha_i, \alpha_j) = 1$ if $i = j$
$= 0$ if $i \neq j$.

So $a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = 1$ if $i = j$
$= 0$ if $i \neq j$.

Therefore $A^tA = I_n$ and this proves that A is orthogonal.

Conversely, let A be an orthogonal matrix. Then $A^tA = I_n$.

So $a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj} = 1$ if $i = j$
$= 0$ if $i \neq j$.

$T(\alpha_i) = a_{1i}\alpha_1 + a_{2i}\alpha_2 + \cdots + a_{ni}\alpha_n$; $T(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n$.

Hence $(T(\alpha_i), T(\alpha_j)) = a_{1i}a_{1j} + a_{2i}a_{2j} + \cdots + a_{ni}a_{nj}$, since $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is an orthonormal set.

Consequently, $(T(\alpha_i), T(\alpha_j)) = 1$ if $i = j$,
$= 0$ if $i \neq j$.

This proves that $\{T(\alpha_1), T(\alpha_2), \ldots, T(\alpha_n)\}$ is an orthonormal set and as this contains n vectors, it is an orthonormal basis of V.

As T maps an orthonormal basis $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ to another orthonormal basis of V, T is orthogonal.

This completes the proof.

Exercises 19

1. Prove that the linear mapping $T : \mathbb{R}^4 \to \mathbb{R}^4$ defined by $T(x, y, z, w) = (-y, x, w, -z), (x, y, z, w) \in \mathbb{R}^4$ is orthogonal.

2. A linear mapping T on the Euclidean space $\mathbb{R}^3$ with standard inner product maps the basis vectors as

$$T(1,0,0) = \left(\tfrac{2}{3}, -\tfrac{2}{3}, -\tfrac{1}{3}\right),\ T(0,1,0) = \left(\tfrac{1}{3}, \tfrac{2}{3}, -\tfrac{2}{3}\right),\ T(0,0,1) = \left(\tfrac{2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\right).$$

Show that T is orthogonal.

3. A linear mapping T on the Euclidean space $\mathbb{R}^3$ with standard inner product maps the basis vectors as follows:

$$T(0,1,1) = (-1,0,1), T(1,0,1) = \tfrac{1}{3}(1,-1,4), T(1,1,0) = (0,1,1).$$

Find the matrix of T relative to the ordered basis $(\epsilon_1, \epsilon_2, \epsilon_3)$, where $\epsilon_1 = (1,0,0), \epsilon_2 = (0,1,0), \epsilon_3 = (0,0,1)$. Deduce that T is orthogonal.

4. The matrix of a linear mapping T on the Euclidean space $\mathbb{R}^3$ with standard inner product relative to the ordered basis $(\alpha_1, \alpha_2, \alpha_3)$, where $\alpha_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \alpha_2 = (0,1,0), \alpha_3 = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ is

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{pmatrix}.$$ Show that T is orthogonal.

4.26. Application to Geometry.

Let V be a Euclidean space of dimension 3 with ordered bases $B = (\alpha_1, \alpha_2, \alpha_3)$, $B' = (\alpha'_1, \alpha'_2, \alpha'_3)$. Let ξ be an arbitrary vector in V with co-ordinates (x_1, x_2, x_3) relative to the B-basis and co-ordinates (x'_1, x'_2, x'_3) relative to the B'-basis.

Then $\xi = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = x'_1\alpha'_1 + x'_2\alpha'_2 + x'_3\alpha'_3$.

$$\begin{aligned} \text{Let } \alpha'_1 &= a_{11}\alpha_1 + a_{21}\alpha_2 + a_{31}\alpha_3 \\ \alpha'_2 &= a_{12}\alpha_1 + a_{22}\alpha_2 + a_{32}\alpha_3 \\ \alpha'_3 &= a_{13}\alpha_1 + a_{23}\alpha_2 + a_{33}\alpha_3. \quad \ldots \quad \text{(i)} \end{aligned}$$

Then $x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3 = x'_1(a_{11}\alpha_1 + a_{21}\alpha_2 + a_{31}\alpha_3) + x'_2(a_{12}\alpha_1 + a_{22}\alpha_2 + a_{32}\alpha_3) + x'_3(a_{13}\alpha_1 + a_{23}\alpha_2 + a_{33}\alpha_3)$.

$$\begin{aligned} \text{Therefore } x_1 &= a_{11}x'_1 + a_{12}x'_2 + a_{13}x'_3 \\ x_2 &= a_{21}x'_1 + a_{22}x'_2 + a_{23}x'_3 \\ x_3 &= a_{31}x'_1 + a_{32}x'_2 + a_{33}x'_3, \end{aligned}$$

since $\{\alpha_1, \alpha_2, \alpha_3\}$ is a linearly independent set.

or, $X = AX'$ where $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $A = (a_{ij})_{3,3}$, $X = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$.

Therefore $X = AX'$ gives a transformation of co-ordinates from the B- basis to the B'-basis. The jth column of A gives co-ordinates of α_j' relative to the B-basis.

A is a non-singular matrix since both the sets $\{\alpha_1, \alpha_2, \alpha_3\}$ and $\{\alpha_1', \alpha_2', \alpha_3'\}$ are linearly independent.

Conversely, if $B = \{\alpha_1, \alpha_2, \alpha_3\}$ be an ordered basis, then for a non-singular matrix A the relation $X = AX'$ yields a transformation of co-ordinates from the B-basis to another ordered basis B', the vectors of B' being determined by (i).

Theorem 4.26.1. If B be an orthonormal basis then B' will be orthonormal if and only if A is an orthogonal matrix.

Proof. Let B' be an orthonormal basis. Then $(\alpha_i', \alpha_j') = 0$ if $i \neq j$
$= 1$ if $i = j$.

$(\alpha_i', \alpha_j') = a_{1i}a_{1j} + a_{2i}a_{2j} + a_{3i}a_{3j}$, since B is orthonormal.

Therefore $(\alpha_i', \alpha_j') = a_{1i}a_{1j} + a_{2i}a_{2j} + a_{3i}a_{3j} \;=\; 0$ if $i \neq j$
$= 1$ if $i = j$.

This implies $A^tA = I_n$ and therefore A is an orthogonal matrix.

Conversely, let A be an orthogonal matrix. Then

$a_{1i}a_{1j} + a_{2i}a_{2j} + a_{3i}a_{3j} = 0$ if $i \neq j$
$= 1$ if $i = j$.

This means $(\alpha_i', \alpha_j') = 0$ if $i \neq j$
$= 1$ if $i = j$.

Therefore B' is an orthonormal basis.

Note 1. The transformation $X = AX'$ is said to be an *orthogonal transformation* if A is an orthogonal matrix.

Note 2. The transformation $X = AX'$ is said to be a *non-singular transformation* if A be a non-singular matrix.

In Euclidean geometry the quadratic forms appear in connection with the equations of conics, quadric surfaces, hyper-surfaces.

In the plane, the equation $ax^2 + 2hxy + by^2 = k$ represents a conic. The left hand side expression is a quadratic form in two variables x, y.

In 3-space, the equation $ax^2 + by^2 + cz^2 + 2hxy + 2fyz + 2gzx = k$ represents a quadric surface. The left hand side expression is a quadratic form in three variables x, y, z.

In matrix notation,

the equation $ax^2 + 2hxy + by^2 = k$ is $X^tAX = kI_1$, where $X = \begin{pmatrix} x \\ y \end{pmatrix}$, $A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$;

the equation $ax^2 + by^2 + cz^2 + 2hxy + 2fyz + 2gzx = k$ is $X^tAX = kI_1$, where $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$.

Our main interest lies in finding a suitable orthogonal transformation $X = PX'$ which will reduce the equation to the simplest form (canonical form) in the new variable X'. Since P is orthogonal, the transformation $X = PX'$ yields a change of orthonormal basis, i.e., a change of the reference frame of the co-ordinate axes.

By the transformation $X = PX'$, the equation $X^tAX = kI_1$ transforms to $X'^t(P^tAP)X' = kI_1$.

Since A is a real symmetric matrix there exists an orthogonal matrix P such that $P^{-1}AP$ is a diagonal matrix. For such a matrix P, the transformed equation will be free from the term containing xy.

We search for an orthogonal matrix P such that the transformation $X = PX'$ will reduce the equation of a conic (or a quadric) to the canonical form.

We consider the equation of a quadric. The two dimensional case is similar and simpler.

Definition. Principal direction of a conicoid

A direction l, m, n (i.e., a line with direction ratios l, m, n) is said to be a *principal direction* of a conicoid if the line with direction ratios l, m, n is perpendicular to the diametral plane conjugate to it.

Definition. The locus of middle points of chords of a conicoid parallel to a given line is a plane, called the *diametral plane* corresponding to (or conjugate to) the given line.

Let the equation of a conicoid be $ax^2 + by^2 + cz^2 + 2hxy + 2fyz + 2gzx + 2ux + 2vy + 2wz + d = 0$.

Let (x_1, y_1, z_1) be the middle point of a chord parallel to the direction l, m, n.

The locus of (x_1, y_1, z_1) is the plane $(al + hm + gn)x + (hl + bm + fn)y + (gl + fm + cn)z + (ul + vm + wn) = 0$.

This is the diametral plane conjugate to the direction l, m, n.

The direction l, m, n will be a principal direction of the conicoid if $\frac{al+hm+gn}{l} = \frac{hl+bm+fn}{m} = \frac{gl+fm+cn}{n} = \lambda$, say.

That is, $\begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix} \begin{pmatrix} l \\ m \\ n \end{pmatrix} = \lambda \begin{pmatrix} l \\ m \\ n \end{pmatrix}$.

Since (l, m, n) is a non-zero triplet, $\begin{pmatrix} l \\ m \\ n \end{pmatrix}$ is an eigen vector of the matrix $A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$.

The homogeneous system of equations
$$\begin{aligned} (a-\lambda)l + hm + gn &= 0 \\ hl + (b-\lambda)m + fn &= 0 \\ gl + fm + (c-\lambda)n &= 0 \end{aligned}$$
admits of non-zero solutions for l, m, n.

Therefore $\begin{vmatrix} a-\lambda & h & g \\ h & b-\lambda & f \\ g & f & c-\lambda \end{vmatrix} = 0$.

This shows that λ is an eigen value of the matrix A.

Thus we observe that a principal direction of a conicoid is an eigen vector corresponding to an eigen value of the real symmetric matrix associated with the quadratic form of the equation of the conicoid.

To determine the principal directions of a quadric it is sufficient to consider the homogeneous quadratic part of the equation of the quadric.

The quadratic form is X^tAX. The matrix A associated with the form is symmetric. There are three real eigen values of A. Three cases arise.

Case 1. Three eigen values are distinct.

Let the eigen values be $\lambda_1, \lambda_2, \lambda_3$. Let us choose one eigen vector corresponding to each eigen value. Let the eigen vectors be α, β, γ. The set $\{\alpha, \beta, \gamma\}$ is an orthogonal set. The corresponding orthonormal set is $\{\frac{\alpha}{\|\alpha\|}, \frac{\beta}{\|\beta\|}, \frac{\gamma}{\|\gamma\|}\}$.

The three mutually perpendicular eigen vectors give three mutually perpendicular principal directions of the quadric. Let P be the matrix whose column vetors are $\frac{\alpha}{\|\alpha\|}, \frac{\beta}{\|\beta\|}, \frac{\gamma}{\|\gamma\|}$. Then P is an orthogonal matrix.

Let us apply the orthogonal transformation $X = PX'$, where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$. Then X^tAX transforms to $(X')^t(P^tAP)X'$. Since P is orthogonal, $(P^tAP) = P^{-1}AP$. But $P^{-1}AP$ is a diagonal matrix D whose diagonal elements are the eigen values of A. $D = \text{diag}\,(\lambda_1, \lambda_2, \lambda_3)$. X^tAX transforms to $(X')^tDX' = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$.

Case 2. Two eigen values are equal.

Let the eigen values be $\lambda_1, \lambda_2, \lambda_2$. Let us choose one eigen vector α corresponding to λ_1. Since λ_2 is a 2-fold eigen value, the dimension of the characteristic subspace of λ_2 is 2. That is, the eigen vectors corresponding to λ_2 together with the null vector form a subspace of dimension 2. Therefore there are two linearly independent vectors forming a basis of this subspace and by Gram-Schmidt method we can have an orthogonal basis $\{\beta, \gamma\}$ of this subspace.

Eigen vectors α and β are orthogonal as they correspond to two distinct eigen values. Similarly, α and γ are orthogonal.

Thus we have three mutually perpendicular eigen vectors α, β, γ. These vectors give three mutually perpendicular principal directions of the quadric. Then we proceed as in Case 1.

Case 3. Three eigen values are equal.

Let the eigen values be $\lambda, \lambda, \lambda$. Since λ is a 3-fold eigen value, the dimension of the characteristic subspace of λ is 3. That is, the subspace of solutions of the homogeneous system of equations
$$\begin{aligned} (a-\lambda)x + hy + gz &= 0 \\ hx + (b-\lambda)y + fz &= 0 \\ gx + fy + (c-\lambda)z &= 0 \end{aligned}$$
is of dimension 3.

It follows that the rank of the co-efficient matrix of the system is 0. Therefore $a = b = c = \lambda$, $h = f = g = 0$.

In this case, the equations for determining l, m, n are identically satisfied. Therefore every direction is a principal direction.

In this case, the equation of the conicoid takes the form $\lambda(x^2 + y^2 + z^2) + 2ux + 2vy + 2wz + d = 0$, which is a sphere.

Worked examples.

1. Reduce the equation $3x^2 + 5y^2 + 3z^2 + 2xy + 2yz + 2zx = 1$ into canonical form.

Let $A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Then the equation takes the form $X^tAX = 1$.

The characteristic equation of A is $\begin{vmatrix} 3-x & 1 & 1 \\ 1 & 5-x & 1 \\ 1 & 1 & 3-x \end{vmatrix} = 0.$

or, $(x-2)(x-3)(x-6) = 0$.
The eigen values of A are 2, 3, 6.

The eigen vectors corresponding to the eigen value 2 are $c\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, where $c \neq 0$.

The eigen vectors corresponding to the eigen value 3 are $c\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, where $c \neq 0$.

The eigen vectors corresponding to the eigen value 6 are $c\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, where $c \neq 0$.

Let us take one eigen vector corresponding to each distinct eigen value.

Let $\alpha = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\beta = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $\gamma = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$. Then the set $\{\alpha, \beta, \gamma\}$ is an orthogonal set. The orthonormal set of eigen vectors is $\left\{ \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$.

Let $P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$. Then P is an orthogonal matrix.

Let us apply the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$. Then the equation transforms to $(X')^t(P^tAP)X' = 1$.

$P^tAP(= P^{-1}AP)$ is a diagonal matrix D which has the same eigen values as those of A.

$$AP = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 5 & 1 \\ 1 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

$$= \begin{pmatrix} 2.\frac{1}{\sqrt{2}} & 3.\frac{1}{\sqrt{3}} & 6.\frac{1}{\sqrt{6}} \\ 2.0 & 3.\frac{-1}{\sqrt{3}} & 6.\frac{2}{\sqrt{6}} \\ 2.\frac{-1}{\sqrt{2}} & 3.\frac{1}{\sqrt{3}} & 6.\frac{1}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{-1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix} = PD,$$

where $D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$.

So $P^{-1}AP = D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}$.

The equation transforms to $(X')^tDX' = 1$, i.e., to $2x'^2 + 3y'^2 + 6z'^2 = 1$.

Note. Three mutually perpendicular principal directions are given by the vectors α, β, γ. We choose co-ordinate axes OX', OY', OZ' along the directions of α, β, γ respectively.

As $\alpha = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\beta = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $\gamma = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, the direction cosines of OX', OY', OZ' are $\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}}$; $\frac{1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$; $\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}$.

2. Reduce the equation $2xy + 2yz + 2zx = 1$ into canonical form.

Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Then the equation takes the form $X^tAX = 1$.

The characteristic equation of A is $\begin{vmatrix} -x & 1 & 1 \\ 1 & -x & 1 \\ 1 & 1 & -x \end{vmatrix} = 0.$

or, $(x-2)(x+1)(x+1) = 0$.
The eigen values of A are $2, -1, -1$.

The eigen vectors corresponding to the eigen value 2 are $c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, where $c \neq 0$.

Let $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ be an eigen vector corresponding to the eigen value -1. Then $AX = -X$. $(A + I_3)X = O$.

The eigen vectors are the non-null solutions of the system of equations

$$\begin{aligned} x + y + z &= 0 \\ x + y + z &= 0 \\ x + y + z &= 0. \end{aligned}$$

The system is equivalent to $x + y + z = 0$.

Let $y = c, z = d$, where c, d are arbitrary real numbers. Then $x = -c - d$.

The eigen vectors are $\begin{pmatrix} -c-d \\ c \\ d \end{pmatrix}$, i.e., $c\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + d\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, where $(c, d) \neq (0, 0)$.

Let $\alpha = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \beta = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. Let us choose c, d such that $\gamma = \begin{pmatrix} -c-d \\ c \\ d \end{pmatrix}$ is orthogonal to β. Then $2c+d=0$. Let $c=1$. Then $d = -2$ and $\gamma = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$. Then the set $\{\alpha, \beta, \gamma\}$ is an orthogonal set. The orthonormal set of eigen vectors is $\left\{ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\}$.

Let $P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix}$. Then P is an orthogonal matrix.

Let us apply the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$. Then the equation transforms to $(X')^t(P^tAP)X' = 1$.

$P^tAP(= P^{-1}AP)$ is a diagonal matrix D which has the same eigen values as those of A.

$$AP = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 2.\frac{1}{\sqrt{3}} & -1.\frac{-1}{\sqrt{2}} & -1.\frac{1}{\sqrt{6}} \\ 2.\frac{1}{\sqrt{3}} & -1.\frac{1}{\sqrt{2}} & -1.\frac{1}{\sqrt{6}} \\ 2.\frac{1}{\sqrt{3}} & -1.0 & -1.\frac{-2}{\sqrt{6}} \end{pmatrix} =$$

$$\begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = PD, \text{ where } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

So $P^tAP = D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

The equation transforms to $(X')^tDX' = 1$, i.e., to $2x'^2 - y'^2 - z'^2 = 1$.

Note. Three mutually perpendicular principal directions are given by the vectors α, β, γ. We choose co-ordinate axes OX', OY', OZ' along the directions of α, β, γ respectively. As β, γ can be chosen in many ways, three mutually perpendicular principal directions are not unique.

Remark. If A is a symmetric matrix, there exists a non-singular matrix P such that P^tAP is a diagonal matrix D. Therefore the equation $X^tAX = 1$ transforms to the canonical form (free from the product terms) by the transformation $X = PX'$. As P is non-singular (but not orthogonal), an orthonormal basis transforms to a non-orthonormal basis by the transformation $X = PX'$. As the deistance between two points does not remain invariant, the shape and size of the conic (or quadric) are changed.

For example, let us consider the equation $x^2 + 6xy - 7y^2 = 16$.

Taking $A = \begin{pmatrix} 1 & 3 \\ 3 & -7 \end{pmatrix}$ and $X = \begin{pmatrix} x \\ y \end{pmatrix}$, the matrix equation is $X^tAX = 16I_1$.

Appling congruence operations on A we have a non-singular matrix $P = \begin{pmatrix} 1 & -3 \\ 0 & 1 \end{pmatrix}$ such that $P^tAP = \begin{pmatrix} 1 & 0 \\ 0 & -16 \end{pmatrix}$.

By the transformation $X = PX'$, where $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ the equation reduces to $X''(P^tAP)X = 16I_1$, i.e., to $x'^2 - 16y'^2 = 16$.

This shows that the conic is a hyperbola.

Let us find an orthogonal matrix P such that P^tAP is a diagonal matrix.

The eigen values of A are $2, -8$.

The eigen vectors corresponding to the erigen value 2 are $c\begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $c \neq 0$. The eigen vectors corresponding to the erigen value -8 are $c\begin{pmatrix} -1 \\ 3 \end{pmatrix}$, $c \neq 0$.

The orthonormal set of eigen vectors is $\{\frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{10}}\begin{pmatrix} -1 \\ 3 \end{pmatrix}\}$.

Let $P = \frac{1}{\sqrt{10}}\begin{pmatrix} 3 & -1 \\ 1 & 3 \end{pmatrix}$. Then $P^tAP(= P^{-1}AP) = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$.

By the transformation $X = PX'$, where $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ the equation reduces to $X''(P^tAP)X = 16I_1$, i.e., to $2x'^2 - 8y'^2 = 16$.

This is a hyperbola whose principal axes are along the directions of the eigen vectors $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$.

Thus we observe that the equation $x^2 + 6xy - 7y^2 = 16$ transforms to two different forms. The non-orthogonal transformation does not preserve the shape of the conic.

4.26.2. Reduction to canonical forms.

A. The general equation of second degree in x, y is
$ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0.$

In matrix notation, $X^tAX + BX + cI_1 = O$ where
$A = \begin{pmatrix} a & h \\ h & b \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$, $B = (2g\ 2f)$.

Since A is a real symmetric matrix, there exists an orthogonal matrix P such that P^tAP is the diagonal matrix $D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ where λ_1, λ_2 are the eigen values of A.

Therefore by the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$, the equation transforms to $X'^tDX' + BPX' + cI_1 = O$

or, $\lambda_1 x'^2 + \lambda_2 y'^2 + 2g_1 x' + 2f_1 y' + c = 0$ ... (A)

The following cases come up for consideration.

Case 1. Rank $A = 2$.

In this case λ_1, λ_2 are both non-zero.

Let us transfer the origin to the point $(\frac{-g_1}{\lambda_1}, \frac{-f_1}{\lambda_2})$. The equation takes the form $\lambda_1 x^2 + \lambda_2 y^2 + d = 0$.

Sub case (i). $d = 0$. The equation represents

a pair of lines if $\lambda_1\lambda_2 < 0$;
a point ellipse if $\lambda_1\lambda_2 > 0$ and $\lambda_1 \neq \lambda_2$;
a point circle if $\lambda_1 = \lambda_2$.

Sub case (ii). $d \neq 0$. The equation represents

an ellipse if $\lambda_1\lambda_2 > 0, \lambda_1 d < 0$;
a conic without real trace if $\lambda_1\lambda_2 > 0, \lambda_1 d > 0$;
a hyperbola if $\lambda_1\lambda_2 < 0$.

Case 2. Rank $A = 1$.

In this case one of λ_1, λ_2 is zero. Let $\lambda_2 = 0$.

Let us transfer the origin to the point $\left(-\frac{g_1}{\lambda_1}, 0\right)$. The equation (A) takes the form $\lambda_1 x^2 + 2f_1 y + d = 0$. ... (B)

Sub case (i). $f_1 \neq 0$.

Let us transfer the origin to the point $\left(0, -\frac{d}{2f_1}\right)$. The equation (B) takes the form $\lambda_1 x^2 + 2f_1 y = 0$. This represents a parabola.

Sub case (ii). $f_1 = 0$. The equation represents

a pair of coincident lines if $d = 0$;
a pair of parallel lines if $\lambda_1 d < 0$;
a pair of imaginary lines if $\lambda_1 d > 0$.

B. The general equation of second degree in x, y, z is
$ax^2 + by^2 + cz^2 + 2hxy + 2fyz + 2gzx + 2ux + 2vy + 2wz + d = 0.$

In matrix notation, $X^tAX + BX + dI_1 = O$ where
$A = \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $B = (2u\ 2v\ 2w)$.

Since A is a real symmetric matrix, there exists an orthogonal matrix P such that P^tAP is the diagonal matrix

$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$, where $\lambda_1, \lambda_2, \lambda_3$ are the eigen values of A.

Therefore by the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$, the equation transforms to $X'^tDX' + BPX' + dI_1 = O$

or, $\lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2 + 2u_1 x' + 2v_1 y' + 2w_1 z' + d = 0.$ ...(A)

The following cases come up for consideration.

Case 1. Rank $A = 3$.

In this case $\lambda_1, \lambda_2, \lambda_3$ are all non-zero.

Let us transfer the origin to the point $(-u_1/\lambda_1, -v_1/\lambda_2, -w_1/\lambda_3)$.

The equation (A) takes the form $\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 + d_1 = 0$.

The equation represents

a cone if $d_1 = 0$;
a central quadric if $d_1 \neq 0$.

Case 2. Rank $A = 2$.

In this case one of $\lambda_1, \lambda_2, \lambda_3$ is zero. Let $\lambda_3 = 0$.

Let us transfer the origin to the point $(-u_1/\lambda_1, -v_1/\lambda_2, 0)$.

The equation (A) takes the form $\lambda_1 x^2 + \lambda_2 y^2 + 2wz + d_1 = 0$.

Sub case (i). $w_1 \neq 0$.

Let us transfer the origin to the point $(0, 0, -d_1/(2w_1))$. The equation takes the form $\lambda_1 x^2 + \lambda_2 y^2 + 2w_1 z = 0$. This represents a paraboloid.

Sub case (ii). $w_1 = 0$. The equation represents

a pair of planes (real or imaginary) if $d_1 = 0$;
a hyperbolic or elliptic cylinder if $d_1 \neq 0$.

Case 3. Rank $A = 1$.

In this case two of $\lambda_1, \lambda_2, \lambda_3$ are zero. Let $\lambda_2 = \lambda_3 = 0$.

Let us transfer the origin to the point $(-u_1/\lambda_1, 0, 0)$. The equation (A) takes the form $\lambda_1 x^2 + 2v_1 y + 2w_1 z + d_1 = 0$. $\quad \cdots$ (B)

Sub case (i). At least one of v_1, w_1 is non zero, say $v_1 \neq 0$.

Let us transfer the origin to the point $(0, -d_1/(2v_1), 0)$. The equation (B) takes the form $\lambda_1 x^2 + 2v_1 y + 2w_1 z = 0$. By the orthogonal transformation $X = PX'$ where $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{v_1}{r} & -\frac{w_1}{r} \\ 0 & \frac{w_1}{r} & \frac{v_1}{r} \end{pmatrix}$, $r = \sqrt{v_1^2 + w_1^2}$

the equation reduces to $\lambda_1 x^2 + 2ry = 0$. This represents a parabolic cylinder.

Sub case (ii). $v_1 = w_1 = 0$. The equation represents

a pair of coincident planes if $d_1 = 0$;
a pair of parallel planes if $d_1 \neq 0$.

Worked Examples (continued).

3. Reduce the equation $7x^2 - 2xy + 7y^2 - 16x + 16y - 8 = 0$ into canonical form and determine the nature of the conic.

Let $A = \begin{pmatrix} 7 & -1 \\ -1 & 7 \end{pmatrix}$, $B = (-16 \;\; 16)$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$.

Then the equation takes the form $X^t AX + BX - 8I_1 = O$.

The eigen values of A are 8, 6.

The eigen vectors corresponding the eigen values 8 and 6 are

$c\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $c \neq 0$; $d\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $d \neq 0$ respectively.

The orthonormal set of eigen vectors is $\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$.

Let $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Then P is an orthogonal matrix.

$P^t AP (= P^{-1}AP)$ is a diagonal matrix which has the same eigen values as those of A.

So $P^t AP = \begin{pmatrix} 8 & 0 \\ 0 & 6 \end{pmatrix}$. $BP = (-16\sqrt{2} \;\; 0)$.

By the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$, the equation transforms to $8x'^2 + 6y'^2 - 16\sqrt{2}x' - 8 = 0$

or, $8(x' - \sqrt{2})^2 + 6y'^2 = 24$.

Let us apply the translation $x'' = x' - \sqrt{2}$
$y'' = y'$.

The equation transforms to $8x''^2 + 6y''^2 = 24$.

The canonical form is $8x^2 + 6y^2 = 24$.

The equation represents an ellipse.

4. Reduce the equations

(i) $x^2 - 6xy + y^2 - 4x - 4y + 12 = 0$,

(ii) $x^2 - 6xy + y^2 - 4x - 4y - 4 = 0$

into canonical form and determine the nature of the conic.

(i) Let $A = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix}$, $B = (-4 \;\; -4)$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$.

Then the equation takes the form $X^t AX + BX + 12I_1 = O$.

The eigen values of A are 4, -2.

The eigen vectors corresponding to the eigen values 4 and -2 are

$c\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $c \neq 0$; $d\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $d \neq 0$ respectively.

The orthonormal set of eigen vectors is $\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$.

Let $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Then P is an orthogonal matrix.

$P^t AP (= P^{-1}AP)$ is a diagonal matrix whose eigen values are same as those of A.

So $P^t AP = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$. Also $BP = (0 \;\; -4\sqrt{2})$.

By the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$, the equation transforms to $4x'^2 - 2y'^2 - 4\sqrt{2}y' + 12 = 0$

or, $4x'^2 - 2(y' + \sqrt{2})^2 = -16$.

Let us apply the translation $x'' = x'$
$y'' = y' + \sqrt{2}$.

The equation transforms to $4x''^2 - 2y''^2 = -16$.
The canonical form is $2x^2 - y^2 = -8$.

The equation represents a hyperbola.

(ii) By the orthogonal transformation $X = PX'$, the equation transforms to $4x'^2 - 2y'^2 - 4\sqrt{2}y' - 4 = 0$
or, $4x'^2 - 2(y' + \sqrt{2})^2 = 0$.

Let us apply the translation $x'' = x'$
$y'' = y' + \sqrt{2}$.

The equation transforms to $4x''^2 - 2y''^2 = 0$.
The canonical form is $2x^2 - y^2 = 0$.

The equation represents a pair of lines.

5. Reduce the equations

(i) $9x^2 - 24xy + 16y^2 + 2x - 11y + 6 = 0$,

(ii) $9x^2 - 24xy + 16y^2 + 6x - 8y + 1 = 0$

into canonical form and determine the nature of the conic.

(i) Let $A = \begin{pmatrix} 9 & -12 \\ -12 & 16 \end{pmatrix}$, $B = (2 \ \ -11)$, $X = \begin{pmatrix} x \\ y \end{pmatrix}$.

Then the equation takes the form $X^tAX + BX + 6I_1 = O$.
The eigen values of A are 25, 0.
The eigen vectors corresponding to the eigen values 25 and 0 are $c\begin{pmatrix} 3 \\ -4 \end{pmatrix}$, $c \neq 0$; $d\begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $d \neq 0$ respectively.

The orthonormal set of eigen vectors is $\left\{\frac{1}{5}\begin{pmatrix} 3 \\ -4 \end{pmatrix}, \frac{1}{5}\begin{pmatrix} 4 \\ 3 \end{pmatrix}\right\}$.

Let $P = \frac{1}{5}\begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix}$. Then P is an orthogonal matrix.

$P^tAP (= P^{-1}AP)$ is a diagonal matrix whose diagonal elements are the eigen values of A.

Therefore $P^tAP = \begin{pmatrix} 25 & 0 \\ 0 & 0 \end{pmatrix}$ and $BP = (10 \ \ -5)$.

By the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$, the equation transforms to $25x'^2 + 10x' - 5y' + 6 = 0$
or, $25(x' + \frac{1}{5})^2 - 5(y' - 1) = 0$.

By the translation $x'' = x' + \frac{1}{5}$
$y'' = y' - 1$,

the equation transforms to $25x''^2 - 5y'' = 0$.
The canonical form is $25x^2 = 5y$.
The equation represents a parabola.

(ii) By the orthogonal transformation $X = PX'$, where $P = \frac{1}{5}\begin{pmatrix} 3 & 4 \\ -4 & 3 \end{pmatrix}$, $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ the equation transforms to $25x'^2 + 10x' + 1 = 0$ or, $25(x' + \frac{1}{5})^2 = 0$.

By the translation $x'' = x' + \frac{1}{5}$
$y'' = y'$,

the equation transforms to $25x''^2 = 0$.
The canonical form is $25x^2 = 0$.

The equation represents a pair of coincident lines.

6. Reduce the equations

(i) $2xy + 2yz + 2zx + x - 2y + z - 2 = 0$;

(ii) $2xy + 2yz + 2zx + x - y - \frac{1}{2} = 0$

into canonical form and determine the nature of the quadric.

(i) Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $B = (1 \ \ -2 \ \ 1)$.

The equation takes the form $X^tAX + BX + I_1 = O$.

The eigen values of A are 2, −1, −1.

The eigen vectors corresponding to the eigen value 2 are $c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $c \neq 0$.

The eigen vectors corresponding to the eigen value −1 are $c\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + d\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, where $(c, d) \neq (0, 0)$.

Let $\alpha = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\beta = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. Let us choose c, d such that $\gamma = \begin{pmatrix} -c-d \\ d \\ c \end{pmatrix}$ is orthogonal to β. Then $2c + d = 0$. Let $c = 1$. Then $d =$

-2 and $\gamma = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. Then the set $\{\alpha, \beta, \gamma\}$ is an orthogonal set. The orthonormal set of eigen vectors is $\left\{ \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$.

Let $P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$. Then P is an orthogonal matrix.

By the orthogonal transformation $X = PX'$, where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$, the equation transforms to $(X')^t(P^tAP)X' + BPX' - 2 = 0$.

$P^tAP(= P^{-1}AP)$ is a diagonal matrix D which has the same eigen values as those of A.

So $P^tAP = D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$. $BP = (0\ 0\ \sqrt{6})$.

The equation transforms to $2x'^2 - y'^2 - z'^2 + \sqrt{6}z' - 2 = 0$

or, $2x'^2 - y'^2 - (z' - \frac{1}{2}\sqrt{6})^2 - \frac{1}{2} = 0$.

Let us apply the translation
$$\begin{aligned} x'' &= x' \\ y'' &= y' \\ z'' &= z' - \tfrac{1}{2}\sqrt{6}. \end{aligned}$$

Then the equation transforms to $2x''^2 - y''^2 - z''^2 = \frac{1}{2}$.
The canonical form is $2x^2 - y^2 - z^2 = \frac{1}{2}$.

This represents a hyperboloid.

(ii) Let $B = (1\ -1\ 0)$. Then $BP = (0\ \ -\frac{1}{\sqrt{2}}\ \ \frac{3}{\sqrt{6}})$.

By the orthogonal transformation $X = PX'$, the equation transforms to $2x'^2 - y'^2 - z'^2 - \frac{1}{\sqrt{2}}y' + \frac{3}{\sqrt{6}}z' - \frac{1}{2} = 0$

or, $2x'^2 - (y' + \frac{1}{2\sqrt{2}})^2 - (z' - \frac{3}{2\sqrt{6}})^2 = 0$.

Let us apply the translation
$$\begin{aligned} x'' &= x' \\ y'' &= y' + \frac{1}{2\sqrt{2}} \\ z'' &= z' - \frac{3}{2\sqrt{6}}. \end{aligned}$$

The equation transforms to $2x''^2 - y''^2 - z''^2 = 0$.

The canonical form is $2x^2 - y^2 - z^2 = 0$.

The equation represents a cone.

7. Reduce the equations
(i) $x^2 + 2yz - 4x + 6y + 2z + 10 = 0$,
(ii) $x^2 + 2yz + 4x + 2y + 2z + 5 = 0$
into canonical form and determine the nature of the quadric.

(i) Let $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, $B = (-4\ 6\ 2)$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Then the equation takes the form $X^tAX + BX + 10I_1 = O$.
The eigen values of A are $1, 1, -1$.
The eigen vectors corresponding to the eigen values 1 and -1 are
$c\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + d\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, (c, d) \neq (0, 0)$; $k\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$, $k \neq 0$ respectively.

Let $\alpha = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and let us determine c, d such that α is orthogonal to $\beta = c\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + d\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} c+d \\ d \\ d \end{pmatrix}$. Then $c + d = 0$.

Taking $c = -1$ and $d = 1$, we have $\beta = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

The three mutually orthogonal eigen vectors are $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

The orthogonal set of eigen vectors is $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \right\}$.

Let $P = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$. Then P is an orthogonal matrix.

$P^tAP(= P^{-1}AP)$ is a diagonal matrix whose diagonal elements are the eigen values of A.

So $P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ and $BP = (-4\ \ 4\sqrt{2}\ \ -2\sqrt{2})$.

By the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$,

the equation transforms to

$x'^2 + y'^2 - z'^2 - 4x' + 4\sqrt{2}y' - 2\sqrt{2}z' + 10 = 0$

or, $(x' - 2)^2 + (y' + 2\sqrt{2})^2 - (z' + \sqrt{2})^2 = 0$.

Let us apply the translation
$$\begin{aligned} x'' &= x' - 2 \\ y'' &= y' + 2\sqrt{2} \\ z'' &= z' + \sqrt{2}. \end{aligned}$$

Then the equation transforms to $x''^2 + y''^2 - z''^2 = 0$.

The canonical form is $x^2 + y^2 - z^2 = 0$.

This represents a cone.

(ii) By the orthogonal transformation $X = PX'$, the equation transforms to $x'^2 + y'^2 - z'^2 + 4x' + 2\sqrt{2}y' + 5 = 0$

or, $(x' + 2)^2 + (y' + \sqrt{2})^2 - z'^2 = 1$.

By the translation
$$\begin{aligned} x'' &= x' + 2 \\ y'' &= y' + \sqrt{2} \\ z'' &= z', \end{aligned}$$

the equation transforms to $x''^2 + y''^2 - z''^2 = 1$.

The canonical form is $x^2 + y^2 - z^2 = 1$.

The equation represents a hyperboloid revolution.

8. Reduce the equations

(i) $2y^2 - 2xy - 2yz + 2zx - x - 2y + 3z - 2 = 0$,

(ii) $2y^2 - 2xy - 2yz + 2zx + x + y + z + 3 = 0$

into canonical form and determine the nature of the quadric.

(i) Let $A = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 0 \end{pmatrix}$, $B = (-1 \ \ -2 \ \ 3)$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Then the equation takes the form $X^tAX + BX - 2I_1 = O$.

The eigen values of A are 3, −1, 0.

The orthonormal eigen vectors corresponding to these are

$\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ respectively.

Let $P = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & \sqrt{3} & \sqrt{2} \\ -2 & 0 & \sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \end{pmatrix}$. Then $P^tAP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

$BP = (\sqrt{6} \ \ -2\sqrt{2} \ \ 0)$.

By the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$, the equation transforms to $3x'^2 - y'^2 + \sqrt{6}x' - 2\sqrt{2}y' - 2 = 0$

or, $3(x' + 1/\sqrt{6})^2 - (y' + 2\sqrt{2})^2 = \frac{1}{2}$.

Let us apply the translation
$$\begin{aligned} x'' &= x' + \tfrac{1}{\sqrt{6}} \\ y'' &= y' + \sqrt{2} \\ z'' &= z'. \end{aligned}$$

The equation transforms to $3x''^2 - y''^2 = \frac{1}{2}$.

The canonical form is $3x^2 - y^2 = \frac{1}{2}$.

The equation represents a hyperbolic cylinder.

(ii) By the orthogonal transformation $X = PX'$, the equation transforms to $3x'^2 - y'^2 + \sqrt{3}z' + 3 = 0$.

Let us apply the translation
$$\begin{aligned} x'' &= x' \\ y'' &= y' \\ z'' &= z' + \sqrt{3} \end{aligned}$$

The equation transforms to $3x''^2 - y''^2 + \sqrt{3}z = 0$.

The canonical form is $3x^2 - y^2 + \sqrt{3}z = 0$.

The equation represents a hyperbolic paraboloid.

9. Reduce the equations

(i) $x^2 + y^2 + z^2 - 2xy - 2yz + 2zx + x - 4y + z + 1 = 0$,

(ii) $x^2 + y^2 + z^2 - 2xy - 2yz + 2zx + 2x - 2y + 2z - 1 = 0$

into canonical form and determine the nature of the quadric.

(i) Let $A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$, $B = (1 \ \ -4 \ \ 1)$, $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

Then the equation takes the form $X^tAX + BX + I_1 = O$.

The eigen values of A are 3, 0, 0.

The eigen vectors corresponding to the eigen value 3 are $k\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, $k \neq 0$,

and those corresponding to the eigen value 0 are $c\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + d\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $(c, d) \neq (0, 0)$.

Let $\alpha = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ and let us find c, d such that $\beta = c\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + d\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ may be orthogonal to α. Then $2c + d = 0$.

Taking $c = 1, d = -2$, we have $\beta = \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$.

Therefore three mutually orthogonal eigen vectors are $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}$ and the corresponding orthonormal set is

$$\left\{ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} \right\}.$$

Let $P = \frac{1}{\sqrt{6}} \begin{pmatrix} \sqrt{2} & \sqrt{3} & 1 \\ -\sqrt{2} & \sqrt{3} & -1 \\ \sqrt{2} & 0 & -2 \end{pmatrix}$.

Then $P^t AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $BP = \left(\frac{6}{\sqrt{3}} \;\; -\frac{3}{\sqrt{2}} \;\; \frac{3}{\sqrt{6}}\right)$.

By the orthogonal transformation $X = PX'$ where $X' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}$, the equation transforms to $3x'^2 + \frac{6}{\sqrt{3}}x' - \frac{3}{\sqrt{2}}y' + \frac{3}{\sqrt{6}}z' + 1 = 0$

or, $3(x' + \frac{1}{\sqrt{3}})^2 - \frac{3}{\sqrt{2}}y' + \frac{3}{\sqrt{6}}z' = 0$.

By the translation
$$x'' = x' + \tfrac{1}{\sqrt{3}}, \quad y'' = y', \quad z'' = z',$$
the equation transforms to $3x''^2 - \frac{3}{\sqrt{2}}y'' + \frac{3}{\sqrt{6}}z'' = 0$.

Let us apply the orthogonal transformation

$x'' = x_1$

$y'' = y_1 \cos\theta - z_1 \sin\theta$

$z'' = y_1 \sin\theta + z_1 \cos\theta$

so that $-\frac{3}{\sqrt{2}}\cos\theta + \frac{3}{\sqrt{6}}\sin\theta = 0$. This gives $\theta = \frac{\pi}{3}$.

The equation transforms to $3x_1^2 + \sqrt{6}z_1 = 0$.
The canonical form is $3x^2 + \sqrt{6}z = 0$.

The equation represents a parabolic cylinder.

(ii) By the orthogonal transformation $X = PX'$, the equation transforms to $3x'^2 + 2\sqrt{3}x' - 1 = 0$

or, $3(x' + \frac{1}{\sqrt{3}})^2 - 2 = 0$.

Let us apply the translation
$$x'' = x' + \tfrac{1}{\sqrt{3}}, \quad y'' = y', \quad z'' = z'.$$

Then the equation transforms to $3x''^2 - 2 = 0$.
The canonical form is $3x^2 - 2 = 0$.

The equation represents a pair of parallel planes.

Exercises 20

1. Reduce the following equations into canonical form. Determine the nature of the conic.

(i) $7x^2 - 2xy + 7y^2 + 6x + 6y - 1 = 0$,

(ii) $x^2 - 3xy + y^2 + 10x - 10y + 50 = 0$,

(iii) $4x^2 - 4xy + y^2 - 12x + 6y + 9 = 0$,

(iv) $x^2 - 3xy + y^2 + 2x + 2y - 4 = 0$,

(v) $2x^2 + 2y^2 - 3x + y + 1 = 0$,

(vi) $4x^2 - 4xy + y^2 + x + 2y + 1 = 0$,

(vii) $7x^2 - 2xy + 7y^2 + 4x - 4y + 1 = 0$,

(viii) $14x^2 - 4xy + 11y^2 + 4x + 8y + 3 = 0$.

2. Reduce the following equations into canonical form. Determine the nature of the quadric.

(i) $2x^2 + 5y^2 + 10z^2 + 4xy + 12yz + 6zx - 1 = 0$,

(ii) $x^2 + y^2 + z^2 - 2xy - 2yz - 2zx - 1 = 0$,

(iii) $2xy + 2yz + 2zx - 1 = 0$,

(iv) $4x^2 - 4xy - 4yz + 4zx - 2x + 2y - 1 = 0$,

(v) $x^2 - y^2 + 4yz + 4zx + 2x + y + 2z + 1 = 0$,

(vi) $x^2 - y^2 + 4yz + 4zx + 2x - 2y - z - 3 = 0$,

(vii) $x^2 - y^2 - z^2 + 2yz + 4y - 4z - 4 = 0$,

(viii) $x^2 - y^2 - z^2 + 2yz + 2x + y + z + 1 = 0$,

(ix) $x^2 + y^2 + z^2 - 2xy - 2yz + 2zx + x - y - 2z + 6 = 0$,

(x) $2x^2 - 2xy - 2yz + 2zx - x - z = 0$.

ANSWERS TO EXERCISES

Part I

Exercises 1 (Page 14)

1. (a) (i) $\{x \in \mathbb{Z} : 0 \leq x \leq 15\}$, (ii) $\{x \in \mathbb{Z} : 5 \leq x \leq 10\}$, (iii) ϕ, (iv) ϕ.

3. (i) $C - A$, (ii) $B \cap C$, (iii) A. **9.** (i) $2^{10}(2^{10} - 1)$, (ii) $2^{20}(2^5 - 1)$.

Exercises 2 (Page 31)

1. (i) reflexive and transitive (ii) reflexive and symmetric (iii) RST (iv) RST.

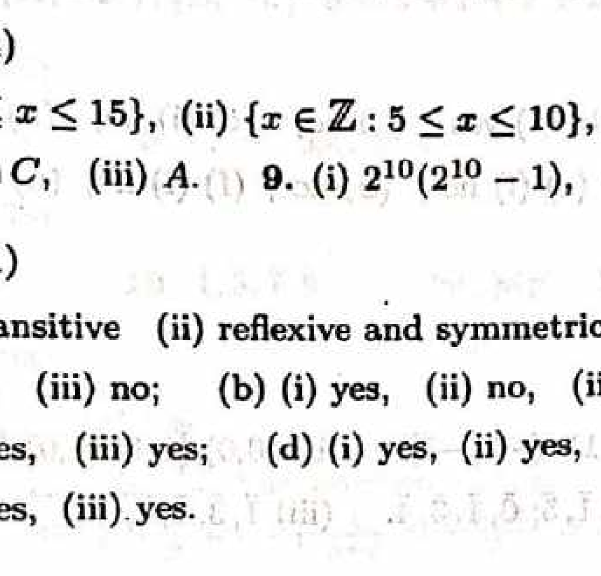
2. (a) (i) no, (ii) yes, (iii) no; (b) (i) yes, (ii) no, (iii) no;
   (c) (i) yes, (ii) yes, (iii) yes; (d) (i) yes, (ii) yes, (iii) no;
   (e) (i) yes, (ii) yes, (iii) yes.

3. (a) yes, (b) no.

4. $cl(0) = \{5n : n \in \mathbb{Z}\}$, $cl(1) = \{5n \pm 1, n \in \mathbb{Z}\}, cl(2) = \{5n \pm 2 : n \in \mathbb{Z}\}$.

7. (i) 16, 4.

10.



11. 6, 7, 8, 9, 10.

Exercises 3 (Page 48)

6. (i) $f^{-1}(x) = \frac{1}{2}(x - 3), x \in \mathbb{R}$; (ii) $f^{-1}(x) = 3\sqrt{x}, x \in \mathbb{R}$;
   (iii) $f^{-1}(x) = \frac{x}{1+|x|}, x \in \mathbb{R}$; (iv) $f^{-1}(x) = \frac{x+1}{x-1}, x \in \mathbb{R} - 1$.

7. (i) $g \circ f(n) = 2n^2$ and $f \circ g(n) = 4n^2, n \in \mathbb{Z}$.
   (ii) $g \circ f(n) = 2(-1)^n$ and $f \circ g(n) = 1, n \in \mathbb{Z}$.
   (iii) $g \circ f(x) = 0, x \in \mathbb{R}$; $f \circ g(x) = 0, x \geq 0$
   $= -4x, x < 0$.

12. Let $f : \mathbb{Z} \to \mathbb{Z}$ be defined by $f(x) = x^2, x \in \mathbb{Z}$. Let $P = \{1, 2\}$. **14.** 36.

Exercises 4 (Page 57)

1. $fg = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 5 & 6 & 4 & 3 \end{pmatrix}$, $gf = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 2 & 3 & 1 \end{pmatrix}$

$f^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 3 & 2 & 4 & 5 \end{pmatrix}$, $g^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 5 & 3 & 4 & 2 \end{pmatrix}$.

2. f is even, g is odd. **3.** (i) $3 \to 5$, $4 \to 2$; (ii) $3 \to 1$, $4 \to 2$.

4. (i) 6, (ii) 6. **5.** (i) (1432), (ii) (12435), (iii) (165432).

6. 20. **7.** If not, each $(a_i - b_i)$ must be odd and since n is odd, $\Sigma(a_i - b_i)$ is odd. But $\Sigma(a_i - b_i) = \Sigma a_i - \Sigma b_i = 0$, an even number.

Exercises 6 (Page 67)

1. (a) (i) yes, (ii) yes; (b) (i) yes, (ii) no; (c) (i) no, (ii) no; (d) (i) yes, (ii) yes; (e) (i) no, (ii) no; (f) (i) no, (ii) no.

3. 16, 8.

Exercises 7 (Page 74)

1. (i) yes, (ii) no.

5. (i) $(1,1), (1,-1), (-1,1), (-1,-1)$. (ii) $(0,0), (0,1), (1,0), (1,1)$.

6. (i) $\bar{1}, \bar{2}, \bar{3}, \bar{4}; \bar{0}, \bar{1}$. (ii) $\bar{1}, \bar{5}; \bar{0}, \bar{1}, \bar{3}, \bar{4}$. (iii) $\bar{1}, \bar{3}, \bar{5}, \bar{7}; \bar{1}$.

Exercises 8 (Page 84)

1. (i) yes, (ii) no, (iii) no, (iv) no, (v) yes, (vi) no.

2. (i) no, (ii) no, (iii) yes. **18.** I_a is the identity mapping on G.

Exercises 9 (Page 94)

8. no. **9.** no. **10.** $b * b = c$, $b * c = a$, $c * b = a$, $c * c = b$.

Exercises 10 (Page 102)

7. (ii) In $GL(2, \mathbb{R})$, $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$. **8.** (i) 25, (ii) 15.

9. 4. **10.** $\bar{3}, \bar{9}, \overline{21}, \overline{27}$. **11.** (i) 10, (ii) 5, (iii) 4. **12.** 7.

Exercises 12 (Page 123)

8. no. **11.** 5. **14.** $\{1, a^5, a^{10}, a^{15}, a^{20}, a^{25}\}$; a^5, a^{25}.

Exercises 13 (Page 134)

2. (i) Two left cosets $eH = aH = \{e, a\}$, $bH = cH = \{b, c\}$; two right cosets $He = Ha = \{e, a\}$, $Hb = Hc = \{b, c\}$.

(ii) Two left cosets $H, \{\rho_3, \rho_4, \rho_5\}$; two right cosets $H, \{\rho_3, \rho_4, \rho_5\}$.

Exercises 14 (Page 148)

25. H, aH, a^2H, a^3H, a^4H, a^5H. $G/H = \langle aH \rangle = \langle a^5H \rangle$.

Exercises 15 (Page 175)

1. (i) yes, $Ker\ \phi = \{0\}$, $Im\ \phi = G$. (ii) yes, $Ker\ \phi = \{\bar{0}, \bar{4}\}$, $Im\ \phi = \{\bar{0}, \bar{2}, \bar{4}, \bar{6}\}$. (iii) no. (iv) Yes, $Ker\ \phi = \{z \in \mathbb{C}^* : |z| = 1\}$, $Im\ \phi = G'$.

13. $\pi_e, \pi_a, \pi_b, \pi_c$, where $\pi_e = \begin{pmatrix} e & a & b & c \\ e & a & b & c \end{pmatrix}$, $\pi_a = \begin{pmatrix} e & a & b & c \\ a & e & c & b \end{pmatrix}$, etc.

15. (i) yes, (ii) yes, (iii) no, (iv) yes, (v) no.

Exercises 16 (Page 181)

3. $\{i\}, \{r_2\}, \{r_1, r_3\}, \{h, v\}, \{d, d'\}$; $8 = 2 + 2 + 2 + 2$, here $o(Z) = 2$.

4. $\{I\}, \{-I\}, \{J, -J\}, \{K, -K\}, \{L, -L\}$; $8 = 2 + 2 + 2 + 2$, here $o(Z) = 2$.

Exercises 17 (Page 189)

6. 24. **7.** 8. **8.** 6. **9.** 6. **14.** 6; 4.

Exercises 18 (Page 208)

4. no. **15.** $\bar{1}, \bar{3}, \bar{5}, \bar{7}$. **16.** $\bar{1}, \bar{3}, \bar{7}, \bar{9}$. **19.** $p(p-1)^2(p+1)$.

Exercises 19 (Page 224)

2. 2, 2^{-1}. **5.** (i) no, (ii) yes.

11. Consider the field $\mathbb{Z}_p$. In $\mathbb{Z}_p$, the inverses of $1, 2, \ldots, p-1$ are just the same elements in some order. So $1 + \frac{1}{2} + \cdots + \frac{1}{p-1} = \frac{p(p-1)}{2} = 0$ in $\mathbb{Z}_p$, since p is odd.

Exercises 20 (Page 231)

4. $\{\bar{0}\}, \{\bar{0}, \bar{3}, \bar{6}, \bar{9}\}, \{\bar{0}, \bar{5}, \overline{10}\}, \mathbb{Z}_{15}$. **9.** (i) no, (ii) yes.

Exercises 21 (Page 250)

7. (i) $\mathbb{Z}$, (ii) $2\mathbb{Z}$, (iii) $6\mathbb{Z}$, (iv) $20\mathbb{Z}$. **11.** $\bar{2}, \bar{3}; \bar{2}, \bar{3}$.

Exercises 22 (Page 261)

1. (i) $1, -1$; (ii) $1, -1$; **7.** (i) $2 + i$, (ii) $2 + i$, (iii) 1.

Exercises 23 (Page 274)

4. (i) $\mathbb{Z}$; (ii) $4\mathbb{Z}$; (iii) $15\mathbb{Z}$; (iv) $24\mathbb{Z}$.

11. (i) $2+3i[= (3+11i).1 + (8-i).(-i)]$, (ii) $2+i[= (11+3i)(-1-i) + (1+8i)(2-i)]$, (iii) $1 + 2i[= (7+4i)(-1) + (4+3i).2]$, (iv) $2 + i[= (4+7i)(-1) + (3+4i).2]$.

Exercises 24 (Page 284)

1. (i) no, (ii) no, (iii) yes, (iv) no. **4.** (i) $6\mathbb{Z}$, (ii) $\{0\}$. **8.** (i) yes, (ii) no.

Exercises 25 (Page 299)

4. (i) yes, (ii) yes, (iii) yes. **5.** (i) no, (ii) no, (iii) no.

Exercises 26 (Page 323)

1. (i) $x'yz' + x'yz + xy'z' + xy'z$, (ii) $x'yz + xy'z' + xy'z + xyz' + xyz$,

(iii) $xyz + x'y'z'$, (iv) $x'yz' + x'yz + xy'z + xyz'$.

2. (i) $(x+y+z)(x+y+z')(x'+y'+z)(x'+y'+z')$,

(ii) $(x+y+z)(x+y+z')(x+y'+z)$,

(iii) $(x+y+z')(x+y'+z)(x+y'+z')(x'+y+z)(x'+y+z')(x'+y'+z)$,

(iv) $(x+y+z)(x+y+z')(x'+y+z)(x'+y'+z')$.

3. (i) $x'yz + xy'z + xy'z'$; $(x+y+z)(x+y+z')(x+y'+z)(x'+y+z)(x'+y'+z')$

(ii) $x'y'z + xy'z' + x'yz' + x'y'z$; $(x'+y'+z')(x+y'+z')(x'+y'+z)(x'+y+z')$.

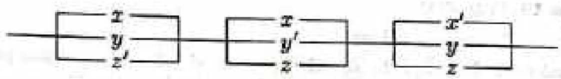
4. (i) $x'y'z' + x'y'z + x'yz' + xy'z'$; $(x'+y'+z)(x+y'+z')(x'+y+z')(x'+y'+z')$

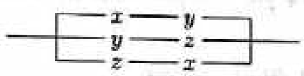
(ii) $x'yz + xy'z + xyz'$; $(x+y+z)(x+y+z')(x+y'+z)(x'+y+z)(x'+y'+z')$

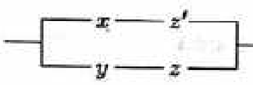
(iii) $x'y'z' + xyz$; $(x+y+z')(x+y'+z)(x'+y+z)(x'+y'+z)(x+y'+z')(x'+y+z')$

(iv) $x'y'z' + x'y'z + x'yz' + xy'z'$; $(x'+y'+z)(x+y'+z')(x'+y+z')(x'+y'+z')$.

Exercises 28 (Page 339)

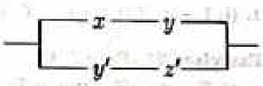
1. (i) 

(ii) 

2 (i) 

(ii) 

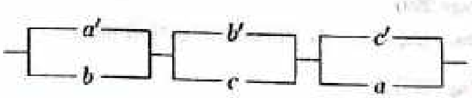
(iii) 

(iv) 

3.(i)

(ii) 

4.(i) 

(ii) 

Part II

Exercises 1 (Page 18)

1. (i) $A = B = \frac{1}{5}I_2$; (ii) $A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$;

(iii) $A = \begin{pmatrix} 1 & 2 \\ 4 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}$. **8.** $\begin{pmatrix} 1 & 0 \\ -50 & 1 \end{pmatrix}$.

10. (i) $\frac{1}{5}\begin{pmatrix} -3 & 2 & 2 \\ 2 & -3 & 2 \\ 2 & 2 & -3 \end{pmatrix}$, (ii) $\frac{1}{9}\begin{pmatrix} -2 & 4 & 1 \\ 1 & -2 & 4 \\ 4 & 1 & -2 \end{pmatrix}$.

11. $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, where $a^2 + bc = 0$. **12.** $\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, where $a^2 + bc = 1, I_2, -I_2$.

13. In all matrices m, n are real numbers.

(i) $\begin{pmatrix} m & 3n \\ n & m+2n \end{pmatrix}$, (ii) $\begin{pmatrix} m & 3n \\ n & m \end{pmatrix}$, (iii) $\begin{pmatrix} m & bn \\ cn & m \end{pmatrix}$,

(iv) $\begin{pmatrix} m & bn \\ cn & m + (d-a)n \end{pmatrix}$. **14.** $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ or $\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$.

17. (i) $P = \begin{pmatrix} 1 & 5 & 3 \\ 5 & 2 & 7 \\ 3 & 7 & 1 \end{pmatrix}$, $Q = \begin{pmatrix} 0 & -2 & 1 \\ 2 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$;

(ii) $P = \begin{bmatrix} 2 & 3 & 2 & 2 \\ 3 & 3 & 2 & 7 \\ 2 & 2 & 1 & 6 \\ 2 & 7 & 6 & 0 \end{bmatrix}$, $Q = \begin{bmatrix} 0 & -2 & -2 & 1 \\ 2 & 0 & 0 & 1 \\ 2 & 0 & 0 & -1 \\ -1 & -1 & 1 & 0 \end{bmatrix}$.

Exercises 2 (Page 63)

19. (i) 1, 2, 3; (ii) $\frac{(b-1)(c-1)}{(b-a)(c-a)}, \frac{(c-1)(a-1)}{(c-b)(a-b)}, \frac{(a-1)(b-1)}{(a-c)(b-c)}$.

20. $\frac{b^2+c^2-a^2}{2bc}, \frac{c^2+a^2-b^2}{2ca}, \frac{a^2+b^2-c^2}{2ab}$.

Exercises 3 (Page 83)

4. $\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$, $\frac{1}{2}\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$.

5. $\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 3 \end{pmatrix}$. **6.** $\begin{pmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 2 & 3 & 4 \end{pmatrix}$, $\begin{pmatrix} -1 & 0 & -1 \\ -3 & -4 & -5 \\ -2 & -3 & -4 \end{pmatrix}$.

7. $\begin{pmatrix} 1 & 1 & -1 \\ 5 & 9 & -8 \\ 3 & 6 & -5 \end{pmatrix}$. **9.** $\frac{1}{8}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}$.

12. (i) 1, 1, 2; (ii) 2, 1, 1. **13.** $\frac{3}{7}$. **14.** $\frac{4}{5+x}$.

15. (i) Hint. Let $A = (a_{ij})_{n,n}, A^{-1} = B = (b_{ij})_{n,n}$. Use $BA = I_n$.
(ii) Hint. Let $A = (a_{ij})_{n,n}, A^{-1} = B = (b_{ij})_{n,n}$. Use $AB = I_n$.

16. 2, 1, 1.

28. (i) $\frac{1}{\sqrt{6}}\begin{pmatrix} \sqrt{2} & -\sqrt{2} & \sqrt{2} \\ \sqrt{3} & \sqrt{3} & 0 \\ -1 & 1 & 2 \end{pmatrix}$, (ii) $\begin{pmatrix} 1/3 & 2/\sqrt{5} & 2/\sqrt{45} \\ -2/3 & 0 & 5/\sqrt{45} \\ 2/3 & -1/\sqrt{5} & 4/\sqrt{45} \end{pmatrix}$.

Exercises 4 (Page 103)

1. (i) 4, (ii) 4, (iii) 4, (iv) 4. **5.** (i) $1, -3$, (ii) $0, -10$, (iii) $1, 2, 3, -10$.

6. (i) 1, (ii) $3, -1$. **7.** (i) 2, (ii) 1.

8. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -2 & 0 \\ -1 & 1 & 1 \end{pmatrix}$, $Q = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

(ii) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 0 & 1 \\ -2 & 1 & 0 \end{pmatrix}$ $Q = \begin{pmatrix} 1 & -2 & 0 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

9. (i) $\frac{1}{12}\begin{pmatrix} 6 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, (ii) $\frac{1}{6}\begin{pmatrix} 3 & 0 & 0 \\ -4 & 2 & 0 \\ -2 & -8 & 6 \end{pmatrix}$, (iii) $\begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.

10. (i) $E_{21}(2)E_{12}(2)$, (ii) $E_2(3)\ E_{12}(2)\ E_3(2)\ E_{13}(-10/3)\ E_{23}(5/3)$,
(iii) $E_{21}(1)\ E_{31}(1)\ E_{32}(1)\ E_{13}(1)$.

11. (ii) **Hint.** Let $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

(iv) **Hint.** Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Exercises 5 (Page 110)

1. (i) 3, 3; (ii) 2, 2; (iii) 2, 2; (iv) 2, 0; (v) $3, -1$; (vi) $3, -1$.

2. (i) yes, (ii) no.

3. (i) $P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$, $P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix}$

(ii) $P^tAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$.

4. (i) $P^tAP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & \frac{5}{2} \end{pmatrix}$, $P = \begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$

(ii) $P^tAP = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $P = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$.

Exercises 6 (Page 125)

2. (i) yes, (ii) no, (iii) no, (iv) yes, (v) no, (vi) yes, (vii) no.

3. (i) yes, (ii) no, (iii) yes. **4.** $\gamma \notin L\{\alpha, \beta\}$, $\delta \in L\{\alpha, \beta\}$.

10. (i) yes, (ii) yes, (iii) no, (iv) - (viii) yes.

Exercises 7 (Page 144)

3. (i) $2, -1$; (ii) $1, -\frac{1}{2}$. **4.** (i) $a = b$, $a = -2b$; (ii) $a + b + 1 = 0$, $a = b = 1$.

9. (i) $k \neq 1, -2$; (ii) $k \neq 1, -3$.

17. A basis if n be odd; not a basis if n be even.

18. (i) 2, (ii) 1. **19.** (i) 3, (ii) 2. **20.** 2.

21. (i) $\left\{\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}\right\}$, (ii) $\left\{\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}\right\}$.

22. (i) $\left\{\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$, 3;

(ii) $\left\{\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right\}$, 2; (iii) $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$, 2;

(iv) $\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}$, 3; (v) $\left\{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right\}$, 1.

24. (i) $(1, -1, 2)$, (ii) $(0, 1, 1)$. **25.** 2. **26.** infinite.

Exercises 8 (Page 152)

1. 2, 2, 1, 3. **2.** 2, 2, 1, 3. **3.** 3, 3, 2, 4. **4.** $P = L\{(1, 0, 0)\}$, $Q = L\{0, 1, 0)\}$.

Exercises 9 (Page 161)

1. (i) 2, (ii) 3, (iii) 3. **2.** $\{(1, 0, -\frac{2}{3}), (0, 1, \frac{7}{3})\}$. **3.** $\{(1, 0, -1), (0, 1, 1)\}$.

4. (i) no, yes; (ii) yes, no. **11.** $\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

Exercises 10 (Page 177)

1. (i) $c(2, -5, 1), c \in \mathbb{R}$, (ii) $c(1, 0, 0, 1) + d(0, 1, 1, 0), c, d \in \mathbb{R}$.

2. (i) $c(1, 0, -2) + d(0, 1, -3)$, $c, d \in \mathbb{Q}$. (ii) $c(1, -1, 3)$, $c \in \mathbb{Q}$.

3. (i) $c(3, 1, -5)$, c is an integer. (ii) $c(1, 7, 5)$, c is an integer.

4. $x_1 - x_2 + x_3 - x_4 = 0$.

5. $x_1 - 2x_2 + x_3 = 0$
$10x_1 - 7x_2 + x_4 = 0$.

6. (i) $(1, 1, -1)$, (ii) inconsistent.

7. (i) $a = 1, \frac{2}{3}$; $c(2, 1, -1) + (1, 0, 0)$, $c(2, 1, -1) + (\frac{8}{9}, -\frac{1}{9}, 0)$, $c \in \mathbb{R}$.
(ii) $a = 1, -2$; $c(-4, 3, 1) + (1, 0, 0)$; $c(-4, 3, 1) + (4, -3, 0)$, $c \in \mathbb{R}$.

8. (i) $k = 3, 0$; $c(1, 1, 1)$, $c(-1, 1, 0) + d(-1, 0, 1)$, $c, d \in \mathbb{R}$
(ii) $k = 6, 2, -1$; $c(1, 1, 1)$, $c(3, 3, -5)$, $c(5, -2, -2)$, $c \in \mathbb{R}$.

9. (i) $a \neq -4$; $a = -4$, $b \neq 4$; $a = -4$, $b = 4$.
(ii) $a \neq 8$; $a = 8$, $(b-3)(b+1) \neq 0$; $a = 8$, $b = 3$ or $a = 8$, $b = -1$.

10. $x_1 = \frac{b+c-a}{2}, x_2 = \frac{c+a-b}{2}, x_3 = \frac{a+b-c}{2}$. $A^{-1} = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$.

11. $x_1 = \frac{b+c}{2}, x_2 = \frac{c+a}{2}, x_3 = \frac{a+b}{2}$. $A^{-1} = \frac{1}{2}\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$.

12. (i) $\begin{bmatrix} 3 \\ -2 \\ -2 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}; \begin{bmatrix} 3 & -2 & -1 \\ -2 & 2 & 1 \\ -2 & 1 & 1 \end{bmatrix}$.

Exercises 11 (Page 183)

1. concurrent, (ii) form a prism, (iii) two are parallel and the other intersects them, (iv) intersect along a line, (v) intersect along a line.

3. (i) $k = 9$, $\frac{x-1}{1} = \frac{y+2}{-1} = \frac{z}{-1}$. (ii) $k = 4$, $\frac{x-1}{3} = \frac{y-1}{-5} = \frac{z}{2}$.

4. (i) $k \neq \frac{9}{2}$; (ii) $k \neq 3, -1$.

Exercises 12 (Page 199)

1. (i) no, (ii) no, (iii) no, (iv) yes. **3.** $\frac{2}{7}(2,3,-1)$, $-\frac{5}{21}(1,-2,4)$, $\frac{1}{3}(2,-1,1)$.

4. (i) $\{(1,2,-2), (2,0,1), \frac{1}{15}(-2,5,4)\}$; (ii) $\{(1,1,0), \frac{1}{2}(-1,-1,2), \frac{1}{3}(2,-2,2)\}$.

5. (i) $\{(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0), (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0,0,1)\}$;
(ii) $\{\frac{1}{\sqrt{3}}(1,1,-1), \frac{1}{\sqrt{6}}(2,-1,1), \frac{1}{\sqrt{2}}(0,1,1)\}$.

6. (i) $\{\frac{1}{\sqrt{3}}(1,1,0,1), \frac{1}{\sqrt{42}}(4,-5,0,1), \frac{1}{\sqrt{105}}(4,2,7,-6)\}$.
(ii) $\{\frac{1}{4}(1,1,1,1), \frac{1}{2}(1,1,-1,-1), \frac{1}{\sqrt{10}}(-1,1,-2,2)\}$.

7. (i) $\{\frac{1}{\sqrt{6}}(1,0,1,2), \frac{1}{2\sqrt{3}}(1,3,1,-1)\}$.
(ii) $\{\frac{1}{\sqrt{2}}(1,0,0,-1), \frac{1}{\sqrt{6}}(1,2,0,1), (0,0,1,0)\}$.

8. (i) $L\{(1,-1,1,0), (4,-2,0,1)\}$; (ii) $L\{(1,1,-1)\}$.

Exercises 13 (Page 215)

1. (i) $A^2 - 5A + 5I_2$, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ -1 & -3 & 2 \end{pmatrix}$

(ii) $\frac{1}{2}A^2 + \frac{1}{9}A - \frac{8}{9}I_3, \frac{1}{9}\begin{pmatrix} -2 & 5 & 3 \\ 3 & -3 & 0 \\ 5 & 1 & -3 \end{pmatrix}$; **2.** $\begin{pmatrix} 1 & 0 & 0 \\ 50 & 1 & 0 \\ 50 & 0 & 1 \end{pmatrix}$.

3. (i) 1, 1, 1; all non-null vectors in $\mathbb{R}^3$. (ii) c, c, c; all non-null vectors in $\mathbb{R}^3$.

(iii) d_1, d_2, d_3; $c\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, c \neq 0$; $c\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, c \neq 0$; $c\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, c \neq 0$.

5. (i) $\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_n^{-1}$; (ii) $c\lambda_1^{-1}, c\lambda_2^{-1}, \ldots, c\lambda_n^{-1}$, where $c = \lambda_1\lambda_2\ldots\lambda_n$.

7. (i) 2, 3, 5; $c\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, c \neq 0$; $c\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, c \neq 0$; $c\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, c \neq 0$.

(ii) 5, 0, 0; $c\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, c \neq 0$; $c\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + d\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, (c,d) \neq (0,0)$.

(iii) 1, 1, 5; $c\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + d\begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, (c,d) \neq (0,0)$; $c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, c \neq 0$.

8. (i) 2, 0; $c\begin{pmatrix} 1 \\ i \end{pmatrix}, c\begin{pmatrix} i \\ 1 \end{pmatrix}, c \neq 0$. (ii) 1, −1; $c\begin{pmatrix} 1 \\ i \end{pmatrix}, c\begin{pmatrix} i \\ 1 \end{pmatrix}, c \neq 0$.

9. eigen value 1, algebraic multiplicity 2, geometric multiplicity 1;
eigen value −1, algebraic multiplicity 1, geometric multiplicity 1.

10. $a = 1$, $b = 2$, $c = 2$.

23. $\begin{pmatrix} 0 & 1 & -1 \\ -2 & 2 & 0 \\ -2 & 1 & 1 \end{pmatrix}$. **24.** $\begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{pmatrix}$.

Exercises 14 (Page 226)

1. (i) $\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$, (iii) $\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

4. $\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

5. (i) $P = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$, (ii) $P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}$

6. (i) $P = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{pmatrix}$, (ii) $P = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}$.

Exercises 15 (Page 235)

3. (i) $x^2 + y^2 + z^2, 3, 3$; (ii) $x^2 + y^2 - z^2, 3, 1$;
(iii) $x^2 + y^2, 2, 2$; (iv) $x^2 - y^2 - z^2, 3, -1$.

4. (i) $\begin{aligned} x_1 &= y_1 - y_2 - y_3 \\ x_2 &= y_2 + y_3 \\ x_3 &= y_3. \end{aligned}$ (ii) $\begin{aligned} x &= x' + z' \\ y &= \tfrac{1}{2}y' + \tfrac{1}{2}z' \\ z &= -\tfrac{1}{2}y' + \tfrac{1}{2}z'. \end{aligned}$

8. (i) $\lambda > 1$, (ii) $\lambda > 1$. **11.** (i) no, (ii) yes, (iii) yes, (iv) no.

12. (i) $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

Exercises 16 (Page 249)

1. (i) yes, $Ker\, T = \{\theta\}$, $Im\, T = \mathbb{R}^2$. (ii) yes, $Ker\, T = \{\theta\}$, $Im\, T = \mathbb{R}^2$.

(iii) no. (iv) yes, $Ker\, T = L\{(1,-2,1)\}$, $Im\, T = L\{(1,3,1)(2,2,1)(3,1,1)\}$.

(v) yes, $Ker\, T = \{\theta\}$, $Im\, T = L\{-1,1,1,1),(1,-1,1,1),(1,1,-1,1)\}$.

(vi) yes, $Ker\, T = L\{(-1,1,0),(-1,0,1)\}$, $Im\, T = \mathbb{R}$

(vii) yes, $Ker\, T$ is the subspace of all 2×2 real skew symmetric matrices. $Im\, T$ is the subspace of all 2×2 real symmetric matrices.

2. $T(x,y,z) = (z,x,y), (x,y,z) \in \mathbb{R}^3$. $Ker\, T = \{\theta\}$, $Im\, T = \mathbb{R}^3$.

3. $T(x,y,z) = (y+z, z+x, x+y), (x,y,z) \in \mathbb{R}^3$. $Ker\, T = \{\theta\}$, $Im\, T = \mathbb{R}^3$.

4. $T(x,y,z) = (\frac{x+y+z}{4}, \frac{x+y+z}{4}, \frac{x+y+z}{4})$, $(x,y,z) \in \mathbb{R}^3$. $Ker\, T = L\{(1,0,-1)$ $(1,-1,0)\}$, $Im\, T = L\{(1,1,1)\}$.

5. $T(x,y,z) = (x,y,z,\frac{x+y+z}{2})$, $(x,y,z) \in \mathbb{R}^3$. $Ker\, T = \{\theta\}$, $Im\, T = L\{(0,1,1,1),(1,0,1,1),(1,1,0,1)\}$.

7. [**Hint.** Let $T(1,0,1) = (0,0,0), T(1,1,0) = (0,0,0), T(1,0,0) = (1,0,0)$].

8. [**Hint.** Let $T(1,0,0) = (1,0,-1), T(0,1,0) = (1,-1,0), T(0,0,1) = (0,0,0)$].

9. [**Hint.** Let $T(1,0,0) = (1,1,1), T(0,1,0) = (1,1,1), T(0,0,1) = (1,1,1)$].

10. [**Hint.** Let $T(1,0,0) = (0,0,0), T(1,0,-1) = (0,1,-1), T(0,1,-1) = (1,0,-1)$].

Exercises 17 (Page 258)

1. (i) $TS(x,y) = (x+y, x+2y)$. (ii) $ST(x,y) = (2x+y, x+y), (x,y) \in \mathbb{R}^2$.

2. (i) $TS(x,y,z) = (x+y+z, x+y, x), (x,y,z) \in \mathbb{R}^3$.

(ii) $ST(x,y,z) = (x, y+z, x+y+z), (x,y,z) \in \mathbb{R}^3$.

4. (iv) **Hint.** Let $V = \mathbb{R}^2$. $T(x,y) = (y,0), (x,y) \in \mathbb{R}^2$.

(v) **Hint.** Let $V = \mathbb{R}^4$. $T(\epsilon_1) = \epsilon_2, T(\epsilon_2) = 0, T(\epsilon_3) = \epsilon_4, T(\epsilon_4) = 0$.

Exercises 18 (Page 288)

1. (i) $\begin{pmatrix} 2 & 1 & 1 \\ -1 & 1 & 3 \end{pmatrix}$, (ii) $\begin{pmatrix} 3 & 3 & -2 \\ 0 & -1 & 3 \end{pmatrix}$. **2.** $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}$.

3. $\begin{pmatrix} 1 & 1 & -1 \\ -1 & 2 & 0 \\ 1 & -2 & 1 \end{pmatrix}$. **4.** (i) $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

5. (i) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$, (ii) $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$.

6. $m(T) = \begin{pmatrix} 2 & 1 & 1 \\ 3 & 2 & 1 \\ 4 & 3 & 1 \end{pmatrix}$.

$T(x,y,z) = (2x+y+z, 3x+2y+z, 4x+3y+z), (x,y,z) \in \mathbb{R}^3$.

7. $m(T) = \frac{1}{2}\begin{pmatrix} 2 & 0 & 2 \\ 5 & -1 & 1 \\ 6 & 0 & 2 \end{pmatrix}$, $m(T^{-1}) = \frac{1}{2}\begin{pmatrix} -1 & 0 & 1 \\ -2 & -4 & 4 \\ 3 & 0 & -1 \end{pmatrix}$.

9. $T(x,y,z) = (2x+3y+3z, 2x+y+3z, 2x+2y), (x,y,z) \in \mathbb{R}^3$.

(i) $\begin{pmatrix} 2 & 3 & 3 \\ 2 & 1 & 3 \\ 2 & 2 & 0 \end{pmatrix}$, (ii) $\begin{pmatrix} 0 & 1 & 1 \\ 2 & 1 & 3 \\ 4 & 4 & 2 \end{pmatrix}$.

Exercises 19 (Page 295)

3. $m(T) = \frac{1}{3}\begin{pmatrix} 2 & -2 & -1 \\ 1 & 2 & -2 \\ 2 & 1 & -2 \end{pmatrix}$.

Exercises 20 (Page 315)

1. (i) $2x^2 + \frac{3}{2}y^2 = 1$, an ellipse;

(ii) $\frac{x^2}{12} - \frac{y^2}{60} = -1$, a hyperbola;

(iii) $(x - \frac{3}{\sqrt{5}})^2 = 0$, a pair of coincident lines;

(iv) $5x^2 - y^2 = 0$, a pair of intersecting lines;

(v) $2(x - \frac{3}{4})^2 + 2(y + \frac{1}{4})^2 = \frac{1}{4}$, a circle;

(vi) $x^2 + \frac{1}{\sqrt{5}}y = 0$, a parabola;

(vii) $4x^2 + 3y^2 = 0$, a point ellipse;

(viii) $5x^2 + 10y^2 + 1 = 0$, a conic without real trace.

2. (i) $x^2 + y^2 + 15z^2 = 1$, an ellipsoid;

(ii) $2x^2 + 2y^2 - z^2 = 1$, a hyperboloid of one sheet;

(iii) $2x^2 - y^2 - z^2 = 1$, a hyperboloid of two sheets;

(iv) $6x^2 - 2y^2 = 1$, a hyperbolic cylinder;

(v) $3x^2 - 3y^2 + \frac{1}{4} = 0$, a hyperbolic cylinder;

(vi) $x^2 - y^2 + z = 0$, a paraboloid;

(vii) $x^2 - 2y^2 = 0$, a pair of planes;

(viii) $x^2 - 2y^2 + \sqrt{2}z = 0$, a hyperbolic paraboloid;

(ix) $3x^2 + \sqrt{6}z = 0$, a parabolic cylinder;

(x) $3x^2 - y^2 = 0$, a pair of planes.