Krishna's

Matrices

First Edition: 1972

Forty Sixth Edition: 2014

Revision, Checking of misprints and Reading done by:

A.K. Vasishtha

I.S.B.N. : 978-81-8283-579-5

Book Code : 236-46

Price : ₹ 240.00 Only

Published by : Satyendra Rastogi *"Mitra"*
for **KRISHNA** Prakashan Media (P) Ltd.
11, Shivaji Road, Meerut – 250 001 (U.P.) India.
Phones: 91.121.2644766, 2642946, 4026111, 4026112; Fax: 91.121.2645855
Website: www.krishnaprakashan.com
E-mail: info@krishnaprakashan.com

Chief Editor : Sugam Rastogi
Printed at : Ganpati Printers, Meerut

Jai Shri Radhey Shyam

Dedicated
to
Lord
Krishna

Authors & Publishers

PREFACE TO THE LATEST EDITION

The authors feel great pleasure in bringing out this enlarged volume of the book on **Matrices**. On repeated demand of several students many more topics have been added to the book so as to make it a complete course for the honours and post-graduate classes of Indian Universities. Besides giving more theorems and problems on eigenvalues and eigenvectors three more chapters have been added to the book. These chapters deal with orthogonal vectors, unitary and orthogonal groups, similarity of matrices, normal matrices and quadratic forms.

Throughout the book, the subject matter has been discussed in such a simple way that the students will not feel any difficulty to understand it. The students are advised to first read the theory portion thoroughly and then they should try to solve the numerical problems themselves taking help from the book whenever necessary.

Suggestions for the improvement of the book will be gratefully received.

—The Authors

SYMBOLS USED IN THE BOOK

$\Rightarrow$	"implies"
iff	"if and only if"
A' or A^T	transpose of a matrix
A^{θ} or A^*	conjugate transpose of a matrix
Adj A	Adjoint of a matrix
$\|A\|$ or det A	Determinant of matrix

BRIEF CONTENTS

DEDICATION .... (III)
PREFACE TO THE LATEST EDITION .... (IV)
BRIEF CONTENTS .... (V)
DETAILED CONTENTS .... (VI-VIII)

Chapter 1: Algebra of Matrices .... (01-65)
Chapter 2: Determinants .... (66-113)
Chapter 3: Inverse of a Matrix .... (114-148)
Chapter 4: Rank of a Matrix .... (149-194)
Chapter 5: Vector Space of *n*-tuples .... (195-208)
Chapter 6: Linear Equations .... (209-240)
Chapter 7: Eigenvalues and Eigenvectors .... (241-267)
Chapter 8: Eigenvalues and Eigenvectors (Continued) .... (268-286)
Chapter 9: Orthogonal Vectors .... (287-299)
Chapter 10: Similarity of Matrices .... (300-336)
Chapter 11: Quadratic forms .... (337-384)

DETAILED CONTENTS

Chapter 1: Algebra of Matrices..........(01-65)

Basic concepts 01
Matrix 02
Square matrix 03
Unit matrix or Identity Matrix 04
Null or zero matrix 04
Submatrices of a matrix 04
Equality of two matrices 05
Addition of matrices 06
Multiplication of a matrix by a scalar 10
Multiplication of two matrices 12
Triangular, Diagonal and Scalar Matrices 20
Trace of a Matrix 45
Transpose of a Matrix 46
Conjugate of a Matrix 48
Transposed conjugate of a Matrix 51
Symmetric and skew-symmetric matrices 52
Hermitian and Skew-Hermitian Matrices 53

Chapter 2: Determinants..........(66-113)

Determinants of order 2 66
Determinants of order 3 66
Minors and cofactors 68
Determinants of order *n* 71
Determinant of a square matrix 71
Properties of Determinants 73
Product of two determinants of the same order 101
System of non-homogeneous linear equations (Cramer's Rule) 109

Chapter 3: Inverse of a Matrix..........(114-148)

Adjoint of a square matrix 114
Inverse or Reciprocal of a Matrix 117
Singular and non-singular matrices 118
Reversal law for the inverse of a product of two matrices 118
Use of the inverse of a matrix to find the solution of a system of linear equations 132
Orthogonal and unitary matrices 137
Partitioning of matrices 142

Chapter 4: Rank of a matrix..........(149-194)

Sub-matrix of a Matrix 149
Minors of a Matrix 149
Rank of a matrix 150
Echelon form of a matrix 151
Elementary transformations of a matrix 156
Elementary Matrices 158
Invariance of rank under elementary transformations 162
Reduction to normal form 164
Equivalence of matrices 166
Row and Column equivalence of matrices 186
Rank of a product of two matrices 188
Computation of the inverse of a non-singular matrix by elementary transformations 191

Chapter 5: Vector Space of *n*-tuples..........(195-208)

Vectors 195
Linear dependence and linear independence of vectors 196
The *n*-vector space 197
Sub-space of an *n*-vector space V_n 197
Basis and dimension of a subspace 198
Row rank of a matrix 199
Left nullity of a matrix 200
Column rank of a matrix 201
Right nullity of a matrix 201
Equality of row rank, column rank and rank 204
Rank of a sum 205

Chapter 6: Linear Equations..........(209-240)

Homogeneous linear equations 209
Fundamental set of solutions 212
System of linear non homogeneous equations 223
Condition for consistency 224

Chapter 7: Eigenvalues and Eigenvectors........................(241-267)

Matrix polynominals 241

Characteristic values and characteristic vectors of a matrix 242

Characteristic roots and characteristic vectors of a matrix 243

Cayley-Hamilton theorem 258

Chapter 8: Eigenvalues and Eigenvectors (Continued)....(268-286)

Characteristic subspaces of a matrix 268

Rank multiplicity Theorem 268

Minimal polynomial and minimal equation of a matrix 279

Chapter 9: Orthogonal Vectors....................................(287-299)

Inner product of two vectors 287

Orthogonal vectors 289

Unitary and orthogonal matrices 292

Orthogonal group 295

Chapter 10: Similarity of Matrices...............................(300-336)

Similarity of matrices 300

Diagonalizable matrix 301

Orthogonally similar matrices 315

Unitarily similar matrices 325

Normal matrices 329

Chapter 11: Quadratic forms..(337-384)

Quadratic Forms 337

Linear transformations 342

Congruence of matrices 343

Reduction of a real quadratic form 350

Canonical or Normal form of a real quadratic form 351

Signature and index of a real quadratic form 353

Sylvester's law of intertia 353

Definite, semi-definite and indefinite real quadratic forms 363

Hermitian forms 383

1

Algebra of Matrices

§ 1. Basic Concepts. Consider the system of equations

$$\begin{aligned} 3x+4y-3z&=5\\ 2x+9y+7z&=4\\ 4x-2y+\ z&=2\\ 6x+8y-3z&=1. \end{aligned}$$

Here x, y and z are unknowns and their coefficients are all numbers. Arranging the coefficients in the order in which they occur in the equations and enclosing them in square brackets, we obtain a rectangular array of the form

$$\begin{bmatrix} 3 & 4 & -3\\ 2 & 9 & 7\\ 4 & -2 & 1\\ 6 & 8 & -3 \end{bmatrix}$$

This rectangular array is an example of a matrix. The horizontal lines ($\rightarrow$) are called rows or *row vectors* and the vertical lines ($\downarrow$) are called columns or *column vectors* of the matrix. There are 4 rows and 3 columns in this matrix. Therefore it is a matrix of the type 4×3. The numbers 3, 4, -3, 2 etc., constituting this matrix are called its elements. The difference between a matrix and a number should be clearly understood. A matrix is not a number. It has got no numerical value. It is a new thing formed with the help of numbers. It is just an ordered collection of numbers arranged in the form of a rectangular array. Simply 5 is a number. But in our notation of matrices [5] is a matrix of the type 1×1 and we cannot have $5=[5]$. We cannot have a relation of equality between a matrix and a number.

We shall use capital letters in bold type to denote matrices.

Thus

$$\mathbf{A}=\begin{bmatrix} 3 & 0 & 1\\ 2 & 1 & 2 \end{bmatrix}_{2\times 3},\quad \mathbf{B}=\begin{bmatrix} 0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0 \end{bmatrix}_{3\times 3},$$

$$C=\begin{bmatrix} 1 & 0 & 2 & 5 \\ 0 & 1 & 3 & 7 \\ -1 & 2 & 4 & 2 \end{bmatrix}_{3\times 4}$$

are all matrices. They are of the type 2×3, 3×3 and 3×4 respectively.

Sometimes we also use the brackets ( ) or the double bars, || ||, in place of the square brackets [ ] to denote matrices. Thus

$$A=\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix},\ B=\begin{pmatrix} 2+3i & 5 \\ -4 & 2-3i \end{pmatrix},\ C=\begin{Vmatrix} 1 & 1 \\ 1 & 1 \end{Vmatrix}$$

are all matrices each of the type 2×2.

We shall now define a matrix.

§ 2. Matrix. Definition.

A set of mn numbers (real or complex) arranged in the form of a rectangular array having m rows and n columns is called an $m\times n$ matrix [to be read as 'm by n' matrix].

[Meerut 1977, Kanpur 87, Rohilkhand 90]

An $m\times n$ matrix is usually written as

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots a_{1n} \\ a_{21} & a_{22} & \ldots a_{2n} \\ a_{31} & a_{32} & \ldots a_{3n} \\ \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots a_{mn} \end{bmatrix}$$

In a compact form the above matrix is represented by $A=[a_{ij}]$, $i=1, 2, \ldots, m$, $j=1, 2, \ldots, n$ or simply by $[a_{ij}]_{m\times n}$. We write the general element of the matrix and enclose it in brackets of the type [ ] or of the type ( )

The numbers a_{11}, a_{12} etc. of this rectangular array are called the **elements** of the matrix. The element a_{ij} belongs to the i^{th} row and the j^{th} column and is sometimes called the $(i, j)^{th}$ element of the matrix.

Thus in the element a_{ij} the first suffix i will always denote the number of the row and the second suffix j, the number of the column in which the element occurs.

In a matrix, the number of rows and columns need not be equal.

A matrix over a field F. Definition. *A set of mn elements of a field F arranged in the form of a rectangular array having m rows and n columns is called an $m\times n$ matrix over the field F.*

If all the elements of a matrix belong to the field of real numbers, the matrix is said to be real. Throughout the present treatment, the elements of a matrix shall be assumed to be complex numbers unless stated otherwise.

Remember. *An $m\times n$ matrix is said to be of the type $m\times n$. It has m rows and n columns.*

A matrix having 4 rows and 2 columns will be of the type 4×2. Similarly a matrix having 3 rows and 1 column will be of the type 3×1 and so on.

Example. *What is the type of the matrix given below :*

$$A=\begin{bmatrix} 3 & 2 & 7 & 8 \\ 5 & -4 & 6 & 11 \\ 4 & 8 & -12 & 10 \end{bmatrix}?$$

Write the elements a_{11}, a_{24}, a_{31}, a_{34}, a_{21} for this matrix.

Solution. The matrix A has 3 rows and 4 columns. Therefore it is a matrix of the type 3×4.

a_{11} = the element belonging to the first row and to the first column = 3.

a_{24} = the element belonging to the second row and to the fourth column = 11.

a_{31} = the element belonging to the third row and to the first column = 4

Similarly $a_{34}=10$, $a_{21}=5$.

§ 3. Special types of Matrices.

(i) Square Matrix. Definition. *An $m\times n$ matrix for which $m=n$ (i.e., the number of rows is equal to the number of columns) is called a square matrix of order n.* It is also called an n-rowed square matrix. Thus in a square matrix, we have the same number of rows and columns. The elements a_{ij} of a square matrix $A=[a_{ij}]_{n\times n}$, for which $i=j$ i.e., the elements $a_{11}, a_{22}, a_{33}, \ldots, a_{nn}$ are called the **diagonal elements** and the line along which they lie is called the **principal diagonal** of the matrix.

Example. *The matrix*

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

is a square matrix of order 4. The elements 0, 3, 1, 2 constitute the principal diagonal of this matrix.

(ii) Unit Matrix or Identity Matrix.

Definition. *A square matrix each of whose diagonal elements is 1 and each of whose non-diagonal elements is equal to zero is called a* **unit matrix** *or an* **identity matrix** *and is denoted by* **I**. $\mathbf{I}_n$ *will denote a unit matrix of order n. Thus a square matrix* $\mathbf{A}=[a_{ij}]$ *is a unit matrix if* $a_{ij}=1$ *when* $i=j$ *and* $a_{ij}=0$ *when* $i\neq j$.

For example,

$$\mathbf{I}_4=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},$$

$$\mathbf{I}_3=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix},\quad \mathbf{I}_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}$$

are unit matrices of orders 4, 3, 2 respectively.

(iii) Null Matrix or Zero Matrix. Definition. *The $m\times n$ matrix whose elements are all 0 is called the* **null matrix** *(or* **zero matrix***) of the type $m\times n$. It is usually denoted by* **O** *or more clearly by* $\mathbf{O}_{m,n}$. *Often a null matrix is simply denoted by the symbol 0 read as 'zero'.*

(Jodhpur 1961, Sagar 65, Gorakhpur 63)

For example,

$$\begin{bmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{bmatrix}_{3\times 5} \text{ and } \begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix}_{3\times 3}$$

are zero matrices of the types 3×5 and 3×3 respectively.

(ix) Row matrices. Column matrices. Definition. *Any $1\times n$ matrix which has only one row and n columns is called a* **row matrix** *or a* **row vector**. *Similarly any $m\times 1$ matrix which has m rows and only one column is a* **column matrix** *or a* **column vector**.

For example, $\mathbf{X}=[2\ \ 7\ \ -8\ \ 5\ \ 11]_{1\times 5}$ is a row matrix of the type 1×5 while

$$\mathbf{Y}=\begin{bmatrix}2\\-9\\11\end{bmatrix}_{3\times 1}$$ is a column matrix of the type 3×1.

§ 4. Submatrices of a matrix. Definition. *Any matrix obtained by omitting some rows and columns from a given $(m\times n)$ matrix* **A** *is called a* **submatrix** *of* **A**.

The matrix **A** itself is a sub-matrix of **A** as it can be obtained from **A** by omitting no rows or columns.

A square submatrix of a square matrix **A** is called a *principal submatrix*, if its diagonal elements are also the diagonal elements of the matrix **A**. Principal submatrices are obtained only by omitting corresponding rows and columns.

Example. The matrix $\begin{bmatrix}1&2&3\\0&2&1\end{bmatrix}$ is a submatrix of the matrix $\mathbf{A}=\begin{bmatrix}1&2&3&9\\7&11&6&5\\0&2&1&8\end{bmatrix}$ as it can be obtained from **A** by omitting the second row and the fourth column.

§ 5. Equality of two matrices. Definition.
Two matrices $\mathbf{A}=[a_{ij}]$ *and* $\mathbf{B}=[b_{ij}]$ *are said to be equal if*
(*i*) *they are of the same size and*
(*ii*) *the elements in the corresponding places of the two matrices are the same i.e.,* $a_{ij}=b_{ij}$ *for each pair of subscripts i and j.*

If two matrices **A** and **B** are equal, we write **A**=**B**. If two matrices **A** and **B** are not equal, we write $\mathbf{A}\neq\mathbf{B}$. If two matrices are not of the same size, they cannot be equal.

Example. *Are the following matrices equal :*

$$\mathbf{A}=\begin{bmatrix}1&3&-1\\2&4&2\\3&0&7\\1&0&9\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}1&3&-1\\2&4&2\\3&0&7\\1&0&9\end{bmatrix}?$$

Solution. The matrix **A** is of the type 4×3 and the matrix **B** is also of the type 4×3. Also the corresponding elements of **A** and **B** are equal. Hence **A**=**B**.

Example 2. *Find the values of a, b, c and d so that the matrices* **A** *and* **B** *may be equal, where*

$$\mathbf{A}=\begin{bmatrix}a&b\\c&d\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}1&-1\\0&3\end{bmatrix}.$$

Solution. We see that the matrices **A** and **B** are of the same size 2×2. If **A**=**B**, then the corresponding elements of **A** and **B** must be equal.

$\therefore$ if $a=1$, $b=-1$, $c=0$, $d=3$, then we will have **A**=**B**.

Example 3. *Are the following matrices equal*

$$\mathbf{A}=\begin{bmatrix}1&7&3\\2&4&0\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}1&7&3\\2&4&1\end{bmatrix}?$$

Solution. Here both the matrices **A** and **B** are of the same size. But $a_{23}=0$ and $b_{23}=-1$. Thus $a_{23}\neq b_{23}$. Therefore $\mathbf{A}\neq\mathbf{B}$.

§ 6. Addition of Matrices.

Definition. *Let* **A** *and* **B** *be two matrices of the same type* $m\times n$. *Then their sum (to be denoted by* **A+B***) is defined to be the matrix of the type* $m\times n$ *obtained by adding the corresponding elements of* **A** *and* **B**. *Thus if*

$\mathbf{A}=[a_{ij}]_{m\times n}$ and $\mathbf{B}=[b_{ij}]_{m\times n}$, then $\mathbf{A}+\mathbf{B}=[a_{ij}+b_{ij}]_{m\times n}$.

Note that **A+B** is also a matrix of the type $m\times n$.

More clearly we can say that if

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}_{m\times n} \text{ and } \mathbf{B}=\begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}_{m\times n}$$

then

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{bmatrix}_{m\times n}.$$

For example, if

$$\mathbf{A}=\begin{bmatrix} 3 & 2 & -1 \\ 4 & -3 & 1 \end{bmatrix}_{2\times 3} \text{ and } \mathbf{B}=\begin{bmatrix} 1 & -2 & 7 \\ 3 & 2 & -1 \end{bmatrix}_{2\times 3}, \text{ then}$$

$$\mathbf{A}+\mathbf{B}=\begin{bmatrix} 3+1 & 2-2 & -1+7 \\ 4+3 & -3+2 & 1-1 \end{bmatrix}=\begin{bmatrix} 4 & 0 & 6 \\ 7 & -1 & 0 \end{bmatrix}_{2\times 3}.$$

Important Note. It should be noted that addition is defined only for matrices which are of the same size. If two matrices **A** and **B** are of the same size, they are said to be **conformable for addition**. If the matrices **A** and **B** are not of the same size, we cannot find their sum.

§ 7. Properties of Matrix Addition.

(i) Matrix addition is Commutative. *If* **A** *and* **B** *be two* $m\times n$ *matrices, then* **A+B=B+A**. **(Meerut 1986, Agra 89)**

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$ and $\mathbf{B}=[b_{ij}]_{m\times n}$. Then

$\mathbf{A}+\mathbf{B}=[a_{ij}]_{m\times n}+[b_{ij}]_{m\times n}$

$=[a_{ij}+b_{ij}]_{m\times n}$ [by definition of addition of two matrices]

$=[b_{ij}+a_{ij}]_{m\times n}$ [Since a_{ij} and b_{ij} are numbers and addition of numbers is commutative]

$=[b_{ij}]_{m\times n}+[a_{ij}]_{m\times n}$ [By definition of addition of two matrices]

$=\mathbf{B}+\mathbf{A}$.

(ii) Matrix addition is associative. *If* **A, B, C** *be three matrices each of the type* $m\times n$, *then* **(A+B)+C=A+(B+C)**. **(Meerut 1986, Agra 87)**

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$, $\mathbf{B}=[b_{ij}]_{m\times n}$, $\mathbf{C}=[c_{ij}]_{m\times n}$.

Then $(\mathbf{A}+\mathbf{B})+\mathbf{C}=([a_{ij}]_{m\times n}+[b_{ij}]_{m\times n})+[c_{ij}]_{m\times n}$

$=[a_{ij}+b_{ij}]_{m\times n}+[c_{ij}]_{m\times n}$ [By definition of **A+B**]

$=[(a_{ij}+b_{ij})+c_{ij}]_{m\times n}$ [By definition of addition of matrices]

$=[a_{ij}+(b_{ij}+c_{ij})]_{m\times n}$ [Since a_{ij}, b_{ij}, c_{ij} are numbers and addition of numbers is associative]

$=[a_{ij}]_{m\times n}+[b_{ij}+c_{ij}]_{m\times n}$ [By definition of addition of two matrices]

$=[a_{ij}]_{m\times n}+([b_{ij}]_{m\times n}+[c_{ij}]_{m\times n})=\mathbf{A}+(\mathbf{B}+\mathbf{C})$.

(iii) Existence of additive identity. *If* **O** *be the* $m\times n$ *matrix each of whose elements is zero, then*

$\mathbf{A}+\mathbf{O}=\mathbf{A}=\mathbf{O}+\mathbf{A}$ *for every* $m\times n$ *matrix* **A**.

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$. Then

$\mathbf{A}+\mathbf{O}=[a_{ij}+0]_{m\times n}=[a_{ij}]_{m\times n}=\mathbf{A}$.

Also $\mathbf{O}+\mathbf{A}=[0+a_{ij}]_{m\times n}=[a_{ij}]_{m\times n}=\mathbf{A}$.

Thus the null matrix **O** of the type $m\times n$ acts as the identity element for addition in the set of all $m\times n$ matrices.

(iv) Existence of the additive inverse.

Negative of a matrix. Definition *Let* $\mathbf{A}=[a_{ij}]_{m\times n}$. *Then the negative of the matrix* **A** *is defined as the matrix* $[-a_{ij}]_{m\times n}$ *and is denoted by* $-\mathbf{A}$.

The matrix $-\mathbf{A}$ is the additive inverse of the matrix **A**. Obviously, $-\mathbf{A}+\mathbf{A}=\mathbf{O}=\mathbf{A}+(-\mathbf{A})$. Here **O** is the null matrix of the type $m\times n$. It is identity element for matrix addition.

Subtraction of two matrices. Definition.

If **A** *and* **B** *are two* $m\times n$ *matrices, then we define*

$$\mathbf{A}-\mathbf{B}=\mathbf{A}+(-\mathbf{B}).$$

Thus the difference $\mathbf{A}-\mathbf{B}$ is obtained by subtracting from each element of **A** the corresponding element of **B**.

(v) Cancellation laws hold good in the case of addition of matrices *i.e., if* **A, B, C** *are three* $m\times n$ *matrices, then*

$\mathbf{A}+\mathbf{B}=\mathbf{A}+\mathbf{C} \Rightarrow \mathbf{B}=\mathbf{C}$ **(left cancellation law)**

and $\mathbf{B}+\mathbf{A}=\mathbf{C}+\mathbf{A} \Rightarrow \mathbf{B}=\mathbf{C}$. **(right cancellation law)**

Proof. We have $\mathbf{A}+\mathbf{B}=\mathbf{A}+\mathbf{C}$

$\Rightarrow -\mathbf{A}+(\mathbf{A}+\mathbf{B})=-\mathbf{A}+(\mathbf{A}+\mathbf{C})$, adding $-\mathbf{A}$ to both sides

$\Rightarrow (-A+A)+B=(-A+A)+C$ [$\because$ matrix addition is associative]

$\Rightarrow O+B=O+C$ [$\because -A+A=O$]

$\Rightarrow B=C.$ [$\because O+B=B$]

Similarly we can prove the right cancellation law.

(vi) The equation $A+X=O$ has a unique solution in the set of all $m \times n$ matrices.

Proof. Let A be an $m \times n$ matrix and let $X=-A$. Then X is also an $m \times n$ matrix. We have

$$A+X=A+(-A)=O.$$

$\therefore$ $X=-A$ is an $m \times n$ matrix such that $A+X=O$.

Now to show that the solution is unique. Let X_1 and X_2 be two solutions of the equation $A+X=O$. Then $A+X_1=O$, and $A+X_2=O$. Therefore, we have

$$A+X_1=A+X_2$$

$\Rightarrow$ $X_1=X_2$, by left cancellation law.

Hence the solution is unique.

Example 1. *If* $A=\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$, $B=\begin{bmatrix} 3 & 7 \\ 4 & 8 \end{bmatrix}$, $C=\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$, *verify that* $A+(B+C)=(A+B)+C$.

Solution. We have $A+B=\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}+\begin{bmatrix} 3 & 7 \\ 4 & 8 \end{bmatrix}$

$$=\begin{bmatrix} 1+3 & 0+7 \\ 2+4 & -1+8 \end{bmatrix}=\begin{bmatrix} 4 & 7 \\ 6 & 7 \end{bmatrix}_{2\times 2}.$$

$$\therefore \quad (A+B)+C=\begin{bmatrix} 4 & 7 \\ 6 & 7 \end{bmatrix}+\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}=\begin{bmatrix} 4-1 & 7+1 \\ 6+0 & 7+0 \end{bmatrix}$$

$$=\begin{bmatrix} 3 & 8 \\ 6 & 7 \end{bmatrix}_{2\times 2}.$$

Also $$B+C=\begin{bmatrix} 3 & 7 \\ 4 & 8 \end{bmatrix}+\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}=\begin{bmatrix} 3-1 & 7+1 \\ 4+0 & 8+0 \end{bmatrix}$$

$$=\begin{bmatrix} 2 & 8 \\ 4 & 8 \end{bmatrix}_{2\times 2}.$$

$$\therefore \quad A+(B+C)=\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}+\begin{bmatrix} 2 & 8 \\ 4 & 8 \end{bmatrix}=\begin{bmatrix} 1+2 & 0+8 \\ 2+4 & -1+8 \end{bmatrix}$$

$$=\begin{bmatrix} 3 & 8 \\ 6 & 7 \end{bmatrix}=(A+B)+C.$$

Example 2. *Find the additive inverse of the matrix*

$$A=\begin{bmatrix} 2 & 3 & -1 & 1 \\ 3 & -1 & 2 & 2 \\ 1 & 2 & 8 & 7 \end{bmatrix}.$$

Solution. The additive inverse of the 3×4 matrix A is the 3×4 matrix each of whose elements is the negative of the corresponding element of A. Therefore if we denote the additive inverse of A by $-A$, we have

$$-A=\begin{bmatrix} -2 & -3 & 1 & -1 \\ -3 & 1 & -2 & -2 \\ -1 & -2 & -8 & -7 \end{bmatrix}.$$

Obviously $A+(-A)=(-A)+A=O$, where O is the null matrix of the type 3×4.

Example 3. *If* $A=\begin{bmatrix} 2 & 7 \\ 9 & 8 \end{bmatrix}$, $B=\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$, *find* $A-B$.

Solution. According to our definition of $A-B$, we have

$$A-B=A+(-B)=\begin{bmatrix} 2 & 7 \\ 9 & 8 \end{bmatrix}+\begin{bmatrix} -1 & -2 \\ 0 & -3 \end{bmatrix}$$

$$=\begin{bmatrix} 2-1 & 7-2 \\ 9+0 & 8-3 \end{bmatrix}=\begin{bmatrix} 1 & 5 \\ 9 & 5 \end{bmatrix}.$$

Example 4. *If* A *and* B *are two* $m \times n$ *matrices and* O *is the null matrix of the type* $m \times n$, *show that*

$$A+B=O \textit{ implies } A=-B \textit{ and } B=-A.$$

Solution. We have $A+B=O$

$\Rightarrow -A+(A+B)=-A+O$ [adding $-A$ to both sides]

$\Rightarrow (-A+A)+B=-A$ [$\because$ matrix addition is associative and the matrix O is the additive identity]

$\Rightarrow O+B=-A$ [$\because -A+A=O$]

$\Rightarrow B=-A$ [$\because O+B=B$]

Similarly $A+B=O \Rightarrow (A+B)+(-B)=O+(-B)$

$\Rightarrow A+[B+(-B)]=-B \Rightarrow A+O=-B$

$\Rightarrow A=-B.$

Example 5. *If* A *is an* $m \times n$ *matrix, then show that*

$$-(-A)=A.$$

Solution. We have

$A+(-A)=O$ [See (iv) on page 7]

$\Rightarrow A=-(-A)$, since by example 4, $A+B=O \Rightarrow A=-B$.

Example 6. *If* $A=\begin{bmatrix} 1 & 4 \\ 3 & 2 \\ 2 & 5 \end{bmatrix}$, $B=\begin{bmatrix} -1 & -2 \\ 0 & 5 \\ 3 & 1 \end{bmatrix}$, *find the matrix* D *such that* $A+B-D=O$.

Solution. We have $A+B-D=O$

$\Rightarrow (A+B)+(-D)=O \Rightarrow A+B=-(-D)=D$

Therefore $\mathbf{D}=\mathbf{A}+\mathbf{B}=\begin{bmatrix}0 & 2\\ 3 & 7\\ 5 & 6\end{bmatrix}$.

§ 8. Multiplication of a matrix by a Scalar. Definition. *Let* $\mathbf{A}$ *be any* $m\times n$ *matrix and* k *any complex number called scalar. The* $m\times n$ *matrix obtained by multiplying every element of the matrix* $\mathbf{A}$ *by* k *is called the scalar multiple of* $\mathbf{A}$ *by* k *and is denoted by* $k\mathbf{A}$ *or* $\mathbf{A}k$. *Symbolically, if* $\mathbf{A}=[a_{ij}]_{m\times n}$, *then* $k\mathbf{A}=\mathbf{A}k=[ka_{ij}]_{m\times n}$.

For example if $k=2$ and $\mathbf{A}=\begin{bmatrix}3 & 2 & -1\\ 4 & -3 & 1\end{bmatrix}_{2\times 3}$,

then $2\mathbf{A}=\begin{bmatrix}2\times 3 & 2\times 2 & 2\times -1\\ 2\times 4 & 2\times -3 & 2\times 1\end{bmatrix}=\begin{bmatrix}6 & 4 & -2\\ 8 & -6 & 2\end{bmatrix}_{2\times 3}$.

Properties of multiplication of a matrix by a scalar.

Theorem 1. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are two matrices each of the type* $m\times n$, *then* $k(\mathbf{A}+\mathbf{B})=k\mathbf{A}+k\mathbf{B}$ *i.e. the scalar multiplication of matrices distributes over the addition of matrices.* **[Allahabad 1976]**

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$ and $\mathbf{B}=[b_{ij}]_{m\times n}$. Then

$k(\mathbf{A}+\mathbf{B})=k([a_{ij}]_{m\times n}+[b_{ij}]_{m\times n})$

$=k[a_{ij}+b_{ij}]_{m\times n}$ [by def. of addition of two matrices]

$=[k(a_{ij}+b_{ij})]_{m\times n}$ [by def. of scalar multiplication]

$=[ka_{ij}+kb_{ij}]_{m\times n}$ [by the distributive law of numbers]

$=[ka_{ij}]_{m\times n}+[kb_{ij}]_{m\times n}=k[a_{ij}]_{m\times n}+k[b_{ij}]_{m\times n}=k\mathbf{A}+k\mathbf{B}$.

Theorem 2. *If* p *and* q *are two scalars and* $\mathbf{A}$ *is any* $m\times n$ *matrix, then* $(p+q)\mathbf{A}=p\mathbf{A}+q\mathbf{A}$.

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$. Then

$(p+q)\mathbf{A}=(p+q)[a_{ij}]_{m\times n}=[(p+q)a_{ij}]_{m\times n}=[pa_{ij}+qa_{ij}]_{m\times n}$

$=[pa_{ij}]_{m\times n}+[qa_{ij}]_{m\times n}=p[a_{ij}]_{m\times n}+q[a_{ij}]_{m\times n}=p\mathbf{A}+q\mathbf{A}$.

Theorem 3. *If* p *and* q *are two scalars and* $\mathbf{A}$ *is any* $m\times n$ *matrix, then* $p(q\mathbf{A})=(pq)\mathbf{A}$.

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$. Then

$p(q\mathbf{A})=p(q[a_{ij}]_{m\times n})=p[qa_{ij}]_{m\times n}=[p(qa_{ij})]_{m\times n}$

$=[(pq)a_{ij}]_{m\times n}$ [$\because$ multiplication of numbers is associative]

$=(pq)[a_{ij}]_{m\times n}=(pq)\mathbf{A}$.

Theorem 4. *If* $\mathbf{A}$ *be any* $m\times n$ *matrix and* k *be any scalar, then* $(-k)\mathbf{A}=-(k\mathbf{A})=k(-\mathbf{A})$.

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$. Then

$(-k)\mathbf{A}=[(-k)a_{ij}]_{m\times n}=[-(ka_{ij})]_{m\times n}$

$=-[ka_{ij}]_{m\times n}=-(k\mathbf{A})$.

Also $(-k)\mathbf{A}=[(-k)a_{ij}]_{m\times n}=[k(-a_{ij})]_{m\times n}$

$=k[-a_{ij}]_{m\times n}=k(-\mathbf{A})$.

Theorem 5. *If* $\mathbf{A}$ *be any* $m\times n$ *matrix, then*

(*i*) $1\mathbf{A}=\mathbf{A}$, (*ii*) $(-1)\mathbf{A}=-\mathbf{A}$.

Proof. Let $\mathbf{A}=[a_{ij}]_{m\times n}$. Then

$1\mathbf{A}=[1a_{ij}]_{m\times n}$, by def. of scalar multiplication

$=[a_{ij}]_{m\times n}$ [$\because 1a_{ij}=a_{ij}$]

$=\mathbf{A}$.

Also $(-1)\mathbf{A}=[(-1)a_{ij}]_{m\times n}=[-a_{ij}]_{m\times n}=-\mathbf{A}$.

Theorem 6. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are two* $m\times n$ *matrices, then* $-(\mathbf{A}+\mathbf{B})=-\mathbf{A}-\mathbf{B}$.

Proof. We have

$-(\mathbf{A}+\mathbf{B})=(-1)(\mathbf{A}+\mathbf{B})$ [by (ii) of theorem 5]

$=(-1)\mathbf{A}+(-1)\mathbf{B}$ [by theorem 1]

$=-\mathbf{A}+(-\mathbf{B})$ [by theorem 5]

$=-\mathbf{A}-\mathbf{B}$.

Example. *If* $\mathbf{A}=\begin{bmatrix}3 & 9 & 0\\ 1 & 8 & -2\\ 7 & 5 & 4\end{bmatrix}$, $\mathbf{B}=\begin{bmatrix}4 & 0 & 2\\ 7 & 1 & 4\\ 2 & 2 & 6\end{bmatrix}$, *verify that* $3(\mathbf{A}+\mathbf{B})=3\mathbf{A}+3\mathbf{B}$.

Solution. We have $\mathbf{A}+\mathbf{B}=\begin{bmatrix}3+4 & 9+0 & 0+2\\ 1+7 & 8+1 & -2+4\\ 7+2 & 5+2 & 4+6\end{bmatrix}$

$=\begin{bmatrix}7 & 9 & 2\\ 8 & 9 & 2\\ 9 & 7 & 10\end{bmatrix}$.

$\therefore\ 3(\mathbf{A}+\mathbf{B})=\begin{bmatrix}3\times 7 & 3\times 9 & 3\times 2\\ 3\times 8 & 3\times 9 & 3\times 2\\ 3\times 9 & 3\times 7 & 3\times 10\end{bmatrix}=\begin{bmatrix}21 & 27 & 6\\ 24 & 27 & 6\\ 27 & 21 & 30\end{bmatrix}$

Again $3\mathbf{A}=3\begin{bmatrix}3 & 9 & 0\\ 1 & 8 & -2\\ 7 & 5 & 4\end{bmatrix}=\begin{bmatrix}3\times 3 & 3\times 9 & 3\times 0\\ 3\times 1 & 3\times 8 & 3\times -2\\ 3\times 7 & 3\times 5 & 3\times 4\end{bmatrix}$

$=\begin{bmatrix}9 & 27 & 0\\ 3 & 24 & -6\\ 21 & 15 & 12\end{bmatrix}$

Also $3\mathbf{B}=3\begin{bmatrix}4 & 0 & 2\\ 7 & 1 & 4\\ 2 & 2 & 6\end{bmatrix}=\begin{bmatrix}3\times 4 & 3\times 0 & 3\times 2\\ 3\times 7 & 3\times 1 & 3\times 4\\ 3\times 2 & 3\times 2 & 3\times 6\end{bmatrix}$

$=\begin{bmatrix}12 & 0 & 6\\ 21 & 3 & 12\\ 6 & 6 & 18\end{bmatrix}$.

$$\therefore \quad 3A+3B=\begin{bmatrix}9 & 27 & 0\\ 3 & 24 & -6\\ 21 & 15 & 12\end{bmatrix}+\begin{bmatrix}12 & 0 & 6\\ 21 & 3 & 12\\ 6 & 6 & 18\end{bmatrix}$$

$$\begin{bmatrix}9+12 & 27+0 & 0+6\\ 3+21 & 24+3 & -6+12\\ 21+6 & 15+6 & 12+18\end{bmatrix}=\begin{bmatrix}21 & 27 & 6\\ 24 & 27 & 6\\ 27 & 21 & 30\end{bmatrix}_{3\times 3}.$$

$\therefore$ 3 (A+B)=3A+3B, *i.e.* the scalar multiplication of matrices distributes over the addition of matrices.

Exercises

1. Can the following two matrices be added :

$$\begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 9 & 8\end{bmatrix}, \begin{bmatrix}6 & 4\\ 4 & 7\\ 3 & 3\end{bmatrix}?$$

2. If $A=\begin{bmatrix}1 & 0 & 5\\ 3 & 2 & 7\\ 5 & 4 & 8\end{bmatrix}$, $B=\begin{bmatrix}3 & 1 & 2\\ 9 & 0 & -6\\ 7 & 4 & 1\end{bmatrix}$, $C=\begin{bmatrix}2 & 0 & -1\\ 7 & 5 & 6\\ 1 & 1 & 4\end{bmatrix}$,

verify that A+(B+C)=(A+B)+C.

3. If $A=\begin{bmatrix}2 & 3 & 1\\ 0 & -1 & 5\end{bmatrix}$, $B=\begin{bmatrix}1 & 2 & -1\\ 0 & -1 & 3\end{bmatrix}$, find 2A—3B.

(Ravi Shankar 1971)

4. If $A=\begin{bmatrix}0 & 1 & 2\\ 2 & 3 & 4\\ 4 & 5 & 6\end{bmatrix}$, $B=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$, find 3A—4B.

(Jiwaji 1970)

Answers

1. No. 3. $\begin{bmatrix}1 & 0 & 5\\ 0 & 1 & 1\end{bmatrix}$. 4. $\begin{bmatrix}-4 & 3 & 6\\ 6 & 5 & 12\\ 12 & 15 & 14\end{bmatrix}$.

§ 9. Multiplication of two matrices. Definition.

(Bombay 1966; Gorakhpur 67; Punjab 71; Jabalpur 68)

Let $A=[a_{ij}]_{m\times n}$ *and* $B=[b_{jk}]_{n\times p}$ *be two matrices such that the number of columns in* A *is equal to the number of rows in* B. *Then the* $m\times p$ *matrix* $C=[c_{ik}]_{m\times p}$ *such that*

$$c_{ik}=\sum_{j=1}^{n} a_{ij}\, b_{jk}$$

[Note that the summation is with respect to the repeated suffix]

is called the product of the matrices A *and* B *in that order and we write* $C=AB$.

In the product AB, the matrix A is called the **pre-factor** and the matrix B is called the **post-factor**. Also we say that the matrix A has been post-multiplied by the matrix B and the matrix B has been pre-multiplied by the matrix A.

Explanation to understand the above definition. The product AB of two matrices A and B exists if and only if the number of columns in A is equal to the number of rows in B. Two such matrices are said to be **comformable for multiplication.** If A is an $m\times n$ matrix and B is an $n\times p$ matrix, then AB is an $m\times p$ matrix. Further if $A=[a_{ij}]_{m\times n}$ and $B=[b_{jk}]_{n\times p}$, then $AB=[c_{ik}]_{m\times p}$ where

$$c_{ik}=\sum_{j=1}^{n} a_{ij}\, b_{jk}=a_{i1}\, b_{1k}+a_{i2}\, b_{2k}+\ldots+a_{in}\, b_{nk}$$

i.e., the $(i, k)^{th}$ element c_{ik} of the matrix AB is obtained by multiplying the corresponding elements of the i^{th} row of A and the k^{th} column of B and then adding the products. The rule of multiplication is **row by column multiplication** *i.e.*, in the process of multiplication we take the rows of A and the columns of B. The element c_{11} of the matrix AB is obtained by adding the products of the corresponding elements of the first row of A and the first column of B. The element c_{12} of the matrix AB is obtained by adding the products of the corresponding elements of the first row of A and the second column of B. Similarly the element c_{21} of the matrix AB is obtained by adding the products of the corresponding elements of the second row of A and the first column of A. In this way we multiply two matrices A and B.

For example, if $A=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\\ a_{31} & a_{32}\end{bmatrix}_{3\times 2}$, $B=\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}_{2\times 2}$

then $AB=\begin{bmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\\ a_{31}b_{11}+a_{32}b_{21} & a_{31}b_{12}+a_{32}b_{22}\end{bmatrix}_{3\times 2}$.

Important Note. If the product AB exists, then it is not necessary that the product BA will also exist. For example if A is a 4×5 matrix and B is a 5×3 matrix, then the product AB exists while the product BA does not exist.

Example 1.

If $A=\begin{bmatrix}2 & 1 & 0\\ 3 & 2 & 1\\ 1 & 0 & 1\end{bmatrix}$ *and* $B=\begin{bmatrix}1 & 2 & 3 & 4\\ 2 & 0 & 1 & 2\\ 3 & 1 & 0 & 5\end{bmatrix}$

then find AB. *Does* BA *exist?*

Solution. The matrix A is of the type 3×3 and the matrix B is of the type 3×4. Since the number of columns of A is equal to the number of rows of B, therefore AB is defined *i.e.*, the product AB exists and it will be a matrix of the type 3×4.

$$\text{Let } AB=\begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14}\\ c_{21} & c_{22} & c_{23} & c_{24}\\ c_{31} & c_{32} & c_{33} & c_{34}\end{bmatrix}.$$

Then c_{11} = the sum of the products of the corresponding elements of the first row of A and the first column of B.

c_{12} = the sum of the products of the corresponding elements of the first row of A and the second column of B.

c_{13} = the sum of the products of the corresponding elements of the first row of A and the third column of B.

c_{23} = the sum of the products of the corresponding elements of the second row of A and the third column of B.

c_{32} = the sum of the products of the corresponding elements of the third row of A and the second column of B, and so on.

Therefore by the row by column rule of multiplication (Rows of A multiplied by the columns of B), we have

$$AB=\begin{bmatrix} 2 & 1 & 0\\ 3 & 2 & 1\\ 1 & 0 & 1\end{bmatrix}_{3\times 3}\times\begin{bmatrix} 1 & 2 & 3 & 4\\ 2 & 0 & 1 & 2\\ 3 & 1 & 0 & 5\end{bmatrix}_{3\times 4}$$

$$=\begin{bmatrix} 2.1+1.2+0.3 & 2.2+1.0+0.1 & 2.3+1.1+0.0 & 2.4+1.2+0.5\\ 3.1+2.2+1.3 & 3.2+2.0+1.1 & 3.3+2.1+1.0 & 3.4+2.2+1.5\\ 1.1+0.2+1.3 & 1.2+0.0+1.1 & 1.3+0.1+1.0 & 1.4+0.2+1.5\end{bmatrix}$$

$$=\begin{bmatrix} 4 & 4 & 7 & 10\\ 10 & 7 & 11 & 21\\ 4 & 3 & 3 & 9\end{bmatrix}_{3\times 4}.$$

Since the number of the columns of B is not equal to the number of rows of A, therefore the product BA does not exist.

Example 2. *If* $A=\begin{bmatrix} 1 & 2\\ 3 & 0\\ 4 & 1\end{bmatrix}$ *and* $B=\begin{bmatrix} 0 & 1 & 0\\ 0 & 2 & 1\\ 2 & 3 & 0\end{bmatrix}$, *find* BA. *Can we find* AB *also* ? **[Meerut 1976]**

Solution. The matrix B has 3 columns and the matrix A has 3 rows, therefore the product BA is defined. By the row by column rule of multiplication, we have

$$BA=\begin{bmatrix} 0 & 1 & 0\\ 0 & 2 & 1\\ 2 & 3 & 0\end{bmatrix}_{3\times 3}\times\begin{bmatrix} 1 & 2\\ 3 & 0\\ 4 & 1\end{bmatrix}_{3\times 2}$$

$$=\begin{bmatrix} 0.1+1.3+0.4 & 0.2+1.0+0.1\\ 0.1+2.3+1.4 & 0.2+2.0+1.1\\ 2.1+3.3+0.4 & 2.2+3.0+0.1\end{bmatrix}=\begin{bmatrix} 3 & 0\\ 10 & 1\\ 11 & 4\end{bmatrix}_{3\times 2}.$$

It should be noted that here we cannot find AB, since the number of columns of A is 2 and the number of rows of B is 3 *i.e.*, they are not equal.

§ 10. Properties of Matrix Multiplication.

***(i) Matrix multiplication is associative if conformability is assured; i.e.,* A(BC)=(AB) C *if* A, B, C *are* $m\times n$, $n\times p$, $p\times q$ *matrices respectively.*

(Meerut 1986; Delhi 81; I.C.S. 86; Gorakhpur 78; Allahabad 67; Rohilkhand 90; Agra 80)

Proof. Let $A=[a_{ij}]_{m\times n}$, $B=[b_{jk}]_{n\times p}$ and $C=[c_{kl}]_{p\times q}$.

Then $AB=[u_{ik}]_{m\times p}$ is an $m\times p$ matrix where

$$u_{ik}=\sum_{j=1}^{n} a_{ij}\, b_{jk} \qquad \ldots(1)$$

Also $BC=[v_{jl}]_{n\times q}$ is an $n\times q$ matrix where

$$v_{jl}=\sum_{k=1}^{p} b_{jk}\, c_{kl}. \qquad \ldots(2)$$

Now A (BC) is an $m\times q$ matrix and (AB) C is also an $m\times q$ matrix.

Let $A(BC)=[w_{il}]_{m\times q}$ where w_{il} is the $(i, l)^{th}$ element of A (BC).

Then $w_{il}=\sum_{j=1}^{n} a_{ij}\, v_{jl}$

$$=\sum_{j=1}^{n}\left[a_{ij}\left\{\sum_{k=1}^{p} b_{jk}\, c_{kl}\right\}\right] \qquad \text{[Putting the value of } v_{jl} \text{ from (2)]}$$

$$=\sum_{k=1}^{p}\left[\left\{\sum_{j=1}^{n} a_{ij}\, b_{jk}\right\} c_{kl}\right] \qquad [\because \text{ finite summations can be interchanged}]$$

$$=\sum_{k=1}^{p} u_{ik}\, c_{kl} \qquad \text{[from (1)]}$$

= the $(i, l)^{th}$ element of (AB) C.

Therefore by the equality of two matrices, we have

A (BC) = (AB) C.

Note. In view of the associative law being true, it is quite legitimate to write ABC for either of the equal products A (BC) or (AB) C.

**(*ii*) *Multiplication of matrices is distributive with respect to addition of matrices i.e.,*

$$A(B+C)=AB+AC,$$

where A, B, C *are any three*, $m\times n$, $n\times p$, $n\times p$ *matrices respectively.*

(Meerut 1986; Agra 70; Poona 70; Gorakhpur 84; Rohilkhand 81)

Proof. Let $A=[a_{ij}]_{m\times n}$, $B=[b_{jk}]_{n\times p}$ and $C=[c_{jk}]_{n\times p}$.

Then both A (B+C) and AB+AC are $m\times p$ matrices.

We have $B+C=[b_{jk}+c_{jk}]_{n\times p}$.

$\therefore$ the $(i, k)^{th}$ element of $A(B+C)=\sum_{j=1}^{n} a_{ij}(b_{jk}+c_{jk})$

$=\sum_{j=1}^{n}\{a_{ij}b_{jk}+a_{ij}c_{jk}\}$ [since the multiplication of numbers is distributive with respect to addition of numbers]

$=\sum_{j=1}^{n} a_{ij}b_{jk}+\sum_{j=1}^{n} a_{ij}c_{jk}$

$=$ the $(i, k)^{th}$ element of AB + the $(i, k)^{th}$ element of AC

$=$ the $(i, k)^{th}$ element of AB+AC.

Hence A (B+C)=AB+AC.

Note 1. It can be shown in a similar manner as above that (B+C) D=BD+CD, where B, C, D are matrices of suitable types so that the above equation is meaningful *i.e.*, if B and C are $m\times n$ matrices then D should be an $n\times p$ matrix.

Note 2. Distributive laws hold unconditionally for square matrices of order n, since conformability is always assured for them.

**(*iii*) *The multiplication of matrices is not always commutative.*

(Gorakhpur 1961)

(a) Whenever AB exists, it is not always necessary that BA should also exist. For example if A be a 5×4 matrix while B be 4×3 matrix then AB exists while BA does not exist.

(b) Whenever AB and BA both exist, it is always not necessary that they should be matrices of the same type. For example if A be a 5×4 matrix while B be a 4×5 matrix, then AB exists and it is a 5×5 matrix. In this case BA also exists and it is a 4×4 matrix. Since the matrices AB and BA are not of the same size therefore we have $AB\neq BA$. **(Gorakhpur 1960)**

(c) Whenever AB and BA both exist and are matrices of the same type, it is not necessary that AB=BA. For example, if

$$A=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \text{ and } B=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \text{ then}$$

$$AB=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$$

$$=\begin{bmatrix}1.0+0.1 & 1.1+0.0\\ 0.0-1.1 & 0.1-1.0\end{bmatrix}=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}$$

$$\text{and } BA=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

$$=\begin{bmatrix}0.1+1.0 & 0.0-1.1\\ 1.1+0.0 & 1.0-0.1\end{bmatrix}=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}.$$

Thus $AB\neq BA$. **(Rohilkhand 1990)**

(d) It however does not imply that AB is never equal to BA. For example if

$$A=\begin{bmatrix}1 & 2 & 1\\ 3 & 4 & 2\\ 1 & 3 & 2\end{bmatrix} \text{ and } B=\begin{bmatrix}10 & -4 & -1\\ -11 & 5 & 0\\ 9 & -5 & 1\end{bmatrix},$$

$$\text{then } AB=\begin{bmatrix}-3 & 1 & 0\\ 4 & -2 & -1\\ -5 & 1 & 1\end{bmatrix}=BA.$$

Definition. ***Whenever* AB=BA, *the matrices* A *and* B *are said to commute*.**

If* AB=−BA, *the matrices* A *and* B *are said to anti-commute.

(*iv*) *If* A *be any* $m\times n$ *matrix and* $O_{n,p}$ *be an* $n\times p$ *null matrix, then* $AO_{n,p}=O_{m,p}$ *where* $O_{m,p}$ *is an* $m\times p$ *null matrix.*

***Similarly if* $O_{m,n}$ *be an* $m\times n$ *null matrix and* A *be any* $n\times p$ *matrix, then* $O_{m,n}A=O_{m,p}$.**

If A be any n-rowed square matrix, and O be an n-rowed null matrix, then AO=OA=O.

(*v*) *The equation* AB=O *does not necessarily imply that at least one of the matrices* A *and* B *must be a zero matrix.*

Or

The product of two matrices can be a zero matrix while neither of them is a zero matrix.

(Delhi 1970; Kanpur 70; Agra 71; Poona 70; Allahabad 79; Rohilkhand 77; Gorakhpur 67)

For example, if $A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$, **and** $B=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$, **then**

$$\mathbf{AB}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}0.1+1.0 & 0.0+1.0\\ 0.1+0.0 & 0.0+0.0\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

Thus AB is a null matrix while neither A nor B is a null matrix.

Thus the product of two matrices can be a zero matrix without either of the matrices being a zero matrix.

***(vi)* *In the case of matrix multiplication if* AB=O, *then it does not necessarily imply that* BA=O. (Lucknow 1978)**

Let $\mathbf{A}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$, **and** $\mathbf{B}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$. **Then AB=O as shown above.**

But $\mathbf{BA}=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}=\begin{bmatrix}1.0+0.0 & 1.1+0.0\\ 0.0+0.0 & 0.1+0.0\end{bmatrix}$

$$=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\neq\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

Thus AB=O while BA≠O.

(vii)* *If* A *be an* $m\times n$ *matrix,* $\mathbf{I}_n$ *denotes the n-rowed unit matrix, it can be easily seen that

$$\mathbf{AI}_n=\mathbf{A}=\mathbf{I}_m\mathbf{A}.$$

§ 11. A useful way of representing matrix products.

Let $\mathbf{A}=[a_{ij}]$, $\mathbf{B}=[b_{jk}]$ **where** $i=1, 2,\ldots, m, j=1, 2,\ldots, n, k=1, 2,\ldots, p$ **be two** $m\times n$ **and** $n\times p$ **matrices respectively.**

Let $\mathbf{R}_1, \mathbf{R}_2,\ldots, \mathbf{R}_m$

and $\mathbf{C}_1, \mathbf{C}_2,\ldots, \mathbf{C}_p$

denote the ordered sets of the rows and columns of A and B respectively. Each of these R's is a $1\times n$ **matrix while each of these C's is a** $n\times 1$ **matrix. Hence the product** $\mathbf{R}_i\mathbf{C}_k$ **is defined for all values of** i **from 1 to** m **and for all values of** k **from 1 to** p**. Moreover all these products are** 1×1 **matrices.**

Now from our definition of matrix product,

$$\begin{bmatrix}\mathbf{R}_1\\ \mathbf{R}_2\\ \vdots\\ \ldots\\ \mathbf{R}_m\end{bmatrix}[\mathbf{C}_1\mathbf{C}_2\ldots\mathbf{C}_p]=\begin{bmatrix}\mathbf{R}_1\mathbf{C}_1 & \mathbf{R}_1\mathbf{C}_2\ldots\ldots\mathbf{R}_1\mathbf{C}_p\\ \mathbf{R}_2\mathbf{C}_1 & \mathbf{R}_2\mathbf{C}_2\ldots\ldots\mathbf{R}_2\mathbf{C}_p\\ \ldots\ldots & \ldots\ldots\ldots\ldots\\ \ldots\ldots & \ldots\ldots\ldots\ldots\\ \mathbf{R}_m\mathbf{C}_1 & \mathbf{R}_m\mathbf{C}_2\ldots\ldots\mathbf{R}_m\mathbf{C}_p\end{bmatrix}$$

is an $m\times p$ **matrix. Also AB is an** $m\times p$ **matrix.**

The $(i, k)^{th}$ **element of AB**

$$=\sum_{j=1}^{n} a_{ij}\, b_{jk}=a_{i1}\, b_{1k}+a_{i2}\, b_{2k}+\ldots+a_{in}\, b_{nk}$$

$$=[a_{i1}\, a_{i2}\ldots a_{in}]\begin{bmatrix}b_{1k}\\ b_{2k}\\ \ldots\\ \ldots\\ b_{nk}\end{bmatrix}$$

$=\mathbf{R}_i\mathbf{C}_k$

$=(i, k)^{th}$ **element of the matrix** $[\mathbf{R}_i\mathbf{C}_k]$.

Hence $\mathbf{AB}=\begin{bmatrix}\mathbf{R}_1\\ \mathbf{R}_2\\ \ldots\\ \ldots\\ \mathbf{R}_m\end{bmatrix}[\mathbf{C}_1\mathbf{C}_2\ldots\mathbf{C}_p]=\begin{bmatrix}\mathbf{R}_1\mathbf{C}_1 & \mathbf{R}_1\mathbf{C}_2 & \ldots & \mathbf{R}_1\mathbf{C}_p\\ \mathbf{R}_2\mathbf{C}_1 & \mathbf{R}_2\mathbf{C}_2 & \ldots & \mathbf{R}_2\mathbf{C}_p\\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots\\ \ldots\ldots & \ldots\ldots & \ldots & \ldots\ldots\\ \mathbf{R}_m\mathbf{C}_1 & \mathbf{R}_m\mathbf{C}_2 & \ldots & \mathbf{R}_m\mathbf{C}_p\end{bmatrix}.$

§ 12. Associative law for the product of four matrices.

If $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{A}_4$ ***be four matrices of suitable sizes for multiplication to be possible, then their product is independent of any particular manner of bracketing, provided that the order of the matrix factors is not changed.***

We have $\mathbf{A}_1\{\mathbf{A}_2(\mathbf{A}_3\mathbf{A}_4)\}$

$=\{(\mathbf{A}_1\mathbf{A}_2)(\mathbf{A}_3\mathbf{A}_4)\}$ **[By the associative law for the product of three matrices** $\mathbf{A}_1$, $\mathbf{A}_2$ **and** $\mathbf{A}_3\mathbf{A}_4$**]**

$=\{(\mathbf{A}_1\mathbf{A}_2)\mathbf{A}_3\}\mathbf{A}_4$ **[By the associative law for the product of three matrices** $\mathbf{A}_1\mathbf{A}_2$, $\mathbf{A}_3$ **and** $\mathbf{A}_4$**]**

$=\{\mathbf{A}_1(\mathbf{A}_2\mathbf{A}_3)\}\mathbf{A}_4$ **[By the associative law for the product of three matrices** $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$**].**

In view of the associative law being true it is quite legitimate to write $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\mathbf{A}_4$ **for each of the above four products.**

Note. By mathematical induction we can generalize the associative law for the product of any number of matrices.

§ 13. Positive Integral powers of matrices.

The product AA is defined only when A is a square matrix. We shall denote this product by $\mathbf{A}^2$**. We shall write**

$$\mathbf{A}^1=\mathbf{A}.$$

Now $\mathbf{A}^2\mathbf{A}=(\mathbf{AA})\,\mathbf{A}=\mathbf{A}\,(\mathbf{AA})$ **[By associative law]**

$$=\mathbf{AA}^2=\mathbf{AAA}.$$

We shall denote each of the above products by A^3 so that

$$A^3 = AAA = AA^2 = A^2A.$$

The product of any number of matrices is associative *i.e.*, it is independent of the manner in which the matrices may be bracketed for multiplication. Therefore it is quite legitimate to denote the product AAA...A, m times by A^m, where m is any +ive integer.

If m and n are any arbitrary positive integers, we have

$A^m.A^n$ = (AA...A, m times).(AAA...A, n times)

= AAA...A, $m+n$ times [since the product of any number of matrices is associative]

$= A^{m+n}$. **(Agra 1979)**

Again $(A^m)^n = A^m.A^m.A^m...A^m$, n times

= (AAA...A, m times) (AAA...A, m times)...upto n times

= AAA...A, mn times $= (A)^{mn}$.

§ 14. Triangular, Diagonal and Scalar Matrices.

(i) Upper Triangular Matrix. Definition. *A square matrix* $A=[a_{ij}]$ *is called an upper triangular matrix if* $a_{ij}=0$ *whenever* $i > j$.

Thus in an upper triangular matrix all the elements below the principal diagonal are zero.

For example $\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$

is an upper triangular matrix of the type $n \times n$. Similarly

$$A=\begin{bmatrix} 1 & 2 & 4 & 2 \\ 0 & 3 & -1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 8 \end{bmatrix}_{4\times 4}, \quad B=\begin{bmatrix} 2 & -9 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}_{3\times 3}$$

are upper triangular matrices.

(ii) Lower Triangular Matrix. Definition. *A square matrix* $A=[a_{ij}]$ *is called a lower triangular matrix if* $a_{ij}=0$ *whenever* $i < j$.

Thus in a lower triangular matrix all the elements above the principal diagonal are zero.

For example $\begin{bmatrix} a_{11} & 0 & 0 & \ldots 0 \\ a_{21} & a_{22} & 0 & \ldots 0 \\ a_{31} & a_{32} & a_{33} & \ldots 0 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots a_{nn} \end{bmatrix}_{n\times n}$

is a lower triangular matrix of the size $n \times n$. Similarly

$$A=\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}_{2\times 2}, \quad B=\begin{bmatrix} 3 & 0 & 0 & 0 \\ 4 & -1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 5 & 7 & 1 & 2 \end{bmatrix}_{4\times 4},$$

are lower triangular matrices.

A triangular matrix $A=[a_{ij}]_{n\times n}$ is called **strictly triangular** if $a_{ii}=0$ for $i=1, 2,..., n$.

(iii) Diagonal Matrix. Definition. *A square matrix* $A=[a_{ij}]_{n\times n}$ *whose elements above and below the principal diagonal are all zero, i.e.* $a_{ij}=0$ *for all* $i \neq j$, *is called a diagonal matrix.* **(Rohilkhand 1990)**

Thus a diagonal matrix is both upper and lower triangular. An n-rowed diagonal matrix whose diagonal elements in order are $d_1, d_2, d_3, ..., d_n$ will often be denoted by the symbol

$$\text{Diag } [d_1, d_2, ..., d_n].$$

For example,

$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ are diagonal matrices.

(iv) Scalar Matrix. Definition. *A diagonal matrix whose diagonal elements are all equal is called a scalar matrix.*

If $S=\begin{bmatrix} k & 0 & \ldots & 0 \\ 0 & k & \ldots & \\ 0 & 0 & \ldots & \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & k \end{bmatrix}$

is an n-rowed scalar matrix each of whose diagonal elements equal to k and A is any n-rowed square matrix, then

$$AS = SA = kA,$$

i.e., **the pre-multiplication or the post-multiplication of A by S has the same effect as the multiplication of A by the scalar k. This is perhaps the motivation behind the name 'scalar matrix'.**

As a particular case, if we take

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ and } \mathbf{S}=\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix}, \text{ then}$$

$$\mathbf{AS}=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & k \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & ka_{13} \\ ka_{21} & ka_{22} & ka_{23} \\ ka_{31} & ka_{32} & ka_{33} \end{bmatrix} = k\mathbf{A}.$$

Similarly $\mathbf{SA}=k\mathbf{A}$. Hence $\mathbf{SA}=\mathbf{AS}=k\mathbf{A}$.

If in place of S, we take $\mathbf{I}_3=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

$$\text{then } \mathbf{AI}_3=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

Similarly $\mathbf{I}_3\mathbf{A}=\mathbf{A}$.

Hence $\mathbf{AI}_3=\mathbf{I}_3\mathbf{A}=\mathbf{A}$.

Solved Examples

Ex. 1. *If* $\mathbf{A}=\begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix}$, $\mathbf{B}=\begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$,

find **AB** *and* **BA** *and show that* **AB≠BA.**

(Meerut 1975, 77)

Solution. We have

$$\mathbf{AB}=\begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix}_{3\times3} \begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}_{3\times3}$$

$$=\begin{bmatrix} 2.1-3.1+4.0 & 2.3+3.2+4.0 & 2.0+3.1+4.2 \\ 1.1-2.1+3.0 & 1.3+2.2+3.0 & 1.0+2.1+3.2 \\ -1.1-1.1+2.0 & -1.3+1.2+2.0 & -1.0+1.1+2.2 \end{bmatrix}_{3\times3}$$

$$=\begin{bmatrix} -1 & 12 & 11 \\ -1 & 7 & 8 \\ -2 & -1 & 5 \end{bmatrix}_{3\times3}.$$

$$\text{Similarly, } \mathbf{BA}=\begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}_{3\times3} \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix}_{3\times3}$$

$$=\begin{bmatrix} 1.2+3.1-0.1 & 1.3+3.2+0.1 & 1.4+3.3+0.2 \\ -1.2+2.1-1.1 & -1.3+2.2+1.1 & -1.4+2.3+1.2 \\ 0.2+0.1-2.1 & 0.3+0.2+2.1 & 0.4+0.3+2.2 \end{bmatrix}$$

$$=\begin{bmatrix} 5 & 9 & 13 \\ -1 & 2 & 4 \\ -2 & 2 & 4 \end{bmatrix}_{3\times3}.$$

The matrix **AB** is of the type 3×3 and the matrix **BA** is also of the type 3×3. But the corresponding elements of these matrices are not equal. Hence **AB≠BA.**

Ex. 2. *If* $\mathbf{A}=\begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}$, $\mathbf{B}=\begin{bmatrix} 1 & -1 \\ 3 & 4 \\ -1 & 5 \end{bmatrix}$

does **AB** *exist ?*

Solution. A is a matrix of the type 2×2 while B is a matrix of the type 3×2. Thus number of rows in B is not equal to the number of columns in A. Hence A and B are not conformable for multiplication and therefore AB does not exist.

Ex. 3. *If* $\mathbf{A}=\begin{bmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \\ -3 & 1 & 2 \end{bmatrix}$ *and* $\mathbf{B}=\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix}$

form the products **AB** *and* **BA**, *and show that* **AB≠BA.**

Solution. Since A and B are both square matrices each of the type 3×3, therefore both the products **AB and BA** will be defined.

$$\mathbf{AB}=\begin{bmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \\ -3 & 1 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix}$$

$$=\begin{bmatrix} 1.1+(-2).0+3.1 & 1.0+(-2).1+3.2 & 1.2+(-2).2+3.0 \\ 2.1+3.0+(-1).1 & 2.0+3.1+(-1).2 & 2.2+3.2+(-1).0 \\ -3.1+1.0+2.1 & -3.0+1.1+2.2 & -3.2+1.2+2.0 \end{bmatrix}$$

$$=\begin{bmatrix} 4 & 4 & -2 \\ 1 & 1 & 10 \\ -1 & 5 & -4 \end{bmatrix}.$$

$$\text{Also } \mathbf{BA}=\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix} \times \begin{bmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \\ -3 & 1 & 2 \end{bmatrix}$$

$$=\begin{bmatrix} 1.1+0.2+2.(-3) & 1.(-2)+0\cdot3+2.1 & 1.3+0.(-1)+2.2 \\ 0.1+1.2+2.(-3) & 0.(-2)+1.3+2.1 & 0\cdot3+1.(-1)+2.2 \\ 1.1+2.2+0.(-3) & 1.(-2)+2.3+0.1 & 1.3+2.(-1)+0.2 \end{bmatrix}$$

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

The matrix AB is of the type 3×3 and the matrix BA is also of the type 3×3. But the corresponding elements of these matrices are not equal. Hence $AB \neq BA$. Thus the multiplication of matrices is not in general commutative.

Ex. 4. *Find the product matrix of the matrices*

$$A=\begin{bmatrix}2 & 1 & 2 & 1\\ 1 & 1 & 1 & 1\end{bmatrix} \text{ and } B=\begin{bmatrix}2 & -1 & 0\\ 0 & 4 & 1\\ -2 & 1 & 0\\ 1 & -3 & 2\end{bmatrix}.$$

(Gorakhpur 1960)

Solution. The matrix A is of the type 2×4 and the matrix B is of the type 4×3. Therefore the product AB is defined and it will be a matrix of the type 2×3.

$$AB=\begin{bmatrix}2 & 1 & 2 & 1\\ 1 & 1 & 1 & 1\end{bmatrix}\times\begin{bmatrix}2 & -1 & 0\\ 0 & 4 & 1\\ -2 & 1 & 0\\ 1 & -3 & 2\end{bmatrix}$$

$$=\begin{bmatrix}2.2+1.0+2.(-2)+1.1 & 2.(-1)+1.4+2.1+1.(-3) & 2.0+1.1+2.0+1.2\\ 1.2+1.0+1.(-2)+1.1 & 1.(-1)+1.4+1.1+1.(-3) & 1.0+1.1+1.0+1.2\end{bmatrix}$$

$$=\begin{bmatrix}1 & 1 & 3\\ 1 & 1 & 3\end{bmatrix}.$$

Ex. 5. *If* $A=\begin{bmatrix}1 & -2 & 3\\ -4 & 2 & 5\end{bmatrix}$ *and* $B=\begin{bmatrix}2 & 3\\ 4 & 5\\ 2 & 1\end{bmatrix}$

find AB and show that $AB \neq BA$.

(Kanpur 1981; Gorakhpur 62; Agra 70)

Solution. The matrix A is of the type 2×3 and the matrix B is of the type 3×2. Therefore the product AB is defined and it will be a matrix of the type 2×2.

$$AB=\begin{bmatrix}1 & -2 & 3\\ -4 & 2 & 5\end{bmatrix}_{2\times 3}\times\begin{bmatrix}2 & 3\\ 4 & 5\\ 2 & 1\end{bmatrix}_{3\times 2}$$

$$=\begin{bmatrix}1.2+(-2).4+3.2 & 1.3+(-2).5+3.1\\ (-4).2+2.4+5.2 & (-4).3+2.5+5.1\end{bmatrix}=\begin{bmatrix}0 & -4\\ 10 & 3\end{bmatrix}_{2\times 2}.$$

Again the matrix B is of the type 3×2 and the matrix A is of the type 2×3. Therefore the product BA is also defined and it will be a matrix of the type 3×3.

We have $BA=\begin{bmatrix}2 & 3\\ 4 & 5\\ \dots\end{bmatrix}_{3\times 2}\times\begin{bmatrix}1 & -2 & 3\\ -4 & 2 & 5\end{bmatrix}_{2\times 3}$

$$=\begin{bmatrix}-10 & 2 & 21\\ -16 & 2 & 37\\ -2 & -2 & 11\end{bmatrix}_{3\times 3}.$$

Since the matrices AB and BA are not of the same type therefore $AB \neq BA$.

Ex. 6. *Find the product of the following two matrices:*

$$\begin{bmatrix}0 & c & -b\\ -c & 0 & a\\ b & -a & 0\end{bmatrix} \text{ and } \begin{bmatrix}a^2 & ab & ac\\ ab & b^2 & bc\\ ac & bc & c^2\end{bmatrix}.$$

(Kanpur 1980; Agra 88; Ravi Shankar 70)

Solution. The required product

$$=\begin{bmatrix}0 & c & -b\\ -c & 0 & a\\ b & -a & 0\end{bmatrix}\times\begin{bmatrix}a^2 & ab & ac\\ ab & b^2 & bc\\ ac & bc & c^2\end{bmatrix}$$

$$=\begin{bmatrix}0.a^2+c.ab+(-b).ac & 0.ab+c.b^2-b.(bc) & 0.ac+c.(bc)+(-b)\ c^2\\ (-c).a^2+0.ab+a.(ac) & -c\ (ab)+0.b^2+a.(bc) & -c.(ac)+0\ (bc)+a.c^2\\ b.a^2+(-a)\ ab+0\ (ac) & b\ (ab)+(-a)\ b^2+0.(bc) & b\ (ac)+(-a)\ bc+0.c^2\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

Ex. 7. *If A, B, C are three matrices such that*

$$A=[x \quad y \quad z],\ B=\begin{bmatrix}a & h & g\\ h & b & f\\ g & f & c\end{bmatrix},\ C=\begin{bmatrix}x\\ y\\ z\end{bmatrix}$$

find ABC. **(Rohilkhand 1981; Gorakhpur 79; Agra 72)**

Solution. Since the product of matrices is associative, therefore we can find ABC either by finding (AB)C or by finding A(BC). Let us find (AB) C.

A is a matrix of the type 1×3 and B is of the type 3×3. So AB will be of the type 1×3.

$$AB=[x \quad y \quad z]_{1\times 3}\times\begin{bmatrix}a & h & g\\ h & b & f\\ g & f & c\end{bmatrix}_{3\times 3}$$

$$=[x.a+y.h+z.g \quad x.h+y.b+z.f \quad x.g+y.f+z.c]_{1\times 3}$$

$$=[ax+hy+gz \quad hx+by+fz \quad gx+fy+cz]_{1\times 3}.$$

Now **AB** is of the type 1×3 and **C** is of the type 3×1. Therefore **(AB) C** will be of the type 1×1.

$$(\mathbf{AB})\,\mathbf{C}=[ax+hy+gz \quad hx+by+fz \quad gx+fy+cz]_{1\times 3}\times\begin{bmatrix}x\\y\\z\end{bmatrix}_{3\times 1}$$

$$=[x\,(ax+hy+gz)+y\,(hx+by+fz)+z\,(gx+fy+cz)]_{1\times 1}$$

$$=[ax^2+by^2+cz^2+2hxy+2gzx+2fyz]_{1\times 1}.$$

Ex. 8. *If* $\mathbf{A}=\begin{bmatrix}1 & 1 & -1\\2 & 0 & 3\\3 & -1 & 2\end{bmatrix}$, $\mathbf{B}=\begin{bmatrix}1 & 3\\0 & 2\\-1 & 4\end{bmatrix}$

and $\mathbf{C}=\begin{bmatrix}1 & 2 & 3 & -4\\2 & 0 & -2 & 1\end{bmatrix}$,

find **A (BC)** *and* **(AB) C** *and show that* **A (BC)=(AB) C.**

Solution. We have

$$\mathbf{BC}=\begin{bmatrix}1 & 3\\0 & 2\\-1 & 4\end{bmatrix}\times\begin{bmatrix}1 & 2 & 3 & -4\\2 & 0 & -2 & 1\end{bmatrix}$$

$$=\begin{bmatrix}1.1+3.2 & 1.2+3.0 & 1.3-3.2 & -1.4+3.1\\0.1+2.2 & 0.2+2.0 & 0.3-2.2 & -0.4+2.1\\-1.1+4.2 & -1.2+4.0 & -1.3-4.2 & +1.4+4.1\end{bmatrix}$$

$$=\begin{bmatrix}7 & 2 & -3 & -1\\4 & 0 & -4 & 2\\7 & -2 & -11 & 8\end{bmatrix}_{3\times 4}.$$

$$\therefore\ \mathbf{A}\,(\mathbf{BC})=\begin{bmatrix}1 & 1 & -1\\2 & 0 & 3\\3 & -1 & 2\end{bmatrix}\times\begin{bmatrix}7 & 2 & -3 & -1\\4 & 0 & -4 & 2\\7 & -2 & -11 & 8\end{bmatrix}$$

$$=\begin{bmatrix}1.7+1.4-1.7 & 1.2+1.0+1.2 & -1.3-1.4+1.11 & -1.1+1.2-1.8\\2.7+0.4+3.7 & 2.2+0.0-3.2 & -2.3-0.4-3.11 & -2.1+0.2+3.8\\3.7-1.4+2.7 & 3.2-1.0-2.2 & -3.3+1.4-2.11 & -3.1-1.2+2.8\end{bmatrix}_{3\times 4}$$

$$=\begin{bmatrix}4 & 4 & 4 & -7\\35 & -2 & -39 & 22\\31 & 2 & -27 & 12\end{bmatrix}_{3\times 4}.$$

Similarly find **AB**. Then post-multiplying **AB** by **C** find **(AB)C.** We shall see that **(AB) C=A (BC)**.

Thus we verify that the product of matrices is associative.

Ex. 9. *If* $\mathbf{A}=\begin{bmatrix}5 & 0 & 0\\0 & 5 & 0\\0 & 0 & 5\end{bmatrix}$ *and* $\mathbf{B}=\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix}$

then show that **AB=BA=5B.**

Solution. $\mathbf{AB}=\begin{bmatrix}5 & 0 & 0\\0 & 5 & 0\\0 & 0 & 5\end{bmatrix}\times\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix}$

$$=\begin{bmatrix}5a_{11} & 5a_{12} & 5a_{13}\\5a_{21} & 5a_{22} & 5a_{23}\\5a_{31} & 5a_{32} & 5a_{33}\end{bmatrix}=5\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix}=5\mathbf{B}.$$

Similarly **BA=5B.** Hence **AB=BA=5B.**

Note. From this example we conclude that if **A** is a scalar matrix of the type $n\times n$ each of whose diagonal elements is a and **B** be any other $n\times n$ matrix, then pre-multiplying or post-multiplying **B** by **A** will be the same as the scalar multiplication of **B** by the scalar a.

***Ex. 10.** *If* **A** *is any* $m\times n$ *matrix such that* **AB** *and* **BA** *are both defined show that* **B** *is an* $n\times m$ *matrix.* **(Allahabad 1979)**

Solution. **A** is an $m\times n$ matrix.

If **A** and **B** are conformable for multiplication, the number of rows in **B** should be equal to the number of columns in **A**. So the number of rows in **B** should be equal to n.

Let **B** be an $n\times p$ matrix.

Since the product **BA** is also defined, the number of rows in **A** should be equal to the number of columns in **B**.

Therefore $m=p$.

Hence **B** is an $n\times m$ matrix.

Ex. 11. **A, B** *are two marices such that* **AB** *and* **A+B** *are both defined; show that* **A, B** *are square matrices of the same order.* **(Allahabad 1979)**

Solution Let **A** be an $m\times n$ matrix.

Since **A+B** is defined, therefore **B** should also be an $m\times n$ matrix.

Further since **AB** is defined, therefore $m=n$.

Hence **A** and **B** are square matrices of the same order.

Ex. 12. *Show that for all values of* p, q, r, s *the matrices,*

$\mathbf{P}=\begin{bmatrix}p & q\\-q & p\end{bmatrix}$, *and* $\mathbf{Q}=\begin{bmatrix}r & s\\-s & r\end{bmatrix}$ *commute.*

Solution. We have $\mathbf{PQ}=\begin{bmatrix}pr-qs & ps+qr\\-qr-ps & -qs+pr\end{bmatrix}$.

Also $\mathbf{QP}=\begin{bmatrix}r & s\\-s & r\end{bmatrix}\begin{bmatrix}p & q\\-q & p\end{bmatrix}=\begin{bmatrix}rp-sq & rq+sp\\-sp-rq & -sq+rp\end{bmatrix}$

$=\begin{bmatrix}pr-qs & ps+qr\\-qr-ps & -qs+pr\end{bmatrix}$, for all values of p, q, r, s.

Hence **PQ=QP**, for all values of p, q, r, s.

Ex. 13. *If A and B are n-rowed square matrices, show that*

(i) $(A+B)^2=A^2+AB+BA+B^2$.

(ii) $(A+B)^3=A^3+BA^2+ABA+B^2A+A^2B+BAB+AB^2+B^3$

(iii) $(A+B)(A-B)=A^2-AB+BA-B^2$

(iv) $(A-B)^2=A^2-AB-BA+B^2$

(v) $(A-B)(A+B)=A^2+AB-BA-B^2$.

What do these formulae become when A and B commute ?

Solution. (i) Since $A+B$ is also an n-rowed square matrix, therefore we have

$(A+B)^2=(A+B)(A+B)=(A+B)A+(A+B)B$

[by distributive law since A and B are both n-rowed square matrices]

$=AA+BA+AB+BB=A^2+BA+AB+B^2$.

If $AB=BA$, we have $(A+B)^2=A^2+2AB+B^2$.

(ii) Again $(A+B)^3=(A+B)(A+B)^2$

$=(A+B)(A^2+BA+AB+B^2)$

$=A^3+ABA+AAB+AB^2+BA^2+BBA+BAB+B^3$

[The distributive law is applicable since A and B are both square matrices of order n and A^2, BA, AB, B^2 are also square matrices of order n].

$=A^3+BA^2+ABA+B^2A+A^2B+BAB+AB^2+B^3$.

If $AB=BA$,

$BA^2=BAA=ABA=AAB=A^2B$

$ABA=AAB=A^2B$

$B^2A=BBA=BAB=ABB=AB^2$

$BAB=ABB=AB^2$

$\therefore \quad (A+B)^3=A^3+3A^2B+3AB^2+B^3$.

(iii) $(A+B)(A-B)=(A+B)\{A+(-B)\}$

$=AA+A(-B)+BA+B(-B)=A^2-AB+BA-B^2$.

If $AB=BA$, then

$(A+B)(A-B)=A^2-B^2+O$ [since $BA-BA$ = a null matrix]

$=A^2-B^2$.

(iv) $(A-B)^2=(A-B)(A-B)=AA-AB-BA+BB$

$=A^2-AB-BA+B^2$.

If $AB=BA$, then $(A-B)^2=A^2-2AB+B^2$.

(v) $(A-B)(A+B)=A^2+AB-BA-B^2$.

If $AB=BA$, then

$(A-B)(A+B)=A^2-B^2+O$, [since $AB-AB$ is a null matrix]

$=A^2-B^2$.

Ex. 14. *Under what conditions is the matrix equation*

$$A^2-B^2=(A-B)(A+B)$$

true ? **(Sagar 1975)**

Solution. We have

$A^2-B^2=(A-B)(A+B) \Rightarrow A^2-B^2=A^2+AB-BA-B^2$

$\Rightarrow A^2-B^2-(A^2+AB-BA-B^2)=O$

$\Rightarrow BA-AB=O \quad [\because \quad A^2-A^2=O,\ B^2-B^2=O]$

$\Rightarrow BA=AB$.

Hence the given matrix equation is true only if the matrices A and B commute.

Ex. 15. *If A and B are square matrices of order n, then prove that A and B will commute if and only if* $A-\lambda I$ *and* $B-\lambda I$ *commute for every scalar* λ.

Solution. Suppose the matrices A and B commute *i.e.*,

$$AB=BA.$$

Then to prove that the matrices $A-\lambda I$ and $B-\lambda I$ commute for every scalar λ. We have

$(A-\lambda I)(B-\lambda I)=AB-\lambda AI-\lambda IB+\lambda^2 I^2=AB-\lambda A-\lambda B+\lambda^2 I$

$=AB-\lambda(A+B)+\lambda^2 I$.

Similarly $(B-\lambda I)(A-\lambda I)=BA-\lambda(A+B)+\lambda^2 I$. But it is given that $AB=BA$. Therefore $(A-\lambda I)(B-\lambda I)=(B-\lambda I)(A-\lambda I)$ for every scalar λ.

Conversely suppose that $A-\lambda I$ and $B-\lambda I$ commute for every scalar λ. Then $(A-\lambda I)(B-\lambda I)=(B-\lambda I)(A-\lambda I)$.

$\therefore \quad AB-\lambda(A+B)+\lambda^2 I=BA-\lambda(A+B)+\lambda^2 I$

or $\quad AB=BA$ and hence A and B commute.

Ex. 16. *Given* $A=\begin{bmatrix}1 & 2 & -3\\ 5 & 0 & 2\\ 1 & -1 & 1\end{bmatrix}$ *and* $B=\begin{bmatrix}3 & -1 & 2\\ 4 & 2 & 5\\ 2 & 0 & 3\end{bmatrix}$;

find the matrix C such that $A+C=B$.

Solution. $A+C=B \Rightarrow -A+(A+C)=-A+B$

$\Rightarrow (-A+A)+C=B-A$.

[$\because$ Matrix addition is associative and commutative]

$\Rightarrow O+C=B-A \Rightarrow C=B-A$.

Therefore $C=\begin{bmatrix}2 & -3 & 5\\ -1 & 2 & 3\\ 1 & 1 & 2\end{bmatrix}$.

Ex. 17. *Prove that the product of two matrices*

$$\begin{bmatrix} \cos^2 \theta & \cos \theta \sin \theta \\ \cos \theta \sin \theta & \sin^2 \theta \end{bmatrix} \text{ and } \begin{bmatrix} \cos^2 \phi & \cos \phi \sin \phi \\ \cos \phi \sin \phi & \sin^2 \phi \end{bmatrix}$$

is a zero matrix when θ *and* ϕ *differ by an odd multiple of* $\pi/2$.

(Raj. 1975; Meerut 87)

Solution. The product of the two given matrices is

$$=\begin{bmatrix} \cos^2 \theta \cos^2 \phi + \cos \theta \sin \theta \cos \phi \sin \phi & \cos^2 \theta \cos \phi \sin \phi + \cos \theta \sin \theta \sin^2 \phi \\ \cos \theta \sin \theta \cos^2 \phi + \sin^2 \theta \cos \phi \sin \phi & \cos \theta \sin \theta \cos \phi \sin \phi + \sin^2 \theta \sin^2 \phi \end{bmatrix}$$

$$=\begin{bmatrix} \cos \theta \cos \phi (\cos \theta \cos \phi + \sin \theta \sin \phi) & \cos \theta \sin \phi (\cos \theta \cos \phi + \sin \theta \sin \phi) \\ \sin \theta \cos \phi (\cos \theta \cos \phi + \sin \theta \sin \phi) & \sin \theta \sin \phi (\cos \theta \cos \phi + \sin \theta \sin \phi) \end{bmatrix}$$

$$=\begin{bmatrix} \cos \theta \cos \phi \cos (\theta - \phi) & \cos \theta \sin \phi \cos (\theta - \phi) \\ \sin \theta \cos \phi \cos (\theta - \phi) & \sin \theta \sin \phi \cos (\theta - \phi) \end{bmatrix}$$

$$=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \text{zero matrix.}$$

[Since θ and ϕ differ by an odd multiple of $\pi/2$, therefore $\theta - \phi$ = an odd multiple of $\pi/2$. Consequently $\cos (\theta - \phi) = 0$].

Ex. 18. *Express the following 3 equations in two unknowns in the form of a single matrix equation :*

$$a_{11}x + a_{12}\, y = b_1$$
$$a_{21}x + a_{22}\, y = b_2$$
$$a_{31}x + a_{32}\, y = b_3.$$

Solution. Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}_{3 \times 2}$ and $\mathbf{X} = \begin{bmatrix} x \\ y \end{bmatrix}_{2 \times 1}$.

Then $\mathbf{AX} = \begin{bmatrix} a_{11}x + a_{12}\, y \\ a_{21}x + a_{22}\, y \\ a_{31}x + a_{32}\, y \end{bmatrix}_{3 \times 1}$. If $\mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}_{3 \times 1}$,

then from our definition of equality of two matrices, the single matrix equation $\mathbf{AX} = \mathbf{B}$ will be equivalent to the given system of 3 equations.

Ex. 19. *If* $\mathbf{A} = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & -3 \\ 4 & -3 & -1 \end{bmatrix}$,

$\mathbf{B} = \begin{bmatrix} 1 & 4 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 2 \end{bmatrix}$ *and* $\mathbf{C} = \begin{bmatrix} 2 & 1 & -1 & -2 \\ 3 & -2 & -1 & -1 \\ 2 & -5 & -1 & 0 \end{bmatrix}$,

show that $\mathbf{AB} = \mathbf{AC}$ *though* $\mathbf{B} \neq \mathbf{C}$. **(I.C.S. 1988)**

Solution. We have

$$\mathbf{AB} = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & -3 \\ 4 & -3 & -1 \end{bmatrix}_{3 \times 3} \times \begin{bmatrix} 1 & 4 & 1 & 0 \\ 2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 2 \end{bmatrix}_{3 \times 4}$$

$$= \begin{bmatrix} 1.1-3.2+2.1 & 1.4-3.1-2.2 & 1.1-3.1+2.1 & 1.0-3.1+2.2 \\ 2.1+1.2-3.1 & 2.4+1.1+3.2 & 2.1+1.1-3.1 & 2.0+1.1-3.2 \\ 4.1-3.2-1.1 & 4.4-3.1+1.2 & 4.1-3.1-1.1 & 4.0-3.1-1.2 \end{bmatrix}$$

$$= \begin{bmatrix} -3 & -3 & 0 & 1 \\ 1 & 15 & 0 & -5 \\ -3 & 15 & 0 & -5 \end{bmatrix}_{3 \times 4}.$$

Also $\mathbf{AC} = \begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & -3 \\ 4 & -3 & -1 \end{bmatrix}_{3 \times 3} \times \begin{bmatrix} 2 & 1 & -1 & -2 \\ 3 & -2 & -1 & -1 \\ 2 & -5 & -1 & 0 \end{bmatrix}_{3 \times 4}$

$$= \begin{bmatrix} 1.2-3.3+2.2 & 1.1+3.2-2.5 & -1.1+3.1-2.1 & -1.2+3.1+2.0 \\ 2.2+1.3-3.2 & 2.1-1.2+3.5 & -2.1-1.1+3.1 & -2.2-1.1-3.0 \\ 4.2-3.3-1.2 & 4.1+3.2+1.5 & -4.1+3.1+1.1 & -4.2+3.1-1.0 \end{bmatrix}$$

$$= \begin{bmatrix} -3 & -3 & 0 & 1 \\ 1 & 15 & 0 & -5 \\ -3 & 15 & 0 & -5 \end{bmatrix}_{3 \times 4}.$$

$\therefore$ $\mathbf{AB} = \mathbf{AC}$, though $\mathbf{B} \neq \mathbf{C}$.

Ex. 20. *If* $\mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ *show that*

$$(a\mathbf{I} + b\mathbf{C})^3 = a^3\mathbf{I} + 3a^2b\mathbf{C}.$$ **(Gorakhpur 1970; Meerut 87)**

Solution. We have

$$a\mathbf{I} + b\mathbf{C} = a \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} = \mathbf{B}, \text{ say.}$$

$$\therefore \quad (a\mathbf{I} + b\mathbf{C})^2 = \mathbf{B}^2 = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} = \begin{bmatrix} a^2 & 2ab \\ 0 & a^2 \end{bmatrix}.$$

$$\therefore \quad (a\mathbf{I} + b\mathbf{C})^3 = \mathbf{B}^3 = \mathbf{B}^2\, \mathbf{B} = \begin{bmatrix} a^2 & 2ab \\ 0 & a^2 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} = \begin{bmatrix} a^3 & 3a^2b \\ 0 & a^3 \end{bmatrix}.$$

Also $a^3\mathbf{I} + 3a^2b\mathbf{C} = a^3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 3a^2b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

$$= \begin{bmatrix} a^3 & 0 \\ 0 & a^3 \end{bmatrix} + \begin{bmatrix} 0 & 3a^2b \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a^3 & 3a^2b \\ 0 & a^3 \end{bmatrix}.$$

Hence $(a\mathbf{I} + b\mathbf{C})^3 = a^3\mathbf{I} + 3a^2b\mathbf{C}$.

Ex. 21. (a) *If* $A=\begin{bmatrix}4 & 2\\ -1 & 1\end{bmatrix}$, *find* $(A-2I)(A-3I)$. **(Burdwan 1976)**

Solution. We have

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

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

$$=\begin{bmatrix}1 & 2\\ -1 & -2\end{bmatrix}.$$

$$\therefore\quad (A-2I)(A-3I)=\begin{bmatrix}2 & 2\\ -1 & -1\end{bmatrix}\begin{bmatrix}1 & 2\\ -1 & -2\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=O.$$

Ex. 21. (b) *If* $A=\begin{bmatrix}-1 & 2\\ 2 & 3\end{bmatrix}, B=\begin{bmatrix}3 & 0\\ 1 & 1\end{bmatrix}$, *verify that* $(A+B)^2=A^2+AB+BA+B^2$. *Can this be put in the simple form* $A^2+2AB+B^2$? **(Meerut 1988P)**

Solution. We have $$A+B=\begin{bmatrix}-1+3 & 2+0\\ 2+1 & 3+1\end{bmatrix}=\begin{bmatrix}2 & 2\\ 3 & 4\end{bmatrix}.$$

$$\therefore\quad (A+B)^2=\begin{bmatrix}2 & 2\\ 3 & 4\end{bmatrix}\begin{bmatrix}2 & 2\\ 2 & 4\end{bmatrix}$$

$$=\begin{bmatrix}2.2+2.3 & 2.2+2.4\\ 3.2+4.3 & 3.2+4.4\end{bmatrix}=\begin{bmatrix}10 & 12\\ 18 & 22\end{bmatrix}. \quad ...(1)$$

Again $$A^2=\begin{bmatrix}-1 & 2\\ 2 & 3\end{bmatrix}\begin{bmatrix}-1 & 2\\ 2 & 3\end{bmatrix}=\begin{bmatrix}5 & 4\\ 4 & 13\end{bmatrix},$$

$$B^2=\begin{bmatrix}3 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 0\\ 1 & 1\end{bmatrix}=\begin{bmatrix}9 & 0\\ 4 & 1\end{bmatrix},$$

$$AB=\begin{bmatrix}-1 & 2\\ 2 & 3\end{bmatrix}\begin{bmatrix}3 & 0\\ 1 & 1\end{bmatrix}=\begin{bmatrix}-1 & 2\\ 9 & 3\end{bmatrix},$$

and $$BA=\begin{bmatrix}3 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}-1 & 2\\ 2 & 3\end{bmatrix}=\begin{bmatrix}-3 & 6\\ 1 & 5\end{bmatrix}.$$

$$\therefore\quad A^2+AB+BA+B^2$$

$$=\begin{bmatrix}5 & 4\\ 4 & 13\end{bmatrix}+\begin{bmatrix}-1 & 2\\ 9 & 3\end{bmatrix}+\begin{bmatrix}-3 & 6\\ 1 & 5\end{bmatrix}+\begin{bmatrix}9 & 0\\ 4 & 1\end{bmatrix}$$

$$=\begin{bmatrix}5-1-3+9 & 4+2+6+0\\ 5+9+1+4 & 13+3+5+1\end{bmatrix}=\begin{bmatrix}10 & 12\\ 18 & 22\end{bmatrix}. \quad ...(2)$$

From (1) and (2), we see that

$$(A+B)^2=A^2+AB+BA+B^2.$$

Since here $AB\neq BA$, therefore the given relation cannot be put in the simple form

$$(A+B)^2=A^2+2AB+B^2.$$

Ex. 21. (c) *Show that* $E^2F+F^2E=E$, *where*

$$E=\begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}, F=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$ **(Meerut 1988)**

Solution. We have $$E^2=\begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}0 & 0 & 1\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix},$$ applying row-by-column rule for multiplication

$=O$ *i.e.*, null matrix of the type 3×3.

$\therefore\quad E^2F=O$ *i.e.*, null matrix of the type 3×3.

Again F is unit matrix of order 3. Therefore $F^2=F$.

$\therefore\quad F^2E=E$.

Now $E^2F+F^2E=O+E=E$, which proves the required result.

Ex. 22. *If* $A=\begin{bmatrix}3 & -4\\ 1 & -1\end{bmatrix}$, *then* $A^k=\begin{bmatrix}1+2k & -4k\\ k & 1-2k\end{bmatrix}$, *where k is any positive integer.* **(Kanpur 1985; Meerut 89, 90)**

Solution. We shall prove the result by induction on k.

We have $$A^1=A=\begin{bmatrix}3 & -4\\ 1 & -1\end{bmatrix}=\begin{bmatrix}1+2.1 & -4.1\\ 1 & 1-2.1\end{bmatrix}.$$

Thus the result is true when $k=1$.

Now suppose that the result is true for any positive integer k *i.e.*, $A^k=\begin{bmatrix}1+2k & -4k\\ k & 1-2k\end{bmatrix}$ where k is any positive integer.

Now we shall show that the result is true for $k+1$ if it is true for k. We have

$$A^{k+1}=AA^k=\begin{bmatrix}3 & -4\\ 1 & -1\end{bmatrix}\begin{bmatrix}1+2k & -4k\\ k & 1-2k\end{bmatrix}$$

$$=\begin{bmatrix}3+6k-4k & -12k-4+8k\\ 1+2k-k & -4k-1+2k\end{bmatrix}$$

$$=\begin{bmatrix}1+2+2k & -4-4k\\ 1+k & -2k-1\end{bmatrix}$$

$$=\begin{bmatrix}1+2(k+1) & -4(1+k)\\ 1+k & 1-2(1+k)\end{bmatrix}.$$

Thus the result is true for $k+1$ if it is true for k. But it is true for $k=1$. Hence by induction it is true for all positive integral values of k.

Ex. 23. *If* $A_\alpha=\begin{bmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{bmatrix}$, *then show that*

(*i*) $(A_\alpha)^n=\begin{bmatrix}\cos n\alpha & \sin n\alpha\\ -\sin n\alpha & \cos n\alpha\end{bmatrix}$, *where n is a positive integer.*

(Meerut 1983; Kanpur 87; Rohilkhand 81; Ravi Shankar 70)

(*ii*) $A_\alpha A_\beta=A_{\alpha+\beta}$.

Solution. (i) We shall prove the result by induction on n.

$$(A_\alpha)^1=A_\alpha=\begin{bmatrix}\cos 1\alpha & \sin 1\alpha\\ -\sin 1\alpha & \cos 1\alpha\end{bmatrix}.$$

Thus the result is true when $n=1$. Now suppose that the result is true for any positive integer n, *i.e.*,

$$(A_\alpha)^n=\begin{bmatrix}\cos n\alpha & \sin n\alpha\\ -\sin n\alpha & \cos n\alpha\end{bmatrix}.$$

We shall show that the result is true for $n+1$ if it is true for n.

We have, $(A_\alpha)^{n+1}=(A_\alpha)^n\, A_\alpha=\begin{bmatrix}\cos n\alpha & \sin n\alpha\\ -\sin n\alpha & \cos n\alpha\end{bmatrix}\begin{bmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{bmatrix}$

$$=\begin{bmatrix}\cos n\alpha\cos\alpha-\sin n\alpha\sin\alpha & \cos n\alpha\sin\alpha+\sin n\alpha\cos\alpha\\ -\sin n\alpha\cos\alpha-\cos n\alpha\sin\alpha & -\sin n\alpha\sin\alpha+\cos n\alpha\cos\alpha\end{bmatrix}$$

$$=\begin{bmatrix}\cos(n+1)\alpha & \sin(n+1)\alpha\\ -\sin(n+1)\alpha & \cos(n+1)\alpha\end{bmatrix}.$$

Thus the result is true for $n+1$ if it is true for n. Now the proof is complete by induction.

(ii) $A_\alpha A_\beta=\begin{bmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{bmatrix}\begin{bmatrix}\cos\beta & \sin\beta\\ -\sin\beta & \cos\beta\end{bmatrix}$

$$=\begin{bmatrix}\cos\alpha\cos\beta-\sin\alpha\sin\beta & \cos\alpha\sin\beta+\sin\alpha\cos\beta\\ -\sin\alpha\cos\beta-\cos\alpha\sin\beta & -\sin\alpha\sin\beta+\cos\alpha\cos\beta\end{bmatrix}$$

$$=\begin{bmatrix}\cos(\alpha+\beta) & \sin(\alpha+\beta)\\ -\sin(\alpha+\beta) & \cos(\alpha+\beta)\end{bmatrix}=A_{\alpha+\beta}.$$

Ex. 24. *If* B, C, *are n-rowed square matrices and if*

$$A=B+C,\ BC=CB,\ C^2=O,$$

then show that for every positive integer p,

$$A^{p+1}=B^p\,[B+(p+1)\,C].$$

Solution. We shall prove the result by induction on p.

To start the induction we see that the result is true for $p=1$.

For $A^{1+1}=A^2=(B+C)^2=(B+C)\,(B+C)=B^2+BC+CB+C^2$

$=B^2+2BC$, since $BC=CB$, $C^2=O$

$=B\,(P+2C)=B^1\,[B+(1+1)\,C]$.

Now suppose that the result is true when $p=k$. Then

$A^{k+2}=A^{k+1}.\,A=B^k\,[B+(k+1)\,C]\,[B+C]$

$=B^k\,[B^2+(k+1)\,CB+BC+(k+1)\,C^2]$

$=B^k\,[B^2+(k+2)\,BC]$, since $BC=CB$, $C^2=O$

$=B^{k+1}\,[B+(k+2)\,C]$, showing that the result is true when $p=k+1$.

Now the proof is complete by induction.

Ex. 25. *If* A *and* B *are matrices such that* $AB=BA$, *then show that for every positive integer n*

(*i*) $AB^n=B^nA$, (*ii*) $(AB)^n=A^nB^n$.

Solution. (i) We shall prove the result by induction on n.

To start the induction we see that the result is true when $n=1$. For $AB^1=AB=BA=B^1A$.

Now suppose that the result is true for any positive integer n.

Then $AB^{n+1}=(AB^n)\,B=(B^nA)\,B=B^n\,(AB)$

$=B^n\,(BA)=(B^nB)\,A=B^{n+1}\,A$,

showing that the result is true for $n+1$.

Now the proof is complete by induction.

(ii) We shall prove the result by induction on n. To start the induction we see that the result is true when $n=1$. For

$$(AB)^1=AB=A^1B^1.$$

Now suppose that the result is true for any positive integer n *i.e.* $(AB)^n=A^nB^n$. Then

$(AB)^{n+1}=(AB)^n\,(AB)=(A^nB^n)\,(AB)=A^n\,(B^nA)\,B$

$=A^n\,(AB^n)\,B$ [by part (i) of the question]

$=A^{n+1}\,B^{n+1}$, showing that the result is true for $n+1$.

The proof is now complete by induction.

Ex. 26. *If* A *and* B *be m-rowed square matrices which commute and n be a positive integer prove the binomial theorem*

$$(A+B)^n={}^nc_0\,A^n+{}^nc_1\,A^{n-1}\,B+\ldots+{}^nc_r\,A^{n-r}\,B^r+\ldots+{}^nc_nB^n.$$

Solution. We have $A+B=A^1+B^1$.

Now $(A+B)^2=(A+B)\,(A+B)$

$=A^2+AB+BA+B^2$, by distributive law

$=A^2+2AB+B^2$, since $AB=BA$

$={}^2c_0\,A^2+{}^2c_1\,AB+{}^2c_2\,B^2$.

Thus the theorem is true for $n=2$.

Now assume that the theorem is true for n *i.e.*

$(A+B)^n={}^nc_0\,A^n+{}^nc_1\,A^{n-1}\,B+\ldots+{}^nc_r\,A^{n-r}\,B^r$

$+{}^nc_{r+1}\,A^{n-r-1}\,B^{r+1}+\ldots+{}^nc_n\,B^n.$

Then $(A+B)^{n+1}=(A+B)\,(A+B)^n$

$$=(A+B)\,({}^nc_0A^n+{}^nc_1A^{n-1}\,B+\ldots+{}^nc_rA^{n-r}B^r+{}^nc_{r+1}A^{n-r-1}\,B^{r+1}+\ldots+{}^nc_nB^n)$$

$$={}^nc_0\,A^{n+1}+({}^nc_0\,BA^n+{}^nc_1\,A^nB)+\ldots+({}^nc_r\,BA^{n-r}\,B^r+{}^nc_{r+1}\,A^{n-r}\,B^{r+1})+\ldots+{}^nc_n\,B^{n+1}.$$

Now $AB=BA$. We can prove by induction that for every positive integer n, $BA^n=A^nB$. [Refer Ex. 25].

Again $BA^{n-r}\,B^r=(BA^{n-r})\,B^r=(A^{n-r}\,B)\,B^r$
$=A^{n-r}\,BB^r=A^{n-r}\,B^{r+1}$.

Also ${}^nc_0={}^{n+1}c_0=1$, ${}^nc_n={}^{n+1}c_{n+1}=1$,

and ${}^nc_r+{}^nc_{r+1}={}^{n+1}c_{r+1}$.

Hence $(A+B)^{n+1}={}^{n+1}c_0\,A^{n+1}+({}^nc_0+{}^nc_1)\,A^nB+\ldots+({}^nc_r+{}^nc_{r+1})\,A^{n-r}\,B^{r+1}+\ldots+{}^{n+1}c_{n+1}\,B^{n+1}$

$$={}^{n+1}c_0\,A^{n+1}+{}^{n+1}c_1\,A^nB+\ldots+{}^{n+1}c_{r+1}\,A^{n-r}\,B^{r+1}+\ldots+{}^{n+1}c_{n+1}\,B^{n+1}.$$

Thus the theorem is true for $n+1$, if it is true for n. But it is true for $n=2$. Hence it is true for all positive integral values of n.

Ex. 27. *If* $A=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$, *prove that*

$$(aI+bA)^n=a^nI+na^{n-1}\,bA,$$

where **I** *is the two rowed unit matrix and n is a positive integer.*

Solution. We shall prove the result by induction on n. To start the induction we see that the result is true for $n=1$. For

$$(aI+bA)^1=aI+bA=a^1I+1a^{1-1}\,bA.$$

Now suppose that the result is true for any positive integer n. Then $(aI+bA)^{n+1}=(aI+bA)^n\,(aI+bA)$

$=(a^nI+na^{n-1}\,bA)\,(aI+bA)$

[$\because$ by assumption the result is true for n]

$=a^{n+1}\,I+a^nbIA+na^nbAI+na^{n-1}\,b^2A^2$

$=a^{n+1}\,I+a^nbA+na^nbA+O$, since $IA=A=AI$, $A^2=O$

$=a^{n+1}\,I+(n+1)\,a^{n+1-1}\,bA$, showing that the result is true for $n+1$. The proof is now complete by induction.

Ex. 28. *If* $A=\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}$, *show that*

$$A^2-4A-5I=O.\quad \textbf{(Meerut 1973, 77; Agra 78)}$$

Solution. We have $A^2=\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}\times\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}$

$$=\begin{bmatrix}1+4+4 & 2+2+4 & 2+4+2\\ 2+2+4 & 4+1+4 & 4+2+2\\ 2+4+2 & 4+2+2 & 4+4+1\end{bmatrix}=\begin{bmatrix}9 & 8 & 8\\ 8 & 9 & 8\\ 8 & 8 & 9\end{bmatrix}.$$

Now $A^2-4A-5I$

$$=\begin{bmatrix}9 & 8 & 8\\ 8 & 9 & 8\\ 8 & 8 & 9\end{bmatrix}-4\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}-5\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix}9 & 8 & 8\\ 8 & 9 & 8\\ 8 & 8 & 9\end{bmatrix}-\begin{bmatrix}4 & 8 & 8\\ 8 & 4 & 8\\ 8 & 8 & 4\end{bmatrix}-\begin{bmatrix}5 & 0 & 0\\ 0 & 5 & 0\\ 0 & 0 & 5\end{bmatrix}$$

$$=\begin{bmatrix}9 & 8 & 8\\ 8 & 9 & 8\\ 8 & 8 & 9\end{bmatrix}-\begin{bmatrix}9 & 8 & 8\\ 8 & 9 & 8\\ 8 & 8 & 9\end{bmatrix}=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}.$$

Ex. 29. *Show that the product of two triangular matrices is itself triangular.*

Solution. Let $A=[a_{ij}]_{n\times n}$ and $B=[b_{jk}]_{n\times n}$ be two triangular matrices each of order n. Then $a_{ij}=0$ when $i>j$. Also $b_{jk}=0$ when $j>k$.

Let $AB=[c_{ik}]_{n\times n}$. Then $c_{ik}=\sum_{j=1}^{n} a_{ij}\,b_{jk}$.

Suppose that $i>k$.

If $j<i$, then $a_{ij}=0$ and therefore $c_{ik}=0$.

If $j>i$, then $j>k$ because $i>k$. In this case $b_{jk}=0$ and therefore $c_{ik}=0$.

Thus $c_{ik}=0$ whenever $i>k$.

Hence the matrix AB is also a triangular matrix.

Ex. 30. *If* **A** *and* **B** *are matrices conformable for multiplication, then show that*

(*i*) $(-A)\,(B)=-(AB)$, (*ii*) $A\,(-B)=-(AB)$.

Solution. (i) We have $(-A+A)\,B=OB$
$\Rightarrow(-A)\,B+AB=O\Rightarrow(-A)\,B=-(AB)$.

(ii) We have $A\,(-B+B)=AO\Rightarrow A\,(-B)+AB=O$
$\Rightarrow A\,(-B)=-(AB)$.

Ex. 31. *Show that the pre-multiplication (or post-multiplication) of a square matrix* **A** *by a diagonal matrix* **D** *multiplies each row (or column) of* **A** *by the corresponding diagonal elements of* **D**.

Deduce that the only matrices commutative with a diagonal matrix with distinct diagonal elements are diagonal matrices.

Solution. Let $A=[a_{ij}]_{n\times n}$ be a square matrix

i.e. $A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$.

Let $D=\begin{bmatrix} b_{11} & 0 & 0 & \dots & 0 \\ 0 & b_{22} & 0 & \dots & 0 \\ 0 & 0 & b_{33} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{nn} \end{bmatrix}$

be a diagonal matrix of order n.

Then $DA=\begin{bmatrix} b_{11}a_{11} & b_{11}a_{12} & \dots & b_{11}a_{1n} \\ b_{22}a_{21} & b_{22}a_{22} & \dots & b_{22}a_{2n} \\ \dots\dots & \dots\dots & \dots & \dots\dots \\ b_{nn}a_{n1} & b_{nn}a_{n2} & \dots & b_{nn}a_{nn} \end{bmatrix}$.

Thus in the product DA the first row of A has been multiplied by the corresponding diagonal element b_{11} of the first row of D and so on.

Again $AD=\begin{bmatrix} a_{11}\,b_{11} & a_{12}\,b_{22} & \dots & a_{1n}\,b_{nn} \\ a_{21}\,b_{11} & a_{22}\,b_{22} & \dots & a_{2n}\,b_{nn} \\ \dots & \dots & \dots & \dots \\ a_{n1}\,b_{11} & a_{n2}\,b_{22} & \dots & a_{nn}\,b_{nn} \end{bmatrix}$.

From this it is obvious that in the product AD the first column of A has been multiplied by the corresponding diagonal element b_{11} of the first column of D and so on. Hence the given result is true for post-multiplication of A by D.

Second part. Let $A=[a_{ij}]_{n\times n}$ commute with the diagonal matrix D of order n having its diagonal elements all distinct. Let $D=\text{diag}\,[b_{11},\ \dots,\ b_{nn}]$.

The $(i,\ j)^{th}$ element of $DA=a_{ij}b_{ii}$.

Also the $(i,\ j)^{th}$ element of $AD=a_{ij}b_{jj}$.

But $AD=DA$.

$\therefore\ a_{ij}\,b_{ii}=a_{ij}\,b_{jj}$ or $a_{ij}\,(b_{ii}-b_{jj})=0$.

But if $i\neq j$, then $b_{ii}\neq b_{jj}$. Therefore $a_{ij}=0$ if $i\neq j$.

Thus each non-diagonal element of A is zero. Therefore A is a diagonal matrix.

Ex. 32. *Show that if a diagonal matrix is commutative with every matrix of the same order, then it is necessarily a scalar matrix.*

Solution. Let $D=\text{diag}\,[b_{11},\ b_{22},\ \dots,\ b_{nn}]$ be a diagonal matrix of order n. Let $A=[a_{ij}]_{n\times n}$ be any square matrix of order n. Then it is given that $AD=DA$.

The $(i,\ j)^{th}$ element of $DA=a_{ij}\ b_{ii}$.

Also the $(i,\ j)^{th}$ element of $AD=a_{ij}\ b_{jj}$.

Since $AD=DA$, therefore

$$a_{ij}\ b_{ii}=a_{ij}\ b_{jj} \quad\text{or}\quad a_{ij}\ (b_{ii}-b_{jj})=0.$$

Since A is any square matrix of order n, therefore we can take $a_{ij}\neq 0$. Therefore we must have

$$b_{ii}-b_{jj}=0 \text{ or } b_{ii}=b_{jj} \text{ for each } i \text{ and } j.$$

Thus the diagonal elements of D are all equal. Therefore D is a scalar matrix.

Ex. 33. *Find the possible square roots of the two rowed unit matrix* **I.**

Solution. Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a square root of the matrix $I=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then $A^2=I$,

i.e. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

i.e. $\begin{bmatrix} a^2+bc & ab+bd \\ ac+cd & cb+d^2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Since the above matrices are equal, therefore

$a^2+bc=1$ ...(i) $\qquad ac+cd=0$ ...(iii)

$ab+bd=0$ ...(ii) $\qquad cb+d^2=1$ ...(iv)

must hold simultaneously.

If $a+d=0$, the above four equations hold simultaneously if $d=-a$ and $a^2+bc=1$.

Hence one possible square root of I is

$A=\begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix}$ where α,β,γ are any three numbers related by the condition $\alpha^2+\beta\gamma=1$.

If $a+d\neq 0$, the above four equations hold simultaneously if $b=0,\ c=0,\ a=1,\ d=1$ or if $b=0,\ c=0,\ a=-1,\ d=-1$.

Hence $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

i.e. $\pm I$ are other possible square roots of I.

Ex. 34. *A matrix A such that* $A^2=A$ *is called* **Idempotent.** *Determine all the idempotent diagonal matrices of order n.*

Solution. Let $A=\text{diag.}\ [d_1, d_2, d_3, \ldots, d_n]$ be an idempotent matrix of order n.

Then $A^2=A$.

$$\text{Now}\quad A^2=\begin{bmatrix} d_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & d_n \end{bmatrix}\begin{bmatrix} d_1 & 0 & 0 & \ldots & 0 \\ 0 & d_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & d_n \end{bmatrix}$$

$$=\begin{bmatrix} d_1^2 & 0 & 0 & \ldots & 0 \\ 0 & d_2^2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 0 & \ldots & d_n^2 \end{bmatrix}$$

$\therefore$ $A^2=A$ gives

$d_1^2=d_1,\ d_2^2=d_2, \ldots, d_n^2=d_n$

i.e. $d_1=0, 1;\ d_2=0, 1; \ldots; d_n=0, 1.$

Hence $\text{Diag.}\ [d_1, d_2, \ldots, d_n]$, $d_1, d_2, \ldots, d_n=0,\ 1$ is the required idempotent diagonal matrix of order n.

Ex. 35. *If* $AB=A$ *and* $BA=B$, *show that A and B are idempotent.*

Solution. We have $AB=A$

$\Rightarrow A(BA)=A$ $[\because BA=B]$

$\Rightarrow (AB)A=A$

$\Rightarrow AA=A$ $[\because AB=A]$

$\Rightarrow A^2=A \Rightarrow A$ is idempotent.

Again $BA=B$

$\Rightarrow B(AB)=B$ $[\because AB=A]$

$\Rightarrow (BA)B=B$

$\Rightarrow BB=B \Rightarrow B^2=B$

$\Rightarrow$ B is idempotent.

Ex. 36. *Show that the matrix*

$$A=\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}\ \textit{is idempotent.}\qquad \textbf{(Agra 1978)}$$

Solution. We have

$$A^2=AA=\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}\times\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$$

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

Since $A^2=A$, therefore the matrix A is an idempotent matrix.

Ex. 37. *If B is an idempotent matrix, show that* $A=I-B$ *is also idempotent and that* $AB=BA=O$.

Solution. Since B is idempotent, therefore $B^2=B$.

Now $A^2=(I-B)^2=(I-B)(I-B)$

$=I-IB-BI+B^2$

$=I-B-B+B^2$ $[\because IB=B=BI]$

$=I-B-B+B$ $[\because B^2=B]$

$=I-B$ $[\because -B+B=O]$

$=A.$

Since $A^2=A$, therefore A is also idempotent.

Again $AB=(I-B)B=IB-B^2=B-B=O$.

Similarly $BA=B(I-B)=BI-B^2=B-B=O$.

Ex. 38. *A matrix A such that* $A^2=I$ *is called* **involutory.** *Show that A is involutory, if and only if* $(I+A)(I-A)=O$.

Solution. Let A be an involutory matrix of order n.

Then $A^2=I$.

$\therefore$ $I-A^2=O$

$\therefore$ $I^2-A^2=O.$ $[\because I^2=I]$

$\therefore$ $(I+A)(I-A)=O.$ $[\because AI=IA]$

Conversely if $(I+A)(I-A)=O$,

then $I^2-IA+AI-A^2=O$

or $I-A^2+AI-AI=O$ $[\because AI=IA]$

or $I-A^2+O=O$

or $I-A^2=O$

or $A^2=I$.

Ex. 39. *A square matrix A is called a* **nilpotent matrix** *if there exists a positive integer m such that* $A^m=O$. *If m is the least positive integer such that* $A^m=O$, *then m is called the* **index** *of the nilpotent matrix A. Show that the matrix*

$$A=\begin{bmatrix} ab & b^2 \\ -a^2 & -ab \end{bmatrix}$$

is nilpotent such that $A^2=O$. *Conversely, show that all 2-rowed nilpotent matrices such that* $A^2=O$ *are of the above form.*

Solution. If $A=\begin{bmatrix} ab & b^2 \\ -a^2 & -ab \end{bmatrix}$,

then $A^2=AA=\begin{bmatrix} ab & b^2 \\ -a^2 & -ab \end{bmatrix}\begin{bmatrix} ab & b^2 \\ -a^2 & -ab \end{bmatrix}$

$$=\begin{bmatrix} a^2b^2-b^2a^2 & ab^3-ab^3 \\ -a^3b+a^3b & -a^2b^2+a^2b^2 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Hence the matrix A is nilpotent of the index 2. Conversely, suppose $A=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$ is a two rowed nilpotent matrix such that $A^2=O$,

i.e. $\begin{bmatrix} \alpha^2+\beta\gamma & \alpha\beta+\beta\delta \\ \alpha\gamma+\gamma\delta & \gamma\beta+\delta^2 \end{bmatrix}=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

If the above equation holds,

$$\alpha^2+\beta\gamma=0 \quad \ldots(i) \qquad \alpha\gamma+\gamma\delta=0 \quad \ldots(iii)$$
$$\alpha\beta+\beta\delta=0 \quad \ldots(ii) \qquad \gamma\beta+\delta^2=0 \quad \ldots(iv)$$

must hold simultaneously.

From (i) and (iv), it is obvious that the above four equations will hold simultaneously if $\alpha^2=\delta^2$ *i.e.* $\alpha=\pm\delta$. If $\alpha=-\delta$, the above four equations will hold simultaneously for all values of the numbers α, β, γ provided they are related by the condition $\alpha^2+\beta\gamma=0$.

If $\alpha=\delta$ and either of them is not equal to 0, we have from (ii) and (iii) $\beta=0$ and $\gamma=0$. Then from (i) and (iv), we shall have $\alpha=\delta=0$. Hence if $\alpha=\delta\neq 0$, the above four equations cannot hold simultaneously.

If $\alpha=\delta=0$, the equations can hold simultaneously, but this solution can be included in the solution $\alpha=-\delta$, $\alpha^2+\beta\gamma=0$.

Thus the matrix A is nilpotent if

$A=\begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix}$ and α, β, γ are any numbers related by the condition $\alpha^2+\beta\gamma=0$.

Obviously the matrix A is of the given form.

Ex. 40. *Show that the matrix*

$$A=\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$$

is nilpotent and find its index. **(Sagar 1964)**

Solution. We have

$$A^2=AA=\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}\times\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$$

$$=\begin{bmatrix} 0 & 0 & 0 \\ 3 & 3 & 9 \\ -1 & -1 & -3 \end{bmatrix}$$

Again

$$A^3=AA^2=\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 3 & 3 & 9 \\ -1 & -1 & -3 \end{bmatrix}$$

$$=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}=O.$$

Thus 3 is the least positive integer such that $A^3=O$. Hence the matrix A is nilpotent of index 3.

Ex. 41. *Show that the matrix*

$$A=\begin{bmatrix} -5 & -8 & 0 \\ 3 & 5 & 0 \\ 1 & 2 & -1 \end{bmatrix}$$ *is involutory.* **(Rohilkhand 1991)**

Sol. Find A^2 *i.e.*, AA. We shall get $A^2=I$ *i.e.*, the unit matrix of order 3. Hence the matrix A is involutory.

Exercises

1. Can the following two matrices be multiplied and if so compute their product

$$\begin{bmatrix} 4 & 2 & -1 & 2 \\ 3 & -7 & 1 & -8 \\ 2 & 4 & -3 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ -3 & 0 \\ 1 & 5 \\ 3 & 1 \end{bmatrix}?$$ **(Lucknow 1965)**

2. If $A=\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B=\begin{bmatrix} 1 & 7 \\ 2 & 3 \\ 5 & 9 \end{bmatrix}$, find the matrices AB and BA.

3. If $A=\begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix}$, $B=\begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix}$ show that AB=BA. **(Meerut 1977)**

4. Examine if AB=BA, where

$$A=\begin{bmatrix} -2 & 3 & -1 \\ -1 & 2 & -1 \\ -6 & 9 & -4 \end{bmatrix} \text{ and } B=\begin{bmatrix} 1 & 3 & -1 \\ 2 & 2 & -1 \\ 3 & 0 & -1 \end{bmatrix}.$$ **(Meerut 1973)**

5. If $\begin{bmatrix} 4 & 1 & 2 \\ 0 & 5 & 3 \end{bmatrix}\begin{bmatrix} 3 & 4 & 5 \\ -1 & 0 & -2 \\ 3 & 4 & 7 \end{bmatrix}=\begin{bmatrix} 8x+3y & 6z & 32 \\ 4 & 12 & 26x-5y \end{bmatrix}$ find the values of x, y and z.

6. Give an example of two matrices A, B such that $AB\neq BA$. Give also an example of matrices A, B such that $AB=O$ but $A\neq O, B\neq O$. **(Poona 1970)**

7. Is the following statement true? Give reasons in favour of your answer:

A, B are n-rowed square matrices.
Then $AB = O \Rightarrow$ at least one of A and B is zero.

(Gujrat 1970)

8. For the 3 matrices A, B, C,

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ B = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

verify the following relations:

$$A^2 = B^2 = C^2 = I,$$

$$AB = -BA;\ AC = -CA;\ BC = -CB.$$

9. Given

$$A = \begin{bmatrix} 1 & 1 & -1 \\ 2 & -3 & 4 \\ 3 & -2 & 3 \end{bmatrix},\ B = \begin{bmatrix} -1 & -2 & -1 \\ 6 & 12 & 6 \\ 5 & 10 & 5 \end{bmatrix}$$

and $C = \begin{bmatrix} -1 & -1 & 1 \\ 2 & 2 & -2 \\ -3 & -3 & 3 \end{bmatrix}$, prove that

$AB = O,\ BA \neq O,\ AC \neq O,\ CA = O.$ **(Rajasthan 1964)**

10. Given

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\ B = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix},\ C = \begin{bmatrix} 5 & 1 \\ 7 & 4 \end{bmatrix}, \text{ verify that}$$

$$A\,(B+C) = AB + AC$$

and $(A+B)\,C = AC + BC.$ **(Agra 1973)**

11. Show that the matrix

$$A = \begin{bmatrix} 1 & -3 & -4 \\ -1 & 3 & 4 \\ 1 & -3 & -4 \end{bmatrix} \text{ is nilpotent of index 2.}$$

12. Let $f(x) = x^2 - 5x + 6$, find $f(A)$ *i.e.* $A^2 - 5A + 6I$

if $A = \begin{bmatrix} 2 & 0 & 1 \\ 2 & 1 & 3 \\ 1 & -1 & 0 \end{bmatrix}$. **(Delhi 1970)**

13. If $A = \begin{bmatrix} \cosh u & \sinh u \\ \sinh u & \cosh u \end{bmatrix}$, then show that

$$A^n = \begin{bmatrix} \cosh nu & \sinh nu \\ \sinh nu & \cosh nu \end{bmatrix}, \text{ where } n \text{ is a positive integer.}$$

(Allahabad 1970)

14. If $AB = A$ and $BA = B$, then show that $A^2 = A,\ B^2 = B$.

(Poona 1972)

15. Show that the sum of the two idempotent matrices A and B is idempotent if $AB = BA = O$. **(Sagar 1968)**

16. Prove that the matrix

$$\begin{bmatrix} \lambda_1^2 & \lambda_1\lambda_2 & \lambda_1\lambda_3 \\ \lambda_1\lambda_2 & \lambda_2^2 & \lambda_2\lambda_3 \\ \lambda_1\lambda_3 & \lambda_2\lambda_3 & \lambda_3^2 \end{bmatrix}$$

is idempotent, where $\lambda_1, \lambda_2, \lambda_3$ are the direction cosines *i.e.*,

$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1.$$ **(Sagar 1968)**

Answers

1. Yes; the product is $\begin{bmatrix} 7 & 9 \\ 4 & 6 \\ -8 & -8 \end{bmatrix}$.

2. AB does not exist and $BA = \begin{bmatrix} 22 & 30 \\ 11 & 16 \\ 32 & 46 \end{bmatrix}$.

4. Yes. 5. $x = 1,\ y = 3,\ z = 4.$

7. No. 12. $\begin{bmatrix} 1 & -1 & -3 \\ -1 & -1 & -10 \\ -5 & 4 & 4 \end{bmatrix}$.

§ 15. Trace of a Matrix. **(Sagar 1964, 68)**

Definition. *Let A be a square matrix of order n. The sum of the elements of A lying along the principal diagonal is called the trace of A.* We shall write the trace of A as tr A. Thus if $A = [a_{ij}]_{n \times n}$, then $\text{tr } A = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn}$.

In the following theorem we have given some fundamental properties of the trace function.

Theorem. *Let A and B be two square matrices of order n and λ be a scalar. Then*

(1) $tr\,(\lambda A) = \lambda\ tr\ A$; **(Sagar 1964, 68)**

(2) $tr\,(A+B) = tr\ A + tr\ B$; **(Sagar 1964, 68)**

(3) $tr\,(AB) = tr\,(BA)$.

Proof. Let $A = [a_{ij}]_{n \times n}$ and $B = [b_{ij}]_{n \times n}$.

(1) We have $\lambda A = [\lambda a_{ij}]_{n \times n}$, by def. of multiplication of a matrix by a scalar.

$$\therefore\ \text{tr}\,(\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda\ \text{tr } A.$$

(2) We have $A + B = [a_{ij} + b_{ij}]_{n \times n}$.

$\therefore \quad \text{tr}\,(A+B)=\sum_{i=1}^{n}(a_{ii}+b_{ii})=\sum_{i=1}^{n}a_{ii}+\sum_{i=1}^{n}b_{ii}=\text{tr}\,A+\text{tr}\,B.$

(3) We have $AB=[c_{ij}]_{n\times n}$ where $c_{ij}=\sum_{k=1}^{n}a_{ik}\,b_{kj}$.

Also $BA=[d_{ij}]_{n\times n}$ where $d_{ij}=\sum_{k=1}^{n}b_{ik}\,a_{kj}$.

Now $\text{tr}\,(AB)=\sum_{i=1}^{n}c_{ii}=\sum_{i=1}^{n}\left(\sum_{k=1}^{n}a_{ik}\,b_{ki}\right)$

$=\sum_{k=1}^{n}\sum_{i=1}^{n}a_{ik}\,b_{ki}$, interchanging the order of summation in the last sum

$=\sum_{k=1}^{n}\left(\sum_{i=1}^{n}b_{ki}\,a_{ik}\right)=\sum_{k=1}^{n}d_{kk}=d_{11}+d_{22}+\ldots+d_{nn}=\text{tr}\,(BA).$

§ 16. Transpose of a Matrix. (Definition).

Let $A=[a_{ij}]_{m\times n}$. *Then the* $n\times m$ *matrix obtained from* A *by changing its rows into columns and its columns into rows is called the transpose of* A *and is denoted by the symbol* A′ *or* A^T.

(Meerut 1989, Sagar 65, Punjab 71, Agra 71, Kolhapur 72, Lucknow 79, Vikram 66, Bombay 66, Allahabad 66)

The operation of interchanging rows with columns is called transposition. Symbolically if

$A=[a_{ij}]_{m\times n}$,

then $A'=[b_{ji}]_{n\times m}$, where $b_{ji}=a_{ij}$,

i.e. the $(j, i)^{th}$ element of A′ is the $(i, j)^{th}$ element of A.

For example, the transpose of the 3×4 matrix

$$A=\begin{bmatrix}1 & 2 & 3 & 4\\ 2 & 3 & 4 & 1\\ 3 & 4 & 2 & 1\end{bmatrix}_{3\times 4}$$

is the 4×3 matrix

$$A'=\begin{bmatrix}1 & 2 & 3\\ 2 & 3 & 4\\ 3 & 4 & 2\\ 4 & 1 & 1\end{bmatrix}_{4\times 3}.$$

The first row of A is the first column of A′. The second row of A is the second column of A′. The third row of A is the third column of A

§ 17. *If* A′ *and* B′ *be the transposes of* A *and* B *respectively, then*

(*i*) $(A')'=A$; **(Meerut 1983)**

(*ii*) $(A+B)'=A'+B'$, *A and B being of the same size.* **(Meerut 1988)**

(*iii*) $(kA)'=kA'$, *k being any complex number.* **(Meerut 1988)**

*(*iv*) $(AB)'=B'A'$, *A and B being conformable to multiplication.*

(Rohilkhand 1977, Sagar 71, Kanpur 89, Kerala 71, Kolhapur 72, Allahabad 76, Punjab 71, Ravi Shanker 71, Agra 74)

Proof.

(i) Let A be an $m\times n$ matrix. Then A′ will be an $n\times m$ matrix. Therefore $(A')'$ will be an $m\times n$ matrix. Thus the matrices A and $(A')'$ are of the same type.

Also the $(i, j)^{th}$ element of $(A')'$

$=$ the $(j, i)^{th}$ element of $A'=$ the $(i, j)^{th}$ element of A.

Hence $A=(A')'$.

(ii) Let $A=[a_{ij}]_{m\times n}$ and $B=[b_{ij}]_{m\times n}$. Then $A+B$ will be a matrix of the type $m\times n$ and consequently $(A+B)'$ will be a matrix of the type $n\times m$.

Again A′ and B′ are both $n\times m$ matrices. Therefore the sum $A'+B'$ exists and will also be a matrix of the type $n\times m$.

Further the $(j, i)^{th}$ element of $(A+B)'$

$=$ the $(i, j)^{th}$ element of $A+B=a_{ij}+b_{ij}$

$=$ the $(i, j)^{th}$ element of A $+$ the $(i, j)^{th}$ element of B

$=$ the $(j, i)^{th}$ element of A′ $+$ the $(j, i)^{th}$ element of B′

$=$ the $(j, i)^{th}$ element of $A'+B'$.

Thus the matrices $(A+B)'$ and $A'+B'$ are of the same type and their $(j, i)^{th}$ elements are equal. Hence $(A+B)'=A'+B'$.

(iii) Let $A=[a_{ij}]_{m\times n}$. If k is any complex number, then kA will also be an $m\times n$ matrix and consequently $(kA)'$ will be an $n\times m$ matrix.

Again A′ will be an $n\times m$ matrix and therefore kA' will also be an $n\times m$ matrix. Further the $(j, i)^{th}$ element of $(kA)'$

$=$ the $(i, j)^{th}$ element of $kA=k.(i, j)^{th}$ element of A

$=k.(j, i)^{th}$ element of $A'=$ the $(j, i)^{th}$ element of kA'

Thus the matrices $(kA)'$ and kA' are of the same size and their $(j, i)^{th}$ elements are equal. Hence $(kA)'=kA'$.

(iv) Let $A=[a_{ij}]_{m\times n}$ and $B=[b_{jk}]_{n\times p}$. Then
$A'=[c_{ji}]_{n\times m}$ where $c_{ji}=a_{ij}$
and $B'=[d_{kj}]_{p\times n}$ where $d_{kj}=b_{jk}$.

The matrix **AB** will be of the type $m\times p$. Therefore the matrix (AB)′ will be of the type $p\times m$.

Again the matrix A′ will be of the type $n\times m$ and the matrix B′ will be of the type $p\times n$. Therefore the product B′A′ exists and will be a matrix of the type $p\times m$. Thus the matrices (AB)′ and B′A′ are of the same type.

Now the $(k, i)^{th}$ element of (AB)′

$$=\text{the } (i, k)^{th} \text{ element of } AB=\sum_{j=1}^{n} a_{ij}\, b_{jk}$$

$$=\sum_{j=1}^{n} c_{ji}\, d_{kj}=\sum_{j=1}^{n} d_{kj}\, c_{ji}=\text{the } (k, i)^{th} \text{ element of } B'A'.$$

Thus the matrices (AB)′ and B′A′ are of the same size and their $(k, i)^{th}$ elements are equal. Hence (AB)′=B′A′.

The above law is called the reversal law for transposes *i.e. the transpose of the product is the product of the transposes taken in the reverse order.*

Example. *If* $A=\begin{bmatrix}2 & 3\\ 0 & 1\end{bmatrix}$, $B=\begin{bmatrix}3 & 4\\ 2 & 1\end{bmatrix}$,

verify that (AB)′=B′A′. **(Agra 1980)**

Solution. We have

$$AB=\begin{bmatrix}2 & 3\\ 0 & 1\end{bmatrix}\times\begin{bmatrix}3 & 4\\ 2 & 1\end{bmatrix}=\begin{bmatrix}2.3+3.2 & 2.4+3.1\\ 0.3+1.2 & 0.4+1.1\end{bmatrix}=\begin{bmatrix}12 & 11\\ 2 & 1\end{bmatrix}.$$

$$\therefore\quad (AB)'=\begin{bmatrix}12 & 2\\ 11 & 1\end{bmatrix}.$$

Also $$B'A'=\begin{bmatrix}3 & 2\\ 4 & 1\end{bmatrix}\times\begin{bmatrix}2 & 0\\ 3 & 1\end{bmatrix}$$

$$=\begin{bmatrix}3.2+2.3 & 3.0+2.1\\ 4.2+1.3 & 4.0+1.1\end{bmatrix}=\begin{bmatrix}12 & 2\\ 11 & 1\end{bmatrix}.$$

Hence (AB)′=B′A′.

§ 18. Conjugate of a Matrix.

If $i=\sqrt{(-1)}$, then $z=x+iy$ is called a complex number where x and y are any real numbers. If $z=x+iy$ then $\bar{z}=x-iy$ is called the *conjugate* of the complex number z.

We have $z\bar{z}=(x+iy)\,(x-iy)=x^2+y^2$ *i.e.*, is real.

Also if $z=\bar{z}$, then $x+iy=x-iy$

i.e., $iy=0$ *i.e.*, $y=0$ *i.e.*, z is real.

Conversely if z is real, then $\bar{z}=z$.

If $z=x+iy$, then $\bar{z}=x-iy$.

$\therefore\quad \overline{(\bar{z})}=x+iy=z.$

If z_1 and z_2 are two complex numbers, then it can be easily seen that

(i) $\overline{(z_1+z_2)}=\bar{z}_1+\bar{z}_2$, and (ii) $\overline{z_1 z_2}=(\bar{z}_1)\,(\bar{z}_2)$.

Conjugate of a matrix. Definition. **(Jodhpur 1965)**

The matrix obtained from any given matrix A *on replacing its elements by the corresponding conjugate complex numbers is called the conjugate of* A *and is denoted by* $\bar{A}$.

Thus if $A=[a_{ij}]_{m\times n}$, then

$\bar{A}=[\bar{a}_{ij}]_{m\times n}$ where $\bar{a}_{ij}$ denotes the conjugate complex of a_{ij}.

If A be a matrix over the field of real numbers, then obviously $\bar{A}$ coincides with A.

Illustration. If

$$A=\begin{bmatrix}2+3i & 4-7i & 8\\ -i & 6 & 9+i\end{bmatrix},\ \text{then } \bar{A}=\begin{bmatrix}2-3i & 4+7i & 8\\ i & 6 & 9-i\end{bmatrix}.$$

Theorem. *If* $\bar{A}$ *and* $\bar{B}$ *be the conjugates of* A *and* B *respectively, then*

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

(*iii*) $\overline{(kA)}=\bar{k}\,\bar{A}$, *k being any complex number* ;

(*iv*) $\overline{(AB)}=\bar{A}\,\bar{B}$, *A and B being conformable to multiplication.*

Proof. (i) Let $A=[a_{ij}]_{m\times n}$.

Then $\bar{A}=\left[\ \bar{a}_{ij}\ \right]_{m\times n}$, where $\bar{a}_{ij}$ is the conjugate complex of a_{ij}.

Obviously both A and $\overline{(\bar{A})}$ are matrices of the same type $m\times n$.

The $(i, j)^{th}$ element of $\overline{(\bar{A})}$

=the conjugate complex of the $(i, j)^{th}$ element of $\bar{A}$

=the conjugate complex of $\bar{a}_{ij}$

$=\overline{(\bar{a}_{ij})}=a_{ij}=$ the $(i, j)^{th}$ element of A.

Hence $\overline{\left(\overline{A}\right)}=A$.

(ii) Let $A=[a_{ij}]_{m\times n}$ and $B=[b_{ij}]_{m\times n}$.

Then $\overline{A}=[\overline{a}_{ij}]_{m\times n}$ and $\overline{B}=[\overline{b}_{ij}]_{m\times n}$.

First we see that both $\overline{(A+B)}$ and $\overline{A}+\overline{B}$ are $m\times n$ matrices.

Again the $(i, j)^{th}$ element of $\overline{(A+B)}$

= the conjugate complex of the $(i, j)^{th}$ element of $A+B$

= the conjugate complex of $a_{ij}+b_{ij}$

$=\overline{(a_{ij}+b_{ij})}=\overline{a}_{ij}+\overline{b}_{ij}$

= the $(i, j)^{th}$ element of $\overline{A}$ + the $(i, j)^{th}$ element of $\overline{B}$

= the $(i, j)^{th}$ element of $\overline{A}+\overline{B}$.

Hence $\overline{(A+B)}=\overline{A}+\overline{B}$.

(iii) Let $A=[a_{ij}]_{m\times n}$. If k is any complex number, then both $\overline{(kA)}$ and $\overline{k}\,\overline{A}$ will be $m\times n$ matrices.

The $(i, j)^{th}$ element of $\overline{(kA)}$

= the conjugate complex of the $(i, j)^{th}$ element of kA

= the conjugate complex of $ka_{ij}=\overline{(ka_{ij})}=\overline{k}\,\overline{a}_{ij}$

$=\overline{k}$. the $(i, j)^{th}$ element of $\overline{A}$ = the $(i, j)^{th}$ element of $\overline{k}\,\overline{A}$.

Hence $\overline{(kA)}=\overline{k}\,\overline{A}$.

(iv) Let $A=[a_{ij}]_{m\times n}$, $B=[b_{jk}]_{n\times p}$.

Then $\overline{A}=[\overline{a}_{ij}]_{m\times n}$ and $\overline{B}=[\overline{b}_{jk}]_{n\times p}$.

First we see that both the matrices $\overline{(AB)}$ and $\overline{A}\,\overline{B}$ are of the type $m\times p$.

Again the $(i, k)^{th}$ element of $\overline{(AB)}$

= the conjugate complex of the $(i, k)^{th}$ element of AB

= the conjugate complex of $\sum_{j=1}^{n} a_{ij}\, b_{jk}$

$$=\overline{\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)}=\sum_{j=1}^{n}\overline{a_{ij}b_{jk}}=\sum_{j=1}^{n}\overline{a}_{ij}\,\overline{b}_{jk}$$

= the $(i, k)^{th}$ element of $\overline{A}\,\overline{B}$.

Hence $\overline{(AB)}=\overline{A}\,\overline{B}$.

§ 19. Transposed conjugate of a matrix. Definition.

The transpose of the conjugate of a matrix A is called transposed conjugate of A and is denoted by A^θ or by A^*.

Obviously the conjugate of the transpose of A is the same as the transpose of the conjugate of A *i.e.*,

$\overline{(A')}=(\overline{A})'=A^\theta$.

If $A=[a_{ij}]_{m\times n}$, then

$A^\theta=[b_{ji}]_{n\times m}$ where $b_{ji}=\overline{a}_{ij}$ *i.e.*, the $(j, i)^{th}$ element of A^θ = the conjugate complex of the $(i, j)^{th}$ element of A.

Example. If

$$A=\begin{bmatrix}1+2i & 2-3i & 3+4i\\ 4-5i & 5+6i & 6-7i\\ 8 & 7+8i & 7\end{bmatrix}, \text{ then } A'=\begin{bmatrix}1+2i & 4-5i & 8\\ 2-3i & 5+6i & 7+8i\\ 3+4i & 6-7i & 7\end{bmatrix}$$

and $$\overline{(A')}=A^\theta=\begin{bmatrix}1-2i & 4+5i & 8\\ 2+3i & 5-6i & 7-8i\\ 3-4i & 6+7i & 7\end{bmatrix}.$$

Theorem. *If A^θ and B^θ be the transposed conjugates of A and B respectively, then*

(i) $(A^\theta)^\theta=A$;

(ii) $(A+B)^\theta=A^\theta+B^\theta$, *A and B being of the same size;* **(Meerut 1990)**

(iii) $(kA)^\theta=\overline{k}A^\theta$, *k being any complex number;*

(iv) $(AB)^\theta=B^\theta A^\theta$, *A and B being conformable to multiplication.* **(Delhi 1980)**

Proof. (i) $(A^\theta)^\theta=\left[\overline{\left\{\left(\overline{A}\right)'\right\}'}\right]=\overline{\left(\overline{A}\right)}$, since $\left\{\left(\overline{A}\right)'\right\}'=\overline{A}$

$=A$.

(ii) $(A+B)^\theta=\overline{\{(A+B)'\}}=\overline{(A'+B')}=\overline{(A')}+\overline{(B')}=A^\theta+B^\theta$.

(iii) $(kA)^\theta=\overline{\{(kA)'\}}=\overline{(kA')}=\overline{k}\;\overline{(A')}=\overline{k}A^\theta$.

(iv) $(AB)^\theta=\overline{\{(AB)'\}}=\overline{(B'A')}=\overline{(B')}\;\overline{(A')}=B^\theta A^\theta$.

Thus the reversal law holds for the transposed conjugate also.

§ 20. Symmetric and skew-symmetric matrices.

Symmetric matrix. Definition. *A square matrix* $A=[a_{ij}]$ *is said to be symmetric if its* $(i, j)^{th}$ *element is the same as its* $(j, i)^{th}$ *element i.e., if* $a_{ij}=a_{ji}$ *for all* i, j.

(Rohilkhand 1990; Lucknow 79: Kolhapur 72; Patna 84)

For example,

$$\begin{bmatrix} a & h & g & p \\ h & b & f & q \\ g & f & c & r \\ p & q & r & s \end{bmatrix}, \begin{bmatrix} 1 & i & -2i \\ i & -2 & 4 \\ -2i & 4 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 4 \\ 4 & 3 \end{bmatrix}$$

are symmetric matrices.

Theorem. *A necessary and sufficient condition for a matrix A to be symmetric is that A and A′ are equal.* **(Kanpur 1986)**

Proof. Let $A=[a_{ij}]$ be an n-rowed symmetric matrix.

Then $a_{ij}=a_{ji}$.

A′ will also be an n-rowed square matrix.

Also the $(i, j)^{th}$ element of A' = the $(j, i)^{th}$ element of $A=a_{ji}$
$=a_{ij}$ = the $(i, j)^{th}$ element of A.

Hence $A'=A$.

Conversely if $A'=A$, then A must be a square matrix.

Also the $(i, j)^{th}$ element of A = the $(i, j)^{th}$ element of A′ [$\because$ $A=A'$]
= the $(j, i)^{th}$ element of A.

Hence A is a symmetric matrix.

Skew-symmetric matrix. Definition. *A square matrix* $A=[a_{ij}]$ *is said to be skew-symmetric if the* $(i, j)^{th}$ *element of A is the negative of the* $(j, i)^{th}$ *element of A i.e., if* $a_{ij}=-a_{ji}$ *for all* i, j.

(Bombay 1968, Kolhapur 72, Patna 79, Meerut 89)

If A is a skew-symmetric matrix, then

$a_{ij}=-a_{ji}$ [by definition]

$\therefore$ $a_{ii}=-a_{ii}$, for all values of i

$\therefore$ $2a_{ii}=0$ or $a_{ii}=0$.

Thus the diagonal elements of a skew-symmetric matrix are all zero. **(Kanpur 1986)**

Illustration.

The matrices $\begin{bmatrix} 0 & h & g \\ -h & 0 & f \\ -g & -f & 0 \end{bmatrix}, \begin{bmatrix} 0 & -3i & -4 \\ 3i & 0 & 8 \\ 4 & -8 & 0 \end{bmatrix}$

are skew-symmetric matrices.

Theorem. *A necessary and sufficient condition for a matrix A to be skew-symmetric is that* $A'=-A$.

Proof. Let A be an n-rowed skew-symmetric matrix. Then

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

Now $-A$ and A' are both n-rowed square matrices.

Also the $(i, j)^{th}$ element of A' = the $(j, i)^{th}$ element of A
$=a_{ji}=-a_{ij}$ = the $(i, j)^{th}$ element of $-A$.

Hence $A'=-A$.

Conversely, if $A'=-A$, then A must be a square matrix.

Also the $(i, j)^{th}$ element of A
= the negative of the $(i, j)^{th}$ element of A′ [$\because$ $A=-A'$]
= the negative of the $(j, i)^{th}$ element of A.

Hence A is a skew-symmetric matrix.

§ 21. Hermitian and Skew-Hermitian matrices.

Hermitian matrix. Definition. *A square matrix* $A=[a_{ij}]$ *is said to be Hermitian if the* $(i, j)^{th}$ *element of A is equal to the conjugate complex of the* $(j, i)^{th}$ *element of A i.e., if* $a_{ij}=\bar{a}_{ji}$ *for all i and j.*

(Sagar 1964; Meerut 89; Delhi 81; Allahabad 79)

For example,

$$\begin{bmatrix} a & b+ic \\ b-ic & d \end{bmatrix}, \begin{bmatrix} 1 & 2-3i & 3+4i \\ 2+3i & 0 & 4-5i \\ 3-4i & 4+5i & 2 \end{bmatrix}$$

are Hermitian matrices.

If A is a Hermitian matrix, then

$a_{ii}=\bar{a}_{ii}$, by definition.

$\therefore$ a_{ii} is real for all i. Thus every diagonal element of a Hermitian matrix must be real. **(Allahabad 1979)**

A Hermitian matrix over the field of real numbers is nothing but a real symmetric matrix.

Obviously a necessary and sufficient condition for a matrix A to be Hermitian is that $A=A^{\theta}$.

Skew-Hermitian matrix. Definition. *A square matrix* $A=[a_{ij}]$ *is said to be Skew-Hermitian if the* $(i, j)^{th}$ *element of A is equal to the negative of the conjugate complex of the* $(j, i)^{th}$ *element of A i.e., if* $a_{ij}=-\bar{a}_{ji}$ *for all i and j.* **(Delhi 1981; Rohilkhand 90)**

If A is a skew-Hermitian matrix, then

$a_{ii}=-\bar{a}_{ii}$, by definition.

$\therefore$ $a_{ii}+\bar{a}_{ii}=0$

i.e., a_{ii} must be either a pure imaginary number or must b *zero. Thus the diagonal elements of a skew-Hermitian matrix must be pure imaginary numbers or zero.* **(Delhi 1980)**

Illustration. $\begin{pmatrix} 0 & -2-i \\ 2-i & 0 \end{pmatrix}, \begin{pmatrix} -i & 3+4i \\ -3+4i & 0 \end{pmatrix}$

are skew-Hermitian matrices. A skew-Hermitian matrix over the field of real numbers is nothing but a real skew-symmetric matrix.

Obviously a necessary and sufficient condition for a matrix **A** *to be skew-Hermitian is that* $\mathbf{A}^\theta=-\mathbf{A}$.

Solved Examples

Ex. 1. *If* **A** *is a symmetric (skew-symmetric) matrix, then show that* $k\mathbf{A}$ *is also symmetric (skew-symmetric).*

Solution. (i) Let **A** be a symmetric matrix. Then $\mathbf{A}'=\mathbf{A}$.

We have $(k\mathbf{A})'=k\mathbf{A}'$

$=k\mathbf{A}.$ $[\because \mathbf{A}'=\mathbf{A}]$

Since $(k\mathbf{A})'=k\mathbf{A}$, therefore $k\mathbf{A}$ is a symmetric matrix.

(ii) Let **A** be a skew-symmetric matrix. Then $\mathbf{A}'=-\mathbf{A}$.

We have $(k\mathbf{A})'=k\mathbf{A}'$

$=k(-\mathbf{A})$ $[\because \mathbf{A}'=-\mathbf{A}]$

$=-(k\mathbf{A}).$

Since $(k\mathbf{A})'=-(k\mathbf{A})$, therefore $k\mathbf{A}$ is a skew-symmetric matrix.

Ex. 2. *If* **A** *is a Hermitian matrix, show that* $i\mathbf{A}$ *is skew-Hermitian.* **(Meerut 1982)**

Solution. Let **A** be a Hermitian matrix. Then $\mathbf{A}^\theta=\mathbf{A}$.

We have $(i\,\mathbf{A})^\theta=\bar{i}\,\mathbf{A}^\theta$ $[\because (k\mathbf{A})^\theta=\bar{k}\,\mathbf{A}^\theta]$

$=(-i)\,\mathbf{A}^\theta$ $[\because \bar{i}=-i]$

$=-(i\mathbf{A}^\theta)$

$=-(i\mathbf{A})$ $[\because \mathbf{A}^\theta=\mathbf{A}].$

Since $(i\mathbf{A})^\theta=-(i\mathbf{A})$, therefore $i\mathbf{A}$ is a skew-Hermitian matrix.

Ex. 3. *If* **A** *is a skew-Hermitian matrix, then show that* $i\mathbf{A}$ *is Hermitian.*

Solution. Let **A** be a skew-Hermitian matrix. Then $\mathbf{A}^\theta=-\mathbf{A}$.

We have $(i\mathbf{A})^\theta=\bar{i}\,\mathbf{A}^\theta=(-i)\,\mathbf{A}^\theta=-(i\mathbf{A}^\theta)$

$=-\{i\,(-\mathbf{A})\}$ $[\because \mathbf{A}^\theta=-\mathbf{A}]$

$=-\{-(i\mathbf{A})\}=i\mathbf{A}.$

Since $(i\mathbf{A})^\theta=i\mathbf{A}$, therefore $i\mathbf{A}$ is a Hermitian matrix.

Ex. 4. *If* **A**, **B** *are symmetric (skew-symmetric), then so is also* $\mathbf{A}+\mathbf{B}$.

Solution. (i) Let **A** and **B** be two symmetric matrices of the

Now $(\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}'=\mathbf{A}+\mathbf{B}.$

Since $(\mathbf{A}+\mathbf{B})'=\mathbf{A}+\mathbf{B}$, therefore $\mathbf{A}+\mathbf{B}$ is a symmetric matrix.

(ii) Let **A** and **B** be two skew-symmetric matrices of the same order. Then $\mathbf{A}'=-\mathbf{A}$ and $\mathbf{B}'=-\mathbf{B}$.

Now $(\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}'=(-\mathbf{A})+(-\mathbf{B})=-(\mathbf{A}+\mathbf{B}).$

Since $(\mathbf{A}+\mathbf{B})'=-(\mathbf{A}+\mathbf{B})$, therefore $\mathbf{A}+\mathbf{B}$ is a skew-symmetric matrix.

Ex. 5. *If* **A**, **B** *are Hermitian or skew-Hermitian, then so is also* $\mathbf{A}+\mathbf{B}$.

Solution. (i) Let **A** and **B** be two Hermitian matrices of the same order. Then $\mathbf{A}^\theta=\mathbf{A}$ and $\mathbf{B}^\theta=\mathbf{B}$.

Now $(\mathbf{A}+\mathbf{B})^\theta=\mathbf{A}^\theta+\mathbf{B}^\theta=\mathbf{A}+\mathbf{B}.$

Since $(\mathbf{A}+\mathbf{B})^\theta=\mathbf{A}+\mathbf{B}$, therefore $\mathbf{A}+\mathbf{B}$ is a Hermitian matrix.

(ii) Let **A** and **B** be two skew-Hermitian matrices of the same order. Then $\mathbf{A}^\theta=-\mathbf{A}$ and $\mathbf{B}^\theta=-\mathbf{B}$.

Now $(\mathbf{A}+\mathbf{B})^\theta=\mathbf{A}^\theta+\mathbf{B}^\theta=-\mathbf{A}+(-\mathbf{B})=-(\mathbf{A}+\mathbf{B}).$

Since $(\mathbf{A}+\mathbf{B})^\theta=-(\mathbf{A}+\mathbf{B})$, therefore $\mathbf{A}+\mathbf{B}$ is a skew-Hermitian matrix.

Ex. 6. *If* **A** *and* **B** *are symmetric matrices, then show that* **AB** *is symmetric if and only if* **A** *and* **B** *commute i.e.* $\mathbf{AB}=\mathbf{BA}$. **(I.C.S. 1987)**

Solution. It is given that **A** and **B** are two symmetric matrices. Therefore $\mathbf{A}'=\mathbf{A}$ and $\mathbf{B}'=\mathbf{B}$.

Now suppose that $\mathbf{AB}=\mathbf{BA}$.

Then to prove that **AB** is symmetric.

We have $(\mathbf{AB})'=\mathbf{B}'\mathbf{A}'$

$=\mathbf{BA}$ $[\because \mathbf{A}'=\mathbf{A}, \mathbf{B}'=\mathbf{B}]$

$=\mathbf{AB}$ $[\because \mathbf{AB}=\mathbf{BA}]$

Since $(\mathbf{AB})'=\mathbf{AB}$, therefore **AB** is a symmetric matrix.

Conversely suppose that **AB** is a symmetric matrix. Then to prove that $\mathbf{AB}=\mathbf{BA}$.

We have $\mathbf{AB}=(\mathbf{AB})'$ $[\because$ **AB** is a symmetric matrix$]$

$=\mathbf{B}'\mathbf{A}'=\mathbf{BA}.$

Ex. 7. *If* **A** *be any matrix, then prove that* $\mathbf{AA}'$ *and* $\mathbf{A}'\mathbf{A}$ *are both symmetric matrices.* **(Rohilkhand 1978; Kanpur 87)**

Solution. Let **A** be any matrix.

We have $(\mathbf{AA}')'=(\mathbf{A}')'\,\mathbf{A}'$ [By the reversal law for transposes]

$=\mathbf{AA}'$ $[\because (\mathbf{A}')'=\mathbf{A}].$

Since $(\mathbf{AA}')'=\mathbf{AA}'$, therefore $\mathbf{AA}'$ is a symmetric matrix.

Ex. 8. *If* A *and* B *are two* $n\times n$ *matrices, then show that*
(*i*) $(-A)'=-(A')$, (*ii*) $(-A)^{\theta}=-(A^{\theta})$,
(*iii*) $(A-B)'=A'-B'$ (*iv*) $(A-B)^{\theta}=A^{\theta}-B^{\theta}$.

Solution (i) We have $(-A)'=\{(-1)\ A\}'=(-1)\ A'=-A'$.

(ii) We have $(-A)^{\theta}=\{(-1)\ A\}^{\theta}=\overline{(-1)}\ A^{\theta}$
$$=(-1)\ A^{\theta}\ [\because\ \overline{(-1)}=-1]$$
$$=-A^{\theta}.$$

(iii) We have $(A-B)'=\{A+(-B)\}'=A'+(-B)'$
$$=A'+(-B')=A'-B'.$$

(iv) We have $(A-B)^{\theta}=\{A+(-B)\}^{\theta}=A^{\theta}+(-B)^{\theta}$
$$=A^{\theta}+(-B^{\theta})=A^{\theta}-B^{\theta}.$$

Ex. 9. A *and* B *are Hermitian ; show that* AB+BA *is Hermitian and* AB−BA *is skew-Hermitian.*

Solution. Let A and B be two Hermitian matrices of the same order. Then $A^{\theta}=A$ and $B^{\theta}=B$.

Now $(AB+BA)^{\theta}=(AB)^{\theta}+(BA)^{\theta}=B^{\theta}A^{\theta}+A^{\theta}B^{\theta}$
$$=BA+AB=AB+BA.$$

Hence AB+BA is Hermitian.

Again $(AB-BA)^{\theta}=(AB)^{\theta}-(BA)^{\theta}=B^{\theta}A^{\theta}-A^{\theta}B^{\theta}$
$$=BA-AB=-(AB-BA).$$

Hence AB−BA is skew-Hermitian.

Ex. 10. *If* A *be any square matrix, then show that* A+A′ *is symmetric and* A−A′ *is skew-symmetric.*

(Rohilkhand 1979; Meerut 82)

Solution. We have $(A+A')'=A'+(A')'=A'+A=A+A'$.

Hence A+A′ is symmetric.

Again $(A-A')'=A'-(A')'=A'-A=-(A-A')$.

Hence A−A′ is skew-symmetric.

Ex. 11. *If* A *be any square matrix, prove that* $A+A^{\theta}$, AA^{θ}, $A^{\theta}A$ *are all Hermitian and* $A-A^{\theta}$ *is skew-Hermitian.*

(Bombay 1968; Gujrat 1970)

Solution. The necessary and sufficient condition for a matrix A to be Hermitian is that A^{θ} and A are equal.

Now (i) $(A+A^{\theta})^{\theta}=A^{\theta}+(A^{\theta})^{\theta}=A^{\theta}+A=A+A^{\theta}$.

Hence $A+A^{\theta}$ is Hermitian.

(ii) $(AA^{\theta})^{\theta}=(A^{\theta})^{\theta}\ A^{\theta}$, by the reversal law for conjugate transposes
$$=AA^{\theta}.$$

Hence AA^{θ} is Hermitian

(iii) $(A^{\theta}A)^{\theta}=A^{\theta}\ (A^{\theta})^{\theta}=A^{\theta}A$. Hence $A^{\theta}A$ is Hermitian.

(iv) The necessary and sufficient condition for a matrix A to be skew-Hermitian is that −A and A^{θ} are equal.

Now $(A-A^{\theta})^{\theta}=A^{\theta}-(A^{\theta})^{\theta}=A^{\theta}-A=-(A-A^{\theta})$.

Hence $A-A^{\theta}$ is skew-Hermitian.

***Ex. 12.** *Prove that* $(A_1A_2...A_n)'=(A'_nA'_{n-1}...A'_1)$, *the matrices* $A_1, A_2,..., A_n$ *being of suitable sizes for* $A_1A_2...A_n$ *to exist.*

Solution. We have $(A_1A_2)'=A_2'A_1'$, the reversal law for the transpose of the product of two matrices being already established.

Let us assume that $(A_1A_2A_3...A_{n-1}A_n)'=A'_nA'_{n-1}...A_2'A_1'$. ...(1)

Then $(A_1A_2...A_{n-1}\ A_n\ A_{n+1})'=\{(A_1A_2A_3...A_{n-1}\ A_n)\ A_{n+1}\}'$, since the product of any number of matrices is associative
$=A'_{n+1}\ (A_1A_2...A_{n-1}\ A_n)'$, by the reversal law for the transpose of the product of two matrices
$=A'_{n+1}\ (A_n'\ A_{n-1}'A_{n-2}'...A_2'\ A_1')$, by virtue of (1)
$=A'_{n+1}\ A_n'\ A_{n-1}'...A_2'\ A_1'$, by the associative law for the product of any number of matrices.

Thus the law is true for the product of $n+1$ matrices, if it is true for the product of n matrices. But the law is true for the product of two matrices. Hence it is true for the product of any number of matrices.

Ex. 13. *Prove that* $(A_1A_2...A_n)^{\theta}=A_n^{\theta}A^{\theta}_{n-1}...A_2^{\theta}A_1^{\theta}$, *the matrices* $A_1, A_2,..., A_n$ *being of suitable sizes for* $A_1A_2...A_n$ *to exist.*

Solution. Proceed exactly on the same lines as in Ex. 12.

Ex. 14. *Show that the matrix* B′AB *is symmetric or skew-symmetric according as* A *is symmetric or skew-symmetric.*

(Sagar 1968)

Solution. Case I. Let A be a symmetric matrix. Then A′=A.

Now $(B'AB)'=B'A'\ (B')'$, by the reversal law for the transposes
$$=B'A'B \qquad [\text{since } (B')'=B]$$
$$=B'AB.$$

Hence B′AB is symmetric.

Case II. Let A be a skew-symmetric matrix.

Then $A'=-A$.

Now $(B'AB)'=B'A'\ (B')'=B'A'B=B'\ (-A)\ B$
$$=-(B'A)\ B=-B'AB.$$

Hence B′AB is a skew-symmetric matrix.

Ex. 15. *Show that the matrix* $B^{\theta}AB$ *is Hermitian or skew-Hermitian according as A is Hermitian or skew-Hermitian.*

(Kanpur 1979; Bombay 66)

Solution. Case I. Let A be a Hermitian matrix. Then $A^{\theta}=A$.

Now $(B^{\theta}AB)^{\theta}=B^{\theta}A^{\theta}(B^{\theta})^{\theta}$, by the reversal law for the conjugate transposes

$=B^{\theta}A^{\theta}B=B^{\theta}AB$.

Hence $B^{\theta}AB$ is a Hermitian matrix.

Case II. Let A be a skew-Hermitian matrix. Then $A^{\theta}=-A$.

Now $(B^{\theta}AB)^{\theta}=B^{\theta}A^{\theta}(B^{\theta})^{\theta}=B^{\theta}A^{\theta}B=B^{\theta}(-A)B$

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

Hence $B^{\theta}AB$ is a skew-Hermitian matrix.

Ex. 16. *Show that every square matrix is uniquely expressible as the sum of a symmetric matrix and a skew-symmetric matrix.*

(Meerut 1988; Kolhapur 72; Delhi 80; Bombay 67)

Solution. Let A be any square matrix. We can write

$A=\frac{1}{2}(A+A')+\frac{1}{2}(A-A')=P+Q$, say,

where $P=\frac{1}{2}(A+A')$ and $Q=\frac{1}{2}(A-A')$.

We have $P'=\{\frac{1}{2}(A+A')\}'$

$=\frac{1}{2}(A+A')'$ $[\because (kA)'=kA']$

$=\frac{1}{2}\{A'+(A')'\}$ $[\because (A+B)'=A'+B']$

$=\frac{1}{2}(A'+A)$ $[\because (A')'=A]$

$=\frac{1}{2}(A+A')=P$.

Therefore P is a symmetric matrix.

Again $Q'=\{\frac{1}{2}(A-A')\}'=\frac{1}{2}(A-A')'=\frac{1}{2}\{A'-(A')'\}$

$=\frac{1}{2}(A'-A)=-\frac{1}{2}(A-A')=-Q$.

Therefore Q is a skew-symmetric matrix.

Thus we have expressed the square matrix A as the sum of a symmetric and a skew-symmetric matrix.

To prove that the representation is unique, let $A=R+S$ be another such representation of A, where R is symmetric and S skew-symmetric. Then to prove that $R=P$ and $S=Q$.

We have $A'=(R+S)'=R'+S'$

$=R-S$ $[\because R'=R$ and $S'=-S]$.

$\therefore$ $A+A'=2R$ and $A-A'=2S$.

This gives $R=\frac{1}{2}(A+A')$ and $S=\frac{1}{2}(A-A')$.

Thus $R=P$ and $S=Q$. Therefore the representation is

****Ex. 17.** *Show that every square matrix is uniquely expressible as the sum of a Hermitian matrix and a skew-Hermitian matrix.*

(Rohilkhand 1980, 81; Delhi 81; Madurai 85)

Solution. If A is any square matrix, then $A+A^{\theta}$ is a Hermitian matrix and $A-A^{\theta}$ is a skew-Hermitian matrix. Therefore $\frac{1}{2}(A+A^{\theta})$ is a Hermitian and $\frac{1}{2}(A-A^{\theta})$ is a skew-Hermitian matrix. Now, we have

$$A=\frac{1}{2}(A+A^{\theta})+\frac{1}{2}(A-A^{\theta})=P+Q, \text{ say,}$$

where P is Hermitian and Q skew-Hermitian. Thus every square matrix can be expressed as the sum of a Hermitian matrix and a skew-Hermitian matrix.

Let, now, $A=R+S$ be another such representation of A, where R is Hermitian and S skew-Hermitian.

Then $A^{\theta}=(R+S)^{\theta}=R^{\theta}+S^{\theta}=R-S$.

$\therefore$ $R=\frac{1}{2}(A+A^{\theta})=P$ and $S=\frac{1}{2}(A-A^{\theta})=Q$.

Thus the representation is unique.

Ex. 18. *Show that all positive integral powers of a symmetric matrix are symmetric.*

Solution. Let A be a symmetric matrix of order n. Then $A^m=AAA\ldots\ldots A$ upto m times, where m is a positive integer.

Now $(A^m)'=(AAA\ldots A$ upto m times$)'$

$=(A'A'A'\ldots A'$ upto m times$)$

$=(AAA\ldots A$ upto m times$)$ $[\because A'=A]$

$=A^m$.

Hence A^m is also a symmetric matrix.

Ex. 19. *Show that +ive odd integral powers of a skew-symmetric matrix are skew-symmetric while +ive even integral powers are symmetric.*

Solution. Let A be a skew-symmetric matrix. Then $A'=-A$.

Now let m be a positive integer. We have

$(A^m)'=(AAA\ldots$upto m times$)'=A'A'A'\ldots$upto m times

$=(-A)(-A)\ldots$upto m times$=(-1)^m A^m$

$=-A^m$ or A^m according as m is odd or even.

Thus if m is an odd+ive integer, then $(A^m)'=-A^m$ and so A^m is skew-symmetric. If m is an even +ive integer, then $(A^m)'=A^m$ and so A^m is symmetric.

Ex. 20. *Show that every real symmetric matrix is Hermitian.*

$a_{ij}=a_{ji}$. Since a_{ji} is a real number, therefore $\bar{a}_{ji}=a_{ji}$. Consequently $a_{ij}=\bar{a}_{ji}$. Hence A is Hermitian.

Ex. 21. *Prove that* $\bar{A}$ *is Hermitian or skew-Hermitian according as* A *is Hermitian or skew-Hermitian.*

Solution. Suppose A is Hermitian. Then $A^\theta=A$. We are to prove that $\bar{A}$ is Hermitian. We have

$$(\bar{A})^\theta=\left[\overline{(\bar{A})}\right]' \quad \text{[by def. of conjugate transpose]}$$
$$=(A)' \quad \left[\because \overline{(\bar{A})}=A\right]$$
$$=(A^\theta)' \quad [\because A \text{ is Hermitian} \Rightarrow A=A^\theta]$$
$$=[(\bar{A})']' \quad [\because A^\theta=(\bar{A})']$$
$$=\bar{A} \quad [\because (A')'=A]$$

Since $(\bar{A})^\theta=\bar{A}$, therefore $\bar{A}$ is Hermitian.

Again suppose that A is skew-Hermitian. Then $A^\theta=-A$.

We have $(\bar{A})^\theta=\left[\overline{(\bar{A})}\right]'=(A)'=(-A^\theta)'$

$$=-(A^\theta)'=-\left[(\bar{A})'\right]'=-\bar{A}.$$

Therefore A is also skew-Hermitian.

Ex. 22. *If* $AB=A$ *and* $BA=B$ *then* $B'A'=A'$ *and* $A'B'=B'$ *and hence prove that* A' *and* B' *are idempotent.*

Solution. We have $AB=A \Rightarrow (AB)'=A' \Rightarrow B'A'=A'$.

Also $BA=B \Rightarrow (BA)'=B' \Rightarrow A'B'=B'$.

Now A' is idempotent if $A'^2=A'$. We have

$$A'^2=A'A'=A'(B'A')=(A'B')A'=B'A'=A'.$$

$\therefore$ A' is idempotent.

Again $B'^2=B'B'=B'(A'B')=(B'A')B'=A'B'=B'$.

$\therefore$ B' is idempotent.

Ex. 23. *Show that every square matrix* A *can be uniquely expressed as* $P+iQ$ *where* P *and* Q *are Hermitian matrices.*

(Punjab 1971; Kanpur 79; Rohilkhand 91)

Solution. Let $P=\frac{1}{2}(A+A^\theta)$ and $Q=\frac{1}{2i}(A-A^\theta)$.

Then $$A=P+iQ. \quad \ldots(1)$$

Now $P^\theta=\{\frac{1}{2}(A+A^\theta)\}^\theta=\frac{1}{2}(A+A^\theta)^\theta$

$$=\tfrac{1}{2}\{A^\theta+(A^\theta)^\theta\}=\tfrac{1}{2}(A^\theta+A)=\tfrac{1}{2}(A+A^\theta)=P.$$

$\therefore$ P is a Hermitian matrix.

Also $$Q^\theta=\left\{\frac{1}{2i}(A-A^\theta)\right\}^\theta=\overline{\left(\frac{1}{2i}\right)}(A-A^\theta)^\theta$$
$$=-\frac{1}{2i}\{A^\theta-(A^\theta)^\theta\}=-\frac{1}{2i}(A^\theta-A)$$
$$=\frac{1}{2i}(A-A^\theta)=Q.$$

$\therefore$ Q is also a Hermitian matrix.

Thus A can be expressed in the form (1) where P and Q are Hermitian matrices.

To show that the expression (1) for A is unique.

Let $A=R+iS$ where R and S are both Hermitian matrices.

We have $A^\theta=(R+iS)^\theta=R^\theta+(iS)^\theta=R^\theta+\bar{i}\,S^\theta=R^\theta-iS^\theta$
$$=R-iS \quad [\because \text{R and S are both Hermitian}]$$

$\therefore$ $A+A^\theta=(R+iS)+(R-iS)=2R$.

This gives $R=\frac{1}{2}(A+A^\theta)=P$.

Also $A-A^\theta=(R+iS)-(R-iS)=2iS$.

This gives $S=\frac{1}{2i}(A-A^\theta)=Q$.

Hence expression (1) for A is unique.

Ex. 24. *Prove that every Hermitian matrix* A *can be written as*

$$A=B+iC$$

where B *is real and symmetric and* C *is real and skew-symmetric.*

(Sagar 1964)

Solution. Let A be a Hermitian matrix. Then $A^\theta=A$. Let us take $B=\frac{1}{2}(A+\bar{A})$ and $C=\frac{1}{2i}(A-\bar{A})$.

Then obviously both B and C are real matrices.

[Note that if $z=x+iy$ is a complex number, then $\frac{1}{2}(z+\bar{z})$ is real and also $\frac{1}{2i}(z-\bar{z})$ is real].

Now we can write $A=\frac{1}{2}(A+\bar{A})+i\left[\frac{1}{2i}(A-\bar{A})\right]=B+iC$.

It remains to show that B is symmetric and C is skew-symmetric. We have

$B'=[\frac{1}{2}(A+\bar{A})]'=\frac{1}{2}(A+\bar{A})'=\frac{1}{2}[A'+(\bar{A})']=\frac{1}{2}[A'+A^\theta]$

$=\frac{1}{2}[(A^\theta)'+A] \quad [\because A^\theta=A]$

$$=\tfrac{1}{2}[\{(\bar{A})'\}'+A]=\tfrac{1}{2}(\bar{A}+A)=B.$$

∴ B is symmetric.

$$\text{Also } C'=\left[\frac{1}{2i}(A-\bar{A})\right]'=\frac{1}{2i}(A-\bar{A})'=\frac{1}{2i}[A'-(\bar{A})']$$

$$=\frac{1}{2i}(A'-A^{\theta})=\frac{1}{2i}[(A^{\theta})'-A]=\frac{1}{2i}(\bar{A}-A)$$

$$=-\frac{1}{2i}(A-\bar{A})=-C.$$

∴ C is skew-symmetric.

Hence the result.

Ex. 25. *If* $A=\begin{bmatrix}3 & 1 & -1\\ 0 & 1 & 2\end{bmatrix}$, *find* AA′ *and* A′A.

Solution. We have $A=\begin{bmatrix}3 & 1 & -1\\ 0 & 1 & 2\end{bmatrix}_{2\times 3}$.

$$\therefore\quad A'=\begin{bmatrix}3 & 0\\ 1 & 1\\ -1 & 2\end{bmatrix}_{3\times 2}.$$

$$\text{Now } AA'=\begin{bmatrix}3 & 1 & -1\\ 0 & 1 & 2\end{bmatrix}_{2\times 3}\begin{bmatrix}3 & 0\\ 1 & 1\\ -1 & 2\end{bmatrix}_{3\times 2}$$

$$=\begin{bmatrix}3.3+1.1+1.1 & 3.0+1.1-1.2\\ 0.3+1.1-2.1 & 0.0+1.1+2.2\end{bmatrix}=\begin{bmatrix}11 & -1\\ -1 & 5\end{bmatrix},$$

which is a symmetric matrix

$$\text{Again}\quad A'A=\begin{bmatrix}3 & 0\\ 1 & 1\\ -1 & 2\end{bmatrix}_{3\times 2}\times\begin{bmatrix}3 & 1 & -1\\ 0 & 1 & 2\end{bmatrix}_{2\times 3}$$

$$=\begin{bmatrix}3.3+0.0 & 3.1+0.1 & -3.1+0.2\\ 1.3+1.0 & 1.1+1.1 & -1.1+1.2\\ -1.3+2.0 & -1.1+2.1 & 1.1+2.2\end{bmatrix}$$

$$=\begin{bmatrix}9 & 3 & -3\\ 3 & 2 & 1\\ -3 & 1 & 5\end{bmatrix},\ \text{which is also a symmetric matrix.}$$

Ex. 26. *If* $A=\begin{bmatrix}1 & 0 & 5 & 3\\ -2 & 1 & 6 & 1\\ 3 & 2 & 7 & 1\\ 4 & -4 & -2 & 0\end{bmatrix}$,

find A+A′ *and* A−A′.

Solution. We have

$$A=\begin{bmatrix}1 & 0 & 5 & 3\\ -2 & 1 & 6 & 1\\ 3 & 2 & 7 & 1\\ 4 & -4 & -2 & 0\end{bmatrix}.\quad\therefore\quad A'=\begin{bmatrix}1 & -2 & 3 & 4\\ 0 & 1 & 2 & -4\\ 5 & 6 & 7 & -2\\ 3 & 1 & 1 & 0\end{bmatrix}.$$

Now

$$A+A'=\begin{bmatrix}1 & 0 & 5 & 3\\ -2 & 1 & 6 & 1\\ 3 & 2 & 7 & 1\\ 4 & -4 & -2 & 0\end{bmatrix}+\begin{bmatrix}1 & -2 & 3 & 4\\ 0 & 1 & 2 & -4\\ 5 & 6 & 7 & -2\\ 3 & 1 & 1 & 0\end{bmatrix}$$

$$=\begin{bmatrix}1+1 & 0-2 & 5+3 & 3+3\\ -2+0 & 1+1 & 6+2 & 1-4\\ 3+5 & 2+6 & 7+7 & 1-2\\ 4+3 & -4+1 & -2+1 & 0+0\end{bmatrix}$$

$$=\begin{bmatrix}2 & -2 & 8 & 7\\ -2 & 2 & 8 & -3\\ 8 & 8 & 14 & -1\\ 7 & -3 & -1 & 0\end{bmatrix},$$

which is a symmetric matrix.

Again

$$A-A'=\begin{bmatrix}1 & 0 & 5 & 3\\ -2 & 1 & 6 & 1\\ 3 & 2 & 7 & 1\\ 4 & -4 & -2 & 0\end{bmatrix}-\begin{bmatrix}1 & -2 & 3 & 4\\ 0 & 1 & 2 & -4\\ 5 & 6 & 7 & -2\\ 3 & 1 & 1 & 0\end{bmatrix}$$

$$=\begin{bmatrix}1-1 & 0+2 & 5-3 & 3-4\\ -2-0 & 1-1 & 6-2 & 1+4\\ 3-5 & 2-6 & 7-7 & 1+2\\ 4-3 & -4-1 & -2-1 & 0-0\end{bmatrix}$$

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

which is a symmetric matrix.

Ex. 27. *Give an example of a matrix which is skew-symmetric but not skew-Hermitian.* **(Kanpur 1987)**

Solution. Let $A=\begin{bmatrix}0 & 2+3i\\ -2-3i & 0\end{bmatrix}$.

$$\text{Then } A'=\begin{bmatrix}0 & -2-3i\\ 2+3i & 0\end{bmatrix}=-\begin{bmatrix}0 & 2+3i\\ -2-3i & 0\end{bmatrix}=-A,$$

so that A is skew-symmetric.

Again $A^\theta=\overline{(A')}=\begin{bmatrix}0 & -2+3i\\ 2-3i & 0\end{bmatrix}\neq -A,$

so that A is not skew-Hermitian.

Ex. 28. *If* **U** *and* **V** *are two symmetric matrices, show that* **UVU** *is also symmetric. Is* **UV** *symmetric always? Explain and illustrate by an example.* **(I.A.S. 1970)**

Solution. Since U and V are symmetric matrices, we have $U'=U$ and $V'=V$.

Now $(UVU)'=U'V'U'=UVU$. Hence UVU is also symmetric.

If U and V are symmetric matrices of the same order, then UV is symmetric if and only if $UV=VU$. In case $UV\neq VU$, UV will not be symmetric.

As an illustration consider the symmetric matrices

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

Here $AB=\begin{bmatrix}8 & 11\\ 13 & 18\end{bmatrix}$ and $BA=\begin{bmatrix}8 & 13\\ 11 & 18\end{bmatrix}$,

so that $AB\neq BA$.

Also we observe that $(AB)'\neq AB$ *i.e.*, AB is not symmetric.

Ex. 29. *If* **A** *is Hermitian, such that* $A^2=O$, *show that* $A=O$, *where* **O** *is the zero matrix.* **(Kanpur 1987)**

Solution. Let $A=[a_{ij}]_{n\times n}$ be a Hermitian matrix of order n, so that $A^\theta=A$. We have

$$A=\begin{bmatrix}a_{11} & a_{12}\dots a_{1n}\\ a_{21} & a_{22}\dots a_{2n}\\ \dots & \dots\dots\dots\\ a_{n1} & a_{n2}\dots a_{nn}\end{bmatrix} \text{ and } A^\theta=\begin{bmatrix}\bar{a}_{11} & \bar{a}_{21}\dots\bar{a}_{n1}\\ \bar{a}_{12} & \bar{a}_{22}\dots\bar{a}_{n2}\\ \dots & \dots\dots\dots\\ \bar{a}_{1n} & \bar{a}_{2n}\dots\bar{a}_{nn}\end{bmatrix}.$$

Now it is given that $A^2=O$. $\therefore$ $AA^\theta=O$.

Let $AA^\theta=[b_{ij}]_{n\times n}$.

If $AA^\theta=O$, then each element of AA^θ is zero and so all the principal diagonal elements of AA^θ are zero.

$\therefore$ $b_{ii}=0$ for all $i=1,\dots,n$.

Now $b_{ii}=a_{i1}\bar{a}_{i1}+a_{i2}\bar{a}_{i2}+\dots+a_{in}\bar{a}_{in}$

$=|a_{i1}|^2+|a_{i2}|^2+\dots+|a_{in}|^2.$

$\therefore$ $b_{ii}=0 \Rightarrow |a_{i1}|^2+|a_{i2}|^2+\dots+|a_{in}|^2=0$

$\Rightarrow |a_{i1}|=0, |a_{i2}|=0, \dots, |a_{in}|=0$

$\Rightarrow a_{i1}=0, a_{i2}=0,\dots,a_{in}=0$

$\Rightarrow$ each element of the i^{th} row of A is zero.

But $b_{ii}=0$ for all $i=1,\dots,n$.

$\therefore$ each element of each row of A is zero.

Hence $A=O$.

Exercises

1. Prove that the matrix A^2 is symmetric if either A is symmetric or A is skew-symmetric.
2. If A and B are skew-symmetric matrices of order n, then show that AB is symmetric if and only if A and B commute.

   Hint. Proceed as in Ex. 6.
3. If A is any square matrix, prove that $A+A^T$ is a symmetric matrix. **(Meerut 1982)**
4. If A and B are symmetric matrices of order n, then show that $AB+BA$ is symmetric and $AB-BA$ is skew-symmetric. **(Bombay 1966)**

   Hint. Proceed as in Ex. 9.
5. Prove that every skew-Hermitian matrix A can be written as $B+iC$ where B is real and skew-symmetric and C is real and symmetric.

   Hint. Proceed as in Ex. 24.
6. If $A=\begin{bmatrix}3 & -4\\ 1 & 1\\ 2 & 0\end{bmatrix}$, $B=\begin{bmatrix}2 & 1 & 2\\ 1 & 3 & 4\end{bmatrix}$,

   verify that $(AB)'=B'A'$. **(Kanpur 1979)**

2

Determinants

§ 1. Determinants of order 2.

Definition. Let $a_{11}, a_{12}, a_{21}, a_{22}$ be any four numbers (real or complex). The symbol

$$\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

represents the number $a_{11}a_{22}-a_{21}a_{12}$ and is called a determinant of order 2. The numbers $a_{11}, a_{12}, a_{21}, a_{22}$ are called elements of the determinant and the number $a_{11}a_{22}-a_{21}a_{12}$ is called the value of the determinant. **The value of a determinant of order 2 is equal to the product of the elements along the principal diagonal minus the product of the off-diagonal elements.** Thus

$$\begin{vmatrix} 4 & -7 \\ 5 & -3 \end{vmatrix} = (4)(-3)-(5)(-7) = -12+35 = 23.$$

There are two rows and two columns in a determinant of order 2. In a matrix the number of rows and the number of columns may be different. But in a determinant the number of rows is equal to the number of columns. Also a determinant is not simply a system of numbers. It has got numerical value. A matrix is just a system of numbers. It has got no numerical value. For example, the value of the determinant $\begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix}$ is the number $4\times3-1\times2$, *i.e.*, the number 10 while the matrix $\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$ has got no numerical value.

Also $\begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix} = 10$ and $\begin{vmatrix} 10 & 6 \\ 5 & 4 \end{vmatrix} = 10\times4-5\times6 = 10.$

$$\therefore \quad \begin{vmatrix} 4 & 2 \\ 1 & 3 \end{vmatrix} = \begin{vmatrix} 10 & 6 \\ 5 & 4 \end{vmatrix}.$$

But $\qquad \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \neq \begin{bmatrix} 10 & 6 \\ 5 & 4 \end{bmatrix}.$

§ 2. Determinants of order 3.

The symbol

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

is called a determinant of order 3 and its value is the number

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

This is called the expansion of the determinant along its first row. To obtain this expansion we multiply each element of the first row by that determinant of the second order which is obtained by leaving the row and the column passing through that element. Starting from the first element, the signs of the products are alternately positive and negative.

There are three rows and three columns in a determinant of order 3. It has got 3×3 *i.e.*, 9 elements. We can find the value of a determinant of order 3 by expanding it along any of its rows or along any of its columns. For example, if we expand Δ along the second column, then

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

$$= (-1)^{1+2}a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + (-1)^{2+2}a_{22}\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix}$$

$$+ (-1)^{3+2}a_{32}\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}.$$

Similarly, if we expand Δ along the third row, then

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

If a_{ij} is the element occurring in the i^{th} row and the jth column, then to fix the positive or negative sign before it we should multiply it by $(-1)^{i+j}$.

Example 1. *Find the value of the determinant*

$$\Delta = \begin{vmatrix} 1 & 3 & 4 \\ 2 & -1 & 3 \\ 2 & 1 & 2 \end{vmatrix}.$$

Solution. Expanding Δ along the first row, we get

$$\Delta = 1\begin{vmatrix} -1 & 3 \\ 1 & 2 \end{vmatrix} - 3\begin{vmatrix} 2 & 3 \\ 2 & 2 \end{vmatrix} + 4\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix}$$

$$= 1(-2-3) - 3(4-6) + 4(2+2) = -5+6+16 = 17.$$

Example 2. *Find the value of the determinant*

$$\Delta = \begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 4 \\ 5 & -6 & 7 \end{vmatrix}.$$

Solution. Expanding Δ along the first row, we get

$$\Delta=1\begin{vmatrix}3 & 4\\ -6 & 7\end{vmatrix}-0\begin{vmatrix}2 & 4\\ 5 & 7\end{vmatrix}+0\begin{vmatrix}2 & 3\\ 5 & -6\end{vmatrix}$$

$$=1\,(21+24)-0+0=45.$$

Note. We see that it is easy to find the value of a determinant if most of the elements along any of its rows or columns are equal to zero.

Ex. 3. *Find the value of the determinant*

$$\Delta=\begin{vmatrix}1 & 2 & 7\\ 5 & 0 & 2\\ 3 & -4 & 6\end{vmatrix}.$$

Solution. Expanding Δ along the second row, we get

$$\Delta=-5\begin{vmatrix}2 & 7\\ -4 & 6\end{vmatrix}+0\begin{vmatrix}1 & 7\\ 3 & 6\end{vmatrix}-2\begin{vmatrix}1 & 2\\ 3 & -4\end{vmatrix}$$

$$=-5\,(12+28)+0-2\,(-4-6)=-200+20=-180.$$

Note. The element 5 occurs in the second row and the first column. We have $(-1)^{2+1}=(-1)^3=-1$. Therefore we have fixed negative sign before 5. Then the sign before 0 will be +ive and the sign before 2 will be —ive.

§ 3. Minors and cofactors.

We shall now define what are called minors and cofactors of the elements of a determinant.

Minors. Consider the determinant

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

If we leave the row and the column passing through the element a_{ij}, then the second order determinant thus obtained is called the minor of the element a_{ij} and we shall denote it by M_{ij}. In this way we can get 9 minors corresponding to the 9 elements of Δ.

For example,

the minor of the element $a_{21}=\begin{vmatrix}a_{12} & a_{13}\\ a_{32} & a_{33}\end{vmatrix}=M_{21}$,

the minor of the element $a_{32}=\begin{vmatrix}a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}=M_{32}$,

the minor of the element $a_{11}=\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}=M_{11}$, and so on.

In terms of the notation of minors, if we expand Δ along the first row, then

$$\Delta=(-1)^{1+1}\,a_{11}M_{11}+(-1)^{1+2}\,a_{12}M_{12}+(-1)^{1+3}\,a_{13}M_{13}$$
$$=a_{11}M_{11}-a_{12}M_{12}+a_{13}M_{13}.$$

Similarly, if we expand Δ along the second column, then

$$\Delta=-a_{12}M_{12}+a_{22}M_{22}-a_{32}M_{32}.$$

Thus we can express the determinant as a linear combination of the minors of the elements of any row or any column.

Cofactors. The minor M_{ij} multiplied by $(-1)^{i+j}$ is called the **cofactor** of the element a_{ij}. We shall denote the cofactor of an element by the corresponding capital letter. With this notation, cofactor of $a_{ij}=A_{ij}=(-1)^{i+j}M_{ij}$. For example,

the cofactor of the element $a_{21}=A_{21}=(-1)^{2+1}\,M_{21}$

$$=-\begin{vmatrix}a_{12} & a_{13}\\ a_{32} & a_{33}\end{vmatrix},$$

the cofactor of the element $a_{32}=A_{32}=-\begin{vmatrix}a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}$,

the cofactor of the element $a_{11}=A_{11}=\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix}$, and so on

Thus the cofactor of any element $a_{ij}=(-1)^{i+j}\times$ the determinant obtained by leaving the row and the column passing through that element. **(Meerut 1989)**

In terms of the notation of cofactors, we have

$$\Delta=a_{11}A_{11}+a_{12}A_{12}+a_{13}A_{13},$$

or $$\Delta=a_{21}A_{21}+a_{22}A_{22}+a_{23}A_{23},$$

or $$\Delta=a_{13}A_{13}+a_{23}A_{23}+a_{33}A_{33},\ \text{and so on.}$$

Therefore, in a determinant the sum of the products of the elements of any row or column with the corresponding cofactors is equal to the value of the determinant.

Example. *Write the cofactors of the elements of the determinant*

$$\Delta=\begin{vmatrix}2 & 3 & 2\\ 1 & 4 & -1\\ 5 & 6 & 8\end{vmatrix}.$$

Solution. $A_{11}=$cofactor of the element 2 which is the first element of the first row$=\begin{vmatrix}4 & -1\\ 6 & 8\end{vmatrix}=32+6=38.$

$A_{22}=$cofactor of the element 3 which is the second element of the first row$=-\begin{vmatrix}1 & -1\\ 5 & 8\end{vmatrix}=-(8+5)=-13.$

A_{13}=cofactor of the element 2 which is the third element of the first row $=+\begin{vmatrix}1 & 4\\ 5 & 6\end{vmatrix}=6-20=-14.$

A_{21}=cofactor of the element 1 which is the first element of the second row $=-\begin{vmatrix}3 & 2\\ 6 & 8\end{vmatrix}=-(24-12)=-12.$

A_{22}=cofactor of the element 4 which is the second element of the second row $=+\begin{vmatrix}2 & 2\\ 5 & 8\end{vmatrix}=16-10=6.$

Similarly,

$$A_{23}=-\begin{vmatrix}2 & 3\\ 5 & 6\end{vmatrix}=-(12-15)=3,\quad A_{31}=+\begin{vmatrix}3 & 2\\ 4 & -1\end{vmatrix}=-3-8=-11$$

$$A_{32}=-\begin{vmatrix}2 & 2\\ 1 & -1\end{vmatrix}=-(-2-2)=4,\quad A_{33}=+\begin{vmatrix}2 & 3\\ 1 & 4\end{vmatrix}=8-3=5.$$

We have

$$\Delta=2A_{11}+3A_{12}+2A_{13}=2\,(38)+3\,(-13)+2\,(-14)=9$$
$$\Delta=1A_{21}+4A_{22}+(-1)\,A_{23}=1\,(-12)+4\,(6)+(-1)\,(3)=9$$
$$\Delta=5A_{31}+6A_{32}+8A_{33}=5\,(-11)+6\,(4)+8\,(5)=9.$$

Also it is interesting to note that the sum of the products of the elements of any row and the cofactors of some other row is zero. For example

$$2A_{21}+3A_{22}+2A_{23}=2\,(-12)+3\,(6)+2\,(3)=0$$
$$5A_{11}+6A_{12}+8A_{13}=5(38)+6\,(-13)+8\,(-14)=0 \text{ and so on.}$$

§ 4. Determinants of order 4. The symbol

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

is called a determinant of order 4 and its value is the number

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

This is the expansion of the determinant Δ along its first row.

A determinant of order four has 4 rows and 4 columns. It can be expanded along any of its rows or columns as is the case of a determinant of order 3.

§ 5. Determinants of order n. A determinant of order n has n rows and n columns. It has $n\times n$ elements.

A determinant of order n is a square array of $n\times n$ quantities (numbers or functions) enclosed between vertical bars,

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

The cofactor A_{ij} of the element a_{ij} in Δ is equal to $(-1)^{i+j}$ times the determinant of order $n-1$ obtained from Δ by leaving the row and the column passing through the element a_{ij}. Then we have

$$\Delta=a_{i1}\,A_{i1}+a_{i2}\,A_{i2}+\dots+a_{in}\,A_{in}\ (i=1,2,3,\dots,\text{ or } n),$$

or $\Delta=a_{1j}\,A_{1j}+a_{2j}\,A_{2j}+\dots+a_{nj}\,A_{nj}\ (j=1,2,3,\dots,\text{ or } n).$

§ 6. Determinant of a square Matrix. Definition.

(Allahabad 1964)

Let $A=[a_{ij}]_{n\times n}$ *be a square matrix of order n. Then the number*

$$\begin{vmatrix}a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ \dots & \dots & \dots & \dots\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{vmatrix}$$

is called the determinant of the matrix A *and is denoted by* $|\,A\,|$ *or by Det.* A *or by* $|\,a_{ij}\,|$. Since in a determinant the number of rows is equal to the number of columns, therefore **only square matrices can have determinants.**

Ex. 1. *Find the value of the determinant of the matrix*

$$A=\begin{bmatrix}a & 0 & 0 & 0\\ 0 & b & 0 & 0\\ 0 & 0 & c & 0\\ 0 & 0 & 0 & d\end{bmatrix}$$

Solution. We have $|A|=\begin{vmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{vmatrix}$

$=a\begin{vmatrix} b & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & d \end{vmatrix}$, on expanding the determinant along the first row

$=ab\begin{vmatrix} c & 0 \\ 0 & d \end{vmatrix}$, on expanding the determinant along the first row

$=ab\,(cd-0)=abcd.$

Important. The value of the determinant of a diagonal matrix is equal to the product of the elements lying along its principal diagonal. In particular if I_n be a unit matrix of order n, then

$$|I_n|=1.$$

Thus the value of the determinant of a unit matrix is always equal to 1. Obviously $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}=1.$

Ex. 2. *Find the value of the determinant of the matrix*

$$A=\begin{bmatrix} a & h & g & f \\ 0 & b & c & e \\ 0 & 0 & d & k \\ 0 & 0 & 0 & l \end{bmatrix}.$$

Solution. We have

$$|A|=\begin{vmatrix} a & h & g & f \\ 0 & b & c & e \\ 0 & 0 & d & k \\ 0 & 0 & 0 & l \end{vmatrix}=a\begin{vmatrix} b & c & e \\ 0 & d & k \\ 0 & 0 & l \end{vmatrix},$$

on expanding the determinant along the first column

$=ab\begin{vmatrix} d & k \\ 0 & l \end{vmatrix}$, on expanding the determinant along the first column

$=ab\,(dl-k0)=abdl.$

Important. The value of the determinant of an upper triangular matrix *i.e.*, in which all the elements below the principal diagonal are zero is equal to the product of the elements along the principal diagonal.

We can show similarly that the value of the determinant of a lower triangular matrix *i.e.*, in which all the elements above the principal diagonal are zero is equal to the product of the elements along the principal diagonal.

Ex. 3. *Find the determinant of the matrix*

$$A=\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}.$$

Solution. We have

$$|A|=\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}=a\begin{vmatrix} b & f \\ f & c \end{vmatrix}-h\begin{vmatrix} h & f \\ g & c \end{vmatrix}+g\begin{vmatrix} h & b \\ g & f \end{vmatrix},$$

on expanding the determinant along the first row

$=a\,(bc-f^2)-h\,(hc-fg)+g\,(hf-bg)$
$=abc+2fgh-af^2-bg^2-ch^2.$

Ex. 4. *Find the value of the determinant of the matrix*

$$A=\begin{bmatrix} 4 & 7 & 8 \\ -9 & 0 & 0 \\ 2 & 3 & 4 \end{bmatrix}.$$

Solution. We have

$$|A|=\begin{vmatrix} 4 & 7 & 8 \\ -9 & 0 & 0 \\ 2 & 3 & 4 \end{vmatrix}$$

$=-(-9)\begin{vmatrix} 7 & 8 \\ 3 & 4 \end{vmatrix}$, on expanding the determinant along the second row

$=9\,(28-24)=36.$

§ 7. Properties of Determinants.

Theorem 1. *The value of a determinant does not change when rows and columns are interchanged, that is to say*

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

Proof. Let

$\Delta=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$ be a determinant of order 3.

Expanding Δ along the first row, we get

$$\Delta=a_1(b_2c_3-b_3c_2)-b_1(a_2c_3-a_3c_2)+c_1(a_2b_3-a_3b_2)$$
$$=a_1(b_2c_3-b_3c_2)-a_2(b_1c_3-b_3c_1)+a_3(b_1c_2-b_2c_1),$$
after rearrangement of terms

$$=\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Note. From this theorem we conclude that if any property is true for the rows (columns) of a determinant, then it will be true for its columns (rows).

Corollary. *If* A *be an n-rowed square matrix, then*
$$|\mathbf{A}|=|\mathbf{A}'|.$$

Theorem 2. *If any two rows (or two columns) of a determinant are interchanged, the value of the determinant is multiplied by* -1.

Let us verify this property for a determinant of order 3.

Let $\Delta=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$

Expanding Δ along the first row, we get
$$\Delta=a_1(b_2c_3-b_3c_2)-b_1(a_2c_3-a_3c_2)+c_1\ a_2b_3-a_3b_2)$$
$$=-\{a_3(b_2c_1-b_1c_2)-b_3(a_2c_1-a_1c_2)+c_3(a_2b_1-a_1b_2)\},$$
after rearrangement of terms

$$=-\begin{vmatrix} a_3 & b_3 & c_3 \\ a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \end{vmatrix}$$

$=(-1)$ times the determinant obtained from Δ by interchanging the first and the third rows.

Theorem 3. *If all the elements of one row (or one column) of a determinant are multiplied by the same number k, the value of the new determinant is k times the value of the given determinant.*

(Gorakhpur 1979)

Proof. Let

$$\Delta=\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{in} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$ be a determinant of order n.

Let $A_{i1}, A_{i2},\dots, A_{in}$ be the cofactors of the elements $a_{i1}, a_{i2}\dots, \dots, a_{in}$ of the ith row of Δ.

Then $\Delta=a_{i1}A_{i1}+a_{i2}A_{i2}+\dots+a_{in}A_{in}$.

Now suppose all the elements of the ith row of Δ are multiplied by the same number k. The value of the new determinant is
$$=ka_{i1}A_{i1}+ka_{i2}A_{i2}+\dots+ka_{in}A_{in}=k\ \Delta.$$

Note 1. We have
$$\begin{vmatrix} ka_1 & b_1 & c_1 \\ ka_2 & b_2 & c_2 \\ ka_3 & b_3 & c_3 \end{vmatrix}=k\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$
i.e., if each element of any column (or row) has k as a common factor, then we can bring k outside the symbol of determinant.

Note 2. If all the elements of a row (or a column) of a determinant are zero, the value of the determinant is zero.

Corollary. *If* A *be an n-rowed square matrix, and k be any scalar, then* $|k\mathbf{A}|=k^n|\mathbf{A}|.$

We can easily prove this result by taking k common from each of the n columns of $|k\mathbf{A}|$.

Theorem 4. Important for Proof. *If two rows (or two columns) of a determinant are identical, the value of the determinant is zero; in particular*
$$\begin{vmatrix} a_1 & b_1 & a_1 \\ a_2 & b_2 & a_2 \\ a_3 & b_3 & a_3 \end{vmatrix}=0.$$
(Sagar 1975; Agra 77; Kanpur 79)

Proof. Suppose Δ is a determinant of order n whose ith and jth rows are identical. If we interchange these two identical rows, then obviously there will be no change in the value of Δ. But by theorem 2, the value of Δ is multiplied by -1 if we interchange two rows. Therefore, we get
$$\Delta=-\Delta \text{ or } 2\Delta=0 \text{ or } \Delta=0.$$

Theorem. 5. *In a determinant the sum of the products of the elements of any row (column) with the cofactors of the corresponding elements of any other row (column) is zero.* **(Kerala 1970)**

Proof. Let $\Delta=|a_{ij}|$ be a determinant of order n. Then
$$\Delta=a_{i1}A_{i1}+a_{i2}A_{i2}+\dots+a_{in}A_{in},$$
where A_{ij} is the cofactor of the element a_{ij} in the determinant Δ.

Suppose we replace the elements of the ith row by the corresponding elements of the k^{th} row ($i \neq k$). Then obviously the value of the new determinant is

$$= a_{k1} A_{i1} + a_{k2} A_{i2} + \dots + a_{kn} A_{in}.$$

Since in the new determinant two rows *i.e.*, the i^{th} and the k^{th} are identical, therefore its value is zero.

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

Some Important results to be remembered.

Suppose $\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$ is a determinant of order 3.

Let A_1, B_1, C_1 etc. be the cofactors of the elements a_1, b_1, c_1 etc. in Δ.

Then we have the following results :

$$a_1 A_1 + b_1 B_1 + c_1 C_1 = \Delta,$$
$$a_1 A_2 + b_1 B_2 + c_1 C_2 = 0,$$
$$a_1 A_3 + b_1 B_3 + c_1 C_3 = 0,$$
$$a_2 A_2 + b_2 B_2 + c_2 C_2 = \Delta,$$
$$a_2 A_1 + b_2 B_1 + c_2 C_1 = 0, \text{ etc.}$$

Theorem 6. *If in a determinant each element in any row (or column) consists of the sum of two terms, then the determinant can be expressed as the sum of two determinants of the same order.*

Proof. Let $\Delta = \begin{vmatrix} a_1+\alpha_1 & b_1 & c_1 \\ a_2+\alpha_2 & b_2 & c_2 \\ a_3+\alpha_3 & b_3 & c_3 \end{vmatrix}.$

Expanding Δ along the first column, we get

$$\Delta = (a_1+\alpha_1) \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - (a_2+\alpha_2) \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + (a_3+\alpha_3) \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix}$$

$$= \left\{ a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \right\}$$

$$+ \left\{ \alpha_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - \alpha_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + \alpha_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \right\}$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} \alpha_1 & b_1 & c_1 \\ \alpha_2 & b_2 & c_2 \\ \alpha_3 & b_3 & c_3 \end{vmatrix}$$

Theorem 7. An Important Property. Also Important for Proof. *If to the elements of a row (or column) of a determinant are added m times the corresponding elements of another row (or column),*

the value of the determinant thus obtained is equal to the value of the original determinant, in particular,

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+mb_1 & b_1 & c_1 \\ a_2+mb_2 & b_2 & c_2 \\ a_3+mb_3 & b_3 & c_3 \end{vmatrix}.$$

(Allahabad 1966; Gorakhpur 62)

Proof. We have

$$\begin{vmatrix} a_1+mb_1 & b_1 & c_1 \\ a_2+mb_2 & b_2 & c_2 \\ a_3+mb_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + \begin{vmatrix} mb_1 & b_1 & c_1 \\ mb_2 & b_2 & c_2 \\ mb_3 & b_3 & c_3 \end{vmatrix}, \text{ by theorem 6}$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + m \begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix}, \text{ by theorem 3}$$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \text{ since } \begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} = 0$$

as two of its columns are identical.

Note. We can similarly prove that

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1+mb_1+nc_1 & b_1 & c_1 \\ a_2+mb_2+nc_2 & b_2 & c_2 \\ a_3+mb_3+nc_3 & b_3 & c_3 \end{vmatrix}.$$

Also it should be marked that the numbers may be positive or negative.

Ex. 1. *Let* **A** *be a square matrix of order n. Show that*

(i) $|\mathbf{A}'| = |\mathbf{A}|$. (ii) $|\bar{\mathbf{A}}| = \overline{|\mathbf{A}|}$. (iii) $|\mathbf{A}^{\theta}| = \overline{|\mathbf{A}|}$.

Solution.

(i) The matrix A' is obtained from the matrix A by changing **rows into columns** and columns into rows. But we know that the **value of a determinant** does not change by this change. Therefore $|\mathbf{A}'| = |\mathbf{A}|$.

(ii) Let $\mathbf{A} = [a_{ij}]_{n \times n}$. Then $\bar{\mathbf{A}} = [\bar{a}_{ij}]_{n \times n}$.

We have $|\bar{\mathbf{A}}| = |\bar{a}_{ij}| = \overline{|a_{ij}|} = \overline{|\mathbf{A}|}$.

(iii) We have $A^{\theta}=\overline{(A')}$.

$\therefore \quad |A^{\theta}|=|\overline{(A')}|=\overline{|A'|}=\overline{|A|}$, since $|A'|=|A|$.

Ex. 2. *Show that the determinant of a Hermitian matrix is always a real number.* **(Gujrat 1970)**

Solution. Let A be a Hermitian matrix. Then $A^{\theta}=A$.

$\therefore \quad |A|=|A^{\theta}|=\overline{|A|}$.

Now we know that if z is a complex number such that $z=\bar{z}$, then z is real. Therefore

$$|A|=\overline{|A|} \text{ implies } |A| \text{ is a real number.}$$

§ 8. Working rule for finding the value of a determinant. If the determinant is of order 2, we can at once write its value. But to find the value of a determinant of order $\geqslant 3$, we should always try to make zeros at maximum number of places in any particular row (or column) and then to expand the determinant along that row (or column). The property given in theorem 7 helps us to make zeros.

For convenience we shall denote 1st, 2nd, 3rd rows of a determinant by R_1, R_2, R_3 and columns by C_1, C_2, C_3 etc. If we change the ith row by adding to it m times the corresponding elements of the jth row, then we shall denote this operation by writing $R_i \to R_i+mR_j$ or simply by writing R_i+mR_j. It should be noted that in this operation only R_i will change while R_j will remain as it is.

Solved Examples

Ex. 1. *Show that*

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

(Marathwada 1971; Meerut 79)

Solution. Applying $R_2 \to R_2-R_1$ and $R_3 \to R_3-R_1$, we get

$$\Delta=\begin{vmatrix} 1 & a & a^2 \\ 0 & b-a & b^2-a^2 \\ 0 & c-a & c^2-a^2 \end{vmatrix}$$

$=1\begin{vmatrix} b-a & (b-a)(b+a) \\ c-a & (c-a)(c+a) \end{vmatrix}$, on expanding the determinant along the first column

$=(b-a)(c-a)\begin{vmatrix} 1 & b+a \\ 1 & c+a \end{vmatrix}$, taking $(b-a)$ common from the first row and $(c-a)$ from the second row

$=(b-a)(c-a)\{(c+a)-(b+a)\}$

$=(b-a)(c-a)(c-b)=(a-b)(b-c)(c-a)$.

Ex. 2. *Evaluate*

$$\Delta=\begin{vmatrix} 3 & 2 & 1 & 4 \\ 15 & 29 & 2 & 14 \\ 16 & 19 & 3 & 17 \\ 33 & 39 & 8 & 38 \end{vmatrix}.$$

(Agra 1978)

Solution. Applying $C_1 \to C_1-3C_3$, $C_2 \to C_2-2C_3$, $C_4 \to C_4-4C_3$, we get

$$\Delta=\begin{vmatrix} 0 & 0 & 1 & 0 \\ 9 & 25 & 2 & 6 \\ 7 & 13 & 3 & 5 \\ 9 & 23 & 8 & 6 \end{vmatrix}$$

$=1\begin{vmatrix} 9 & 25 & 6 \\ 7 & 13 & 5 \\ 9 & 23 & 6 \end{vmatrix}$, on expanding the determinant along the first row

$=\begin{vmatrix} 0 & 2 & 0 \\ 7 & 13 & 5 \\ 9 & 23 & 6 \end{vmatrix}$, applying $R_1 \to R_1-R_3$

$=-2\begin{vmatrix} 7 & 5 \\ 9 & 6 \end{vmatrix}$, on expanding the determinant along the first row

$=-2(42-45)=-2(-3)=6$.

Ex. 3. *Evaluate*

$$\Delta=\begin{vmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 2^2 & 3^2 & 4^2 & 5^2 \\ 3^2 & 4^2 & 5^2 & 6^2 \\ 4^2 & 5^2 & 6^2 & 7^2 \end{vmatrix}.$$

(Meerut 1980; Agra 79)

Solution. We have

$$\Delta=\begin{vmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{vmatrix}$$

$=\begin{vmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 5 & 7 & 9 & 11 \\ 7 & 9 & 11 & 13 \end{vmatrix}$ applying $R_4 \to R_4-R_3$, $R_3 \to R_3-R_2$, $R_2 \to R_2-R_1$

$=\begin{vmatrix} 1 & 4 & 9 & 16 \\ 3 & 5 & 7 & 9 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{vmatrix}$ applying $R_4 \to R_4-R_3$, $R_3 \to R_3-R_2$

$=0$, the last two rows being identical.

Ex. 4. *Evaluate* $\Delta=\begin{vmatrix} a-b & m-n & x-y \\ b-c & n-p & y-z \\ c-a & p-m & z-x \end{vmatrix}$.

Solution. Applying $R_1 \to R_1+R_2+R_3$, we get

$$\Delta=\begin{vmatrix} 0 & 0 & 0 \\ b-c & n-p & y-z \\ c-a & p-m & z-x \end{vmatrix}=0.$$

Ex. 5. *Evaluate*

$$\Delta=\begin{vmatrix} 265 & 240 & 219 \\ 240 & 225 & 198 \\ 219 & 198 & 181 \end{vmatrix}.$$ **(Allahabad 1960)**

Solution. Applying $C_1 \to C_1+C_3-2C_2$, we get

$$\Delta=\begin{vmatrix} 4 & 240 & 219 \\ -12 & 225 & 198 \\ 4 & 198 & 181 \end{vmatrix}$$

$$=4\begin{vmatrix} 1 & 240 & 219 \\ -3 & 225 & 198 \\ 1 & 198 & 181 \end{vmatrix}$$, taking 4 common from the first column

$$=4\times 3\begin{vmatrix} 1 & 240 & 219 \\ -1 & 75 & 66 \\ 1 & 198 & 181 \end{vmatrix}$$, taking 3 common from the second row

$$=12\begin{vmatrix} 1 & 240 & 219 \\ 0 & 315 & 285 \\ 0 & -42 & -38 \end{vmatrix}$$ applying $R_2 \to R_2+R_1$, $R_3 \to R_3-R_1$

$$=12\begin{vmatrix} 315 & 285 \\ -42 & -38 \end{vmatrix}$$, expanding the determinant along the first column

$$=12\times 21\times 19\begin{vmatrix} 15 & 15 \\ -2 & -2 \end{vmatrix}$$, taking 21 common from the first column and 19 from the second column

$=0.$

Ex. 6. *Evaluate*

$$\Delta=\begin{vmatrix} 1 & \omega & \omega^2 \\ \omega & \omega^2 & 1 \\ \omega^2 & 1 & \omega \end{vmatrix}$$, *where ω is one of the imaginary cube roots of unity.* **(Allahabad 1967)**

Solution. Applying $C_1 \to C_1+C_2+C_3$, we get

$$\Delta=\begin{vmatrix} 1+\omega+\omega^2 & \omega & \omega^2 \\ 1+\omega+\omega^2 & \omega^2 & 1 \\ 1+\omega+\omega^2 & 1 & \omega \end{vmatrix}$$

$$=\begin{vmatrix} 0 & \omega & \omega^2 \\ 0 & \omega^2 & 1 \\ 0 & 1 & \omega \end{vmatrix} \qquad [\because\ 1+\omega+\omega^2=0]$$

$=0.$

Ex. 7. *Evaluate* $\Delta=\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}$.

Solution. Applying $C_1 \to C_1+C_2+C_3$, we get

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

$$=(a+b+c)\begin{vmatrix} 1 & b & c \\ 1 & c & a \\ 1 & a & b \end{vmatrix}$$, taking $(a+b+c)$ common from the first column

$$=(a+b+c)\begin{vmatrix} 1 & b & c \\ 0 & c-b & a-c \\ 0 & a-b & b-c \end{vmatrix}$$, applying $R_3 \to R_3-R_1$, $R_2 \to R_2-R_1$,

$=(a+b+c)\{(c-b)(b-c)-(a-b)(a-c)\}$

$=(a+b+c)(bc+ca+ab-a^2-b^2-c^2)$

$=-(a+b+c)(a^2+b^2+c^2-bc-ca-ab)$

$=-(a^3+b^3+c^3-3abc).$

Ex. 8. *Prove that*

$$\Delta\equiv\begin{vmatrix} x & a & a & a \\ a & x & a & a \\ a & a & x & a \\ a & a & a & x \end{vmatrix}=(x+3a)(x-a)^3.$$

(Lucknow 1980; Bihar 66)

Solution. Applying $C_1 \to C_1+C_2+C_3+C_4$ and taking $x+3a$ common from the first column, we get

$$\Delta=(x+3a)\begin{vmatrix} 1 & a & a & a \\ 1 & x & a & a \\ 1 & a & x & a \\ 1 & a & a & x \end{vmatrix}$$

$$=(x+3a)\begin{vmatrix} 1 & a & a & a \\ 0 & x-a & 0 & 0 \\ 0 & 0 & x-a & 0 \\ 0 & 0 & 0 & x-a \end{vmatrix}$$ **applying $R_2 \to R_2-R_1$, $R_3 \to R_3-R_1$, $R_4 \to R_4-R_1$**

$$=(x+3a).1\begin{vmatrix} x-a & 0 & 0 \\ 0 & x-a & 0 \\ 0 & 0 & x-a \end{vmatrix}$$, **on expanding the determinant along the first column**

$$=(x+3a)(x-a)\begin{vmatrix} x-a & 0\\ 0 & x-a\end{vmatrix}, \text{ expanding along the first column}$$

$$=(x+3a)(x-a)^3.$$

Ex. 19. *Prove that*

$$\Delta\equiv\begin{vmatrix} 1+a & b & c & d\\ a & 1+b & c & d\\ a & b & 1+c & d\\ a & b & c & 1+d\end{vmatrix}=1+a+b+c+d.$$

(Delhi 1965; Poona 70)

Solution. Applying $C_1\to C_1+C_2+C_3+C_4$ and taking $1+a+b+c+d$ common from the first column, we get

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

$$=(1+a+b+c+d)\begin{vmatrix} 1 & b & c & d\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{vmatrix}, \text{ applying } R_2\to R_2-R_1,\ R_3\to R_3-R_1,\ R_4\to R_4-R_1,$$

$$=(1+a+b+c+d)\begin{vmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{vmatrix}, \text{ expanding along the first column}$$

$$=(1+a+b+c+d).1=1+a+b+c+d.$$

Ex. 10. *Prove that*

$$\Delta\equiv\begin{vmatrix} 1+a & 1 & 1 & 1\\ 1 & 1+b & 1 & 1\\ 1 & 1 & 1+c & 1\\ 1 & 1 & 1 & 1+d\end{vmatrix}=abcd\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right).$$

(Meerut 1989, 91; Gorakhpur 85; Poona 70; Sagar 66)

Solution. Taking a, b, c, d common from the first, second, third and fourth columns respectively, we get

$$\Delta=abcd\begin{vmatrix} \frac{1}{a}+1 & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}\\ \frac{1}{a} & \frac{1}{b}+1 & \frac{1}{c} & \frac{1}{d}\\ \frac{1}{a} & \frac{1}{b} & \frac{1}{c}+1 & \frac{1}{d}\\ \frac{1}{a} & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}+1\end{vmatrix}$$

$$=abcd\begin{vmatrix} 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d} & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}\\ 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d} & \frac{1}{b}+1 & \frac{1}{c} & \frac{1}{d}\\ 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d} & \frac{1}{b} & \frac{1}{c}+1 & \frac{1}{d}\\ 1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d} & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}+1\end{vmatrix}$$

applying $C_1\to C_1+C_2+C_3+C_4$

$$=(abcd)\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)\begin{vmatrix} 1 & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}\\ 1 & \frac{1}{b}+1 & \frac{1}{c} & \frac{1}{d}\\ 1 & \frac{1}{b} & \frac{1}{c}+1 & \frac{1}{d}\\ 1 & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}+1\end{vmatrix}$$

$$=(abcd)\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)\begin{vmatrix} 1 & \frac{1}{b} & \frac{1}{c} & \frac{1}{d}\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{vmatrix},$$

applying $R_2\to R_2-R_1$, $R_3\to R_3-R_1$, $R_4\to R_4-R_1$

$$=(abcd)\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right).$$

Ex. 11. *Prove that*

$$\Delta=\begin{vmatrix} a^2+1 & ab & ac\\ ab & b^2+1 & bc\\ ac & bc & c^2+1\end{vmatrix}=1+a^2+b^2+c^2.$$

Solution. Multiplying the 1st, 2nd and 3rd columns by a, b and c respectively, we get

$$\Delta=\frac{1}{abc}\begin{vmatrix} a(a^2+1) & ab^2 & ac^2\\ a^2b & b(b^2+1) & bc^2\\ a^2c & b^2c & c(c^2+1)\end{vmatrix}$$

$$=\frac{abc}{abc}\begin{vmatrix} a^2+1 & b^2 & c^2\\ a^2 & b^2+1 & c^2\\ a^2 & b^2 & c^2+1\end{vmatrix}, \text{ taking } a, b, c \text{ common from the first, second and third rows respectively.}$$

$$=\begin{vmatrix} 1+a^2+b^2+c^2 & b^2 & c^2 \\ 1+a^2+b^2+c^2 & b^2+1 & c^2 \\ 1+a^2+b^2+c^2 & b^2 & c^2+1 \end{vmatrix}, \text{ applying } C_1\to C_1+C_2+C_3$$

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

$$=(1+a^2+b^2+c^2)\begin{vmatrix} 1 & b^2 & c^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

applying $R_2\to R_2-R_1$, $R_3\to R_3-R_1$

$$=(1+a^2+b^2+c^2).1\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}, \text{ expanding by first column}$$

$$=(1+a^2+b^2+c^2).$$

Ex. 12. *Prove that*

$$\begin{vmatrix} a^2+1 & ab & ac & ad \\ ab & b^2+1 & bc & bd \\ ac & bc & c^2+1 & cd \\ ad & bd & cd & d^2+1 \end{vmatrix}=1+a^2+b^2+c^2+d^2.$$

Solution. Proceed as in Ex. 11.

Ex. 13. *Show that*

$$\Delta\equiv\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \beta+\gamma & \gamma+\delta & \delta+\alpha & \alpha+\beta \\ \delta & \alpha & \beta & \gamma \end{vmatrix}=0.$$

Solution. Applying $R_2\to R_2+R_3+R_4$, we get

$$\Delta=\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha+\beta+\gamma+\delta & \alpha+\beta+\gamma+\delta & \alpha+\beta+\gamma+\delta & \alpha+\beta+\gamma+\delta \\ \beta+\gamma & \gamma+\delta & \delta+\alpha & \alpha+\beta \\ \delta & \alpha & \beta & \gamma \end{vmatrix}$$

$$=(\alpha+\beta+\gamma+\delta)\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ \beta+\gamma & \gamma+\delta & \delta+\alpha & \alpha+\beta \\ \delta & \alpha & \beta & \gamma \end{vmatrix}, \text{ taking } \alpha+\beta+\gamma+\delta \text{ common from the second row}$$

$=(\alpha+\beta+\gamma+\delta).0$, since the first two rows are identical

$=0.$

Ex. 14. *Prove that*

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

(Agra 1980; Rohilkhand 81; Meerut 85; Kanpur 86)

Solution. Let us denote the given determinant by Δ. Applying $R_1\to R_1+R_2+R_3$, we get

$$\Delta=\begin{vmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}$$

$$=(a+b+c)\begin{vmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{vmatrix}, \text{ taking } (a+b+c) \text{ common from the first row}$$

$$=(a+b+c)\begin{vmatrix} 1 & 0 & 0 \\ 2b & -b-c-a & 0 \\ 2c & 0 & -c-a-b \end{vmatrix} \quad \text{applying } C_2\to C_2-C_1,\ C_3\to C_3-C_1$$

$=(a+b+c)^3.$

Ex. 15. *Prove that*

$$\begin{vmatrix} a+b+2c & a & b \\ c & b+c+2a & b \\ c & a & c+a+2b \end{vmatrix}=2\,(a+b+c)^3.$$

(Gorakhpur 1978; Meerut 87; Kanpur 85)

Solution. Let us denote the given determinant by Δ. Applying $C_1\to C_1+C_2+C_3$, we get

$$\Delta=\begin{vmatrix} 2a+2b+2c & a & b \\ 2a+2b+2c & b+c+2a & b \\ 2a+2b+2c & a & c+a+2b \end{vmatrix}$$

$$=2\,(a+b+c)\begin{vmatrix} 1 & a & b \\ 1 & b+c+2a & b \\ 1 & a & c+a+2b \end{vmatrix}, \text{ taking } (2a+2b+2c) \text{ common from the first column}$$

$$=2\,(a+b+c)\begin{vmatrix} 1 & a & b \\ 0 & b+c+a & 0 \\ 0 & 0 & c+a+b \end{vmatrix}, \text{ applying } R_2\to R_2-R_1,\ R_3\to R_3-R_1$$

$$=2\,(a+b+c)\begin{vmatrix} b+c+a & 0 \\ 0 & c+a+b \end{vmatrix}, \text{ expanding with respect to first column}$$

$=2\,(a+b+c)\,[(b+c+a)\,(c+a+b)]=2\,(a+b+c)^3.$

Ex. 16. *Evaluate* $\begin{vmatrix} y+z & x & y \\ z+x & z & x \\ x+y & y & z \end{vmatrix}.$

Solution. Let us denote the given determinant by Δ. Applying $R_1\to R_1+R_2+R_3$, we get

$$\Delta=\begin{vmatrix} 2x+2y+2z & x+y+z & x+y+z \\ z+x & z & x \\ x+y & y & z \end{vmatrix}$$

$$=(x+y+z)\begin{vmatrix} 2 & 1 & 1 \\ z+x & z & x \\ x+y & y & z \end{vmatrix}$$

$$=(x+y+z)\begin{vmatrix} 0 & 1 & 1 \\ 0 & z & x \\ x-z & y & z \end{vmatrix}, \text{ applying } C_1\to C_1-C_2-C_3$$

$$=(x+y+z)(x-z)\begin{vmatrix} 1 & 1 \\ z & x \end{vmatrix}, \text{ expanding with respect to the first column}$$

$$=(x+y+z)(x-z)(x-z)=(x+y+z)(x-z)^2.$$

Ex. 17. *Evaluate* $\begin{vmatrix} b+c & a+b & a \\ c+a & b+c & b \\ a+b & c+a & c \end{vmatrix}$

Solution. Let us denote the given determinant by Δ. Applying $C_1\to C_1+C_3$, we get

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

$$=(a+b+c)\begin{vmatrix} 1 & a+b & a \\ 1 & b+c & b \\ 1 & c+a & c \end{vmatrix}$$

$$=(a+b+c)\begin{vmatrix} 1 & a+b & a \\ 0 & c-a & b-a \\ 0 & c-b & c-a \end{vmatrix}, \text{ applying } R_2\to R_2-R_1,\ R_3\to R_3-R_1$$

$$=(a+b+c)\begin{vmatrix} c-a & b-a \\ c-b & c-a \end{vmatrix}, \text{ expanding with respect to the first column}$$

$$=(a+b+c)[(c-a)^2-(c-b)(b-a)]$$

$$=(a+b+c)[c^2+a^2-2ca-(cb-ca-b^2+ba)]$$

$$=(a+b+c)[a^2+b^2+c^2-bc-ca-ab]$$

$$=a^3+b^3+c^3-3abc.$$

Ex. 18. *Evaluate*

$$\Delta=\begin{vmatrix} 1 & bc & a(b+c) \\ 1 & ca & b(c+a) \\ 1 & ab & c(a+b) \end{vmatrix}.$$

Solution. Applying $C_2\to C_2+C_3$, we get

$$\Delta=\begin{vmatrix} 1 & bc+ca+ab & a(b+c) \\ 1 & bc+ca+ab & b(c+a) \\ 1 & bc+ca+ab & c(a+b) \end{vmatrix}$$

$$=(bc+ca+ab)\begin{vmatrix} 1 & 1 & a(b+c) \\ 1 & 1 & b(c+a) \\ 1 & 1 & c(a+b) \end{vmatrix}$$

$=(bc+ca+ab)\times 0$, since the first two columns are identical

$=0.$

Ex. 19. *Prove that*

$$\Delta\equiv\begin{vmatrix} 1 & bc+ad & b^2c^2+a^2d^2 \\ 1 & ca+bd & c^2a^2+b^2d^2 \\ 1 & ab+cd & a^2b^2+c^2d^2 \end{vmatrix}=(a-b)(a-c)(a-d)(b-c)(b-d)(c-d).$$

Solution. Applying $R_2\to R_2-R_1$, $R_3\to R_3-R_1$, we get

$$\Delta=\begin{vmatrix} 1 & bc+ad & b^2c^2+a^2d^2 \\ 0 & (a-b)(c-d) & (a^2-b^2)(c^2-d^2) \\ 0 & (a-c)(b-d) & (a^2-c^2)(b^2-d^2) \end{vmatrix}$$

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

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

$$=(a-b)(c-d)(a-c)(b-d)[(a+c)(b+d)-(a+b)(c+d)]$$

$$=(a-b)(c-d)(a-c)(b-d)[ab+ad+cb+cd-ac-ad-bc-bd]$$

$$=(a-b)(c-d)(a-c)(b-d)(ab+cd-ac-bd)$$

$$=(a-b)(c-d)(a-c)(b-d)(b-c)(a-d)$$

$$=(a-b)(a-c)(a-d)(b-c)(b-d)(c-d).$$

Ex. 20. *Prove that*

$$\Delta=\begin{vmatrix} a^3 & 3a^2 & 3a & 1 \\ a^2 & a^2+2a & 2a+1 & 1 \\ a & 2a+1 & a+2 & 1 \\ 1 & 3 & 3 & 1 \end{vmatrix}=(a-1)^6.$$

(Kerala 1965; Kolhapur 79; Lucknow 84)

Solution. Applying $C_1\to C_1-C_2+C_3-C_4$, we get

$$\Delta=\begin{vmatrix} a^3-3a^2+3a-1 & 3a^2 & 3a & 1 \\ 0 & a^2+2a & 2a+1 & 1 \\ 0 & 2a+1 & a+2 & 1 \\ 0 & 3 & 3 & 1 \end{vmatrix}$$

$$=(a^3-3a^2+3a-1)\begin{vmatrix} a^2+2a & 2a+1 & 1 \\ 2a+1 & a+2 & 1 \\ 3 & 3 & 1 \end{vmatrix}, \text{ expanding with respect to first column}$$

$$=(a-1)^3\begin{vmatrix} a^2+2a & 2a+1 & 1\\ 1-a^2 & 1-a & 0\\ 2-2a & 1-a & 0\end{vmatrix}, \text{ applying } \begin{matrix} R_3\to R_3-R_2\\ R_2\to R_2-R_1\end{matrix}$$

$$=(a-1)^3\begin{vmatrix} 1-a^2 & 1-a\\ 2(1-a) & 1-a\end{vmatrix}, \text{ expanding with respect to the third column}$$

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

$$=(a-1)^3(a-1)^2(a-1)=(a-1)^6.$$

Ex. 21. *Show that*

$$\Delta\equiv\begin{vmatrix} 4 & 5 & 6 & x\\ 5 & 6 & 7 & y\\ 6 & 7 & 8 & z\\ x & y & z & 0\end{vmatrix}=(x-2y+z)^2.$$

(Sagar 1972)

Solution. Applying $C_1\to C_1+C_3-2C_2$, we get

$$\Delta=\begin{vmatrix} 0 & 5 & 6 & x\\ 0 & 6 & 7 & y\\ 0 & 7 & 8 & z\\ x-2y+z & y & z & 0\end{vmatrix}$$

$$=-(x-2y+z)\begin{vmatrix} 5 & 6 & x\\ 6 & 7 & y\\ 7 & 8 & z\end{vmatrix}, \text{ expanding along the first column}$$

$$=-(x-2y+z)\begin{vmatrix} 0 & 0 & x-2y+z\\ 6 & 7 & y\\ 7 & 8 & z\end{vmatrix}, \text{ applying } R_1\to R_1+R_3-2R_2$$

$$=-(x-2y+z)(x-2y+z)\begin{vmatrix} 6 & 7\\ 7 & 8\end{vmatrix}, \text{ expanding along the first row}$$

$$=-(x-2y+z)(x-2y+z)(48-49)=(x-2y+z)^2.$$

Ex. 22. *Show that*

$$\Delta\equiv\begin{vmatrix} -a^2 & ab & ac\\ ab & -b^2 & bc\\ ac & bc & -c^2\end{vmatrix}=4a^2b^2c^2.$$

(Kanpur 1989)

Solution. Taking a, b, c common from the first, second and third columns respectively, we get

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

$$=a^2b^2c^2\begin{vmatrix} -1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\end{vmatrix} \text{ taking } a, b, c \text{ common from the 1st, 2nd and 3rd rows respectively}$$

$$=a^2b^2c^2\begin{vmatrix} -1 & 1 & 1\\ 0 & 0 & 2\\ 0 & 2 & 0\end{vmatrix} \text{ applying } R_2\to R_2+R_1,\ R_3\to R_3+R_1$$

$$=(a^2b^2c^2)(-1)(-4), \text{ on expanding along the first column}$$

$$=4a^2b^2c^2.$$

Ex. 23. *Prove that*

$$\Delta\equiv\begin{vmatrix} x & y & z\\ x^2 & y^2 & z^2\\ yz & zx & xy\end{vmatrix}=\begin{vmatrix} 1 & 1 & 1\\ x^2 & y^2 & z^2\\ x^3 & y^3 & z^3\end{vmatrix}=(y-z)(z-x)(x-y)(yz+zx+xy).$$

(Meerut 1986, Kanpur 81, Delhi 79)

Solution. Multiplying the first, second and third columns of the determinant on the left hand side by x, y and z respectively, we get

$$\Delta=\frac{1}{xyz}\begin{vmatrix} x^2 & y^2 & z^2\\ x^3 & y^3 & z^3\\ xyz & xyz & xyz\end{vmatrix}=\frac{xyz}{xyz}\begin{vmatrix} x^2 & y^2 & z^2\\ x^3 & y^3 & z^3\\ 1 & 1 & 1\end{vmatrix}$$

$$=-\begin{vmatrix} x^2 & y^2 & z^2\\ 1 & 1 & 1\\ x^3 & y^3 & z^3\end{vmatrix}=\begin{vmatrix} 1 & 1 & 1\\ x^2 & y^2 & z^2\\ x^3 & y^3 & z^3\end{vmatrix}.$$

Applying $C_2\to C_2-C_1$, $C_3\to C_3-C_1$ to the determinant on the right hand side, we get

$$\Delta=\begin{vmatrix} 1 & 0 & 0\\ x^2 & y^2-x^2 & z^2-x^2\\ x^3 & y^3-x^3 & z^3-x^3\end{vmatrix}$$

$$=1\begin{vmatrix} (y-x)(y+x) & (z-x)(z+x)\\ (y-x)(y^2+xy+x^2) & (z-x)(z^2+zx+x^2)\end{vmatrix}$$

$$=(y-x)(z-x)\begin{vmatrix} y+x & z+x\\ y^2+xy+x^2 & z^2+zx+x^2\end{vmatrix},$$

taking $(y-x)$ common from the first column and $z-x$ from the second column

$$=(y-x)(z-x)\begin{vmatrix} y+x & z-y\\ y^2+xy+x^2 & (z^2-y^2)+zx-xy\end{vmatrix},$$

applying $C_2\to C_2-C_1$

$$=(y-x)(z-x)\begin{vmatrix} y+x & z-y \\ y^2+xy+x^2 & (z-y)(x+y+z)\end{vmatrix}$$

$$=(y-x)(z-x)(z-y)\begin{vmatrix} y+x & 1 \\ y^2+xy+x^2 & x+y+z\end{vmatrix},$$

taking $z-y$ common from the second column

$$=(y-x)(z-x)(z-y)\{(y+x)(x+y+z)-(y^2+xy+x^2)\}$$

$$=(x-y)(y-z)(z-x)(xy+yz+zx).$$

Ex. 24. *Prove that*

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

(Agra 1966; Vikram 61)

Solution. Applying $C_2\to C_2-C_1$, we get

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

$$=-\begin{vmatrix} a^2 & b^2+c^2 & bc \\ b^2 & c^2+a^2 & ca \\ c^2 & a^2+b^2 & ab \end{vmatrix}$$ applying $C_2\to C_2+2C_3$

$$=-\begin{vmatrix} a^2 & a^2+b^2+c^2 & bc \\ b^2 & a^2+b^2+c^2 & ca \\ c^2 & a^2+b^2+c^2 & ab \end{vmatrix}$$ applying $C_2\to C_2+C_1$

$$=-(a^2+b^2+c^2)\begin{vmatrix} a^2 & 1 & bc \\ b^2 & 1 & ca \\ c^2 & 1 & ab \end{vmatrix}$$ taking $a^2+b^2+c^2$ common from the second column

$$=(a^2+b^2+c^2)\begin{vmatrix} 1 & a^2 & bc \\ 1 & b^2 & ca \\ 1 & c^2 & ab \end{vmatrix}$$ interchanging first and second columns

$$=(a^2+b^2+c^2)\begin{vmatrix} 1 & a^2 & bc \\ 0 & b^2-a^2 & ca-bc \\ 0 & c^2-a^2 & ab-bc \end{vmatrix}$$ by $R_2\to R_2-R_1$, $R_3\to R_3-R_1$

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

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

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

by $C_2\to C_2-C_1$

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

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

Ex. 25. *Prove that*

$$\Delta\equiv\begin{vmatrix} b+c & c+a & a+b \\ q+r & r+p & p+q \\ y+z & z+x & x+y \end{vmatrix}=2\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$$

(Meerut 1990)

Solution. Applying $C_1\to C_1+C_2-2C_3$, we get

$$\Delta=\begin{vmatrix} 2c & c+a & a+b \\ 2r & r+p & p+q \\ 2z & z+x & x+y \end{vmatrix}=2\begin{vmatrix} c & c+a & a+b \\ r & r+p & p+q \\ z & z+x & x+y \end{vmatrix}$$

$$=2\begin{vmatrix} c & a & a+b \\ r & p & p+q \\ z & x & x+y \end{vmatrix}$$ applying $C_2\to C_2-C_1$

$$=2\begin{vmatrix} c & a & b \\ r & p & q \\ z & x & y \end{vmatrix}$$ by $C_3\to C_3-C_2$

$$=-2\begin{vmatrix} a & c & b \\ p & r & q \\ x & z & y \end{vmatrix}=2\begin{vmatrix} a & b & c \\ p & q & r \\ x & y & z \end{vmatrix}$$

Ex. 26. *If x, y, z are all different and if*

$$\begin{vmatrix} x & x^2 & 1+x^3 \\ y & y^2 & 1+y^3 \\ z & z^2 & 1+z^3 \end{vmatrix}=0,$$ *prove that $xyz=-1$.*

(Meerut 1989)

Solution. We have

$$\begin{vmatrix} x & x^2 & 1+x^3 \\ y & y^2 & 1+y^3 \\ z & z^2 & 1+z^3 \end{vmatrix}=\begin{vmatrix} x & x^2 & 1 \\ y & y^2 & 1 \\ z & z^2 & 1 \end{vmatrix}+\begin{vmatrix} x & x^2 & x^3 \\ y & y^2 & y^3 \\ z & z^2 & z^3 \end{vmatrix}$$

$$=\begin{vmatrix} x & x^2 & 1 \\ y & y^2 & 1 \\ z & z^2 & 1 \end{vmatrix}+xyz\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix},$$

taking x, y, z common from R_1, R_2 and R_3 of the second determinant

$$=\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}+xyz\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$$

$$=(1+xyz)\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix}$$

$=(1+xyz)(x-y)(y-z)(z-x)$. [See Ex. 1]

Since x, y, z are all different, therefore $x-y\neq 0$, $y-z\neq 0$, $z-x\neq 0$.

Hence $(1+xyz)(x-y)(y-z)(z-x)=0$ implies $1+xyz=0$ *i.e.*, $xyz=-1$.

Ex. 27. *Prove that*

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

(Gorakhpur 1979)

Solution. We have

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

Now $\Delta_2=\dfrac{1}{abcd}\begin{vmatrix} a & a^2 & a^3 & abcd \\ b & b^2 & b^3 & abcd \\ c & c^2 & c^3 & abcd \\ d & d^2 & d^3 & abcd \end{vmatrix}$, multiplying R_1, R_2, R_3, R_4 of Δ_2 by a, b, c, d

$$=\frac{abcd}{abcd}\begin{vmatrix} a & a^2 & a^3 & 1 \\ b & b^2 & b^3 & 1 \\ c & c^2 & c^3 & 1 \\ d & d^2 & d^3 & 1 \end{vmatrix}$$

$$=-\begin{vmatrix} 1 & a & a^2 & a^3 \\ 1 & b & b^2 & b^3 \\ 1 & c & c^2 & c^3 \\ 1 & d & d^2 & d^3 \end{vmatrix}=-\Delta_1.$$

$\therefore\quad \Delta=\Delta_1+(-\Delta_1)=0.$

Ex. 28. *Show that the value of the determinant of a skew-symmetric matrix of odd order is always zero.*

(Nagarjuna 1981, Kanpur 86)

Solution. Let

$$\mathbf{A}=\begin{bmatrix} 0 & -h & -g \\ h & 0 & -f \\ g & f & 0 \end{bmatrix}$$

be a skew-symmetric matrix of order 3.

We have $|\mathbf{A}|=\begin{vmatrix} 0 & -h & -g \\ h & 0 & -f \\ g & f & 0 \end{vmatrix}$.

Taking -1 common from each of the three rows of $|\mathbf{A}|$, we get

$$|\mathbf{A}|=(-1)^3\begin{vmatrix} 0 & h & g \\ -h & 0 & f \\ -g & -f & 0 \end{vmatrix}$$

$=-\begin{vmatrix} 0 & -h & -g \\ h & 0 & -f \\ g & f & 0 \end{vmatrix}$, interchanging the rows and columns

$=-|\mathbf{A}|$.

$\therefore\quad 2|\mathbf{A}|=0$ *i.e.* $|\mathbf{A}|=0$.

Ex. 29. *Show that*

$$\Delta\equiv\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix}=(b-c)(c-a)(a-b)(a+b+c).$$

(Gorakhpur 1981, Meerut 75)

Solution. The students can solve the question easily by applying $C_2\to C_2-C_1$, $C_3\to C_3-C_1$. But we shall give here an alternative method which is also interesting.

The way in which the elements occur in Δ shows that the value of Δ will be a symmetrical expression in a, b, c. Also the leading diagonal term $1\,b.c^3$ (the term obtained by multiplying the elements along the principal diagonal) shows that each term in the expression for Δ will be of degree 4 in a, b, c.

If we put $a=b$ in Δ, we see that the first two columns become identical. Therefore Δ becomes zero and thus $(a-b)$ must be a factor of Δ. Similarly $b-c$ and $c-a$ must also be factors of Δ. Thus $(a-b)(b-c)(c-a)$ is a factor of Δ of degree 3. But Δ is of degree 4 in a, b, c. Therefore Δ must have one more factor of first degree and it should be symmetrical in a, b, c. Let it be $k(a+b+c)$, where k is a constant. We have then the identity

$$\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix}\equiv k(a-b)(b-c)(c-a)(a+b+c).$$

Putting $a=0$, $b=1$, $c=2$ in this identity, we have

$$\begin{vmatrix}1&1&1\\0&1&2\\0&1&8\end{vmatrix}=k\,(-1)\,(-1)\,(2)\,(3)$$

or $\qquad 6k=6 \qquad$ or $\qquad k=1.$

$\therefore \quad \Delta=(a-b)\,(b-c)\,(c-a)\,(a+b+c).$

Ex. 30. *Show that*

$$\Delta\equiv\begin{vmatrix}a&b&c&d\\b&a&d&c\\c&d&a&b\\d&c&b&a\end{vmatrix}=(a+b+c+d)\,(a+d-b-c)\times(a+c-b-d)\,(a+b-c-d).$$

(Nagpur 1980)

Solution. Applying $C_1\to C_1+C_2+C_3+C_4$ and taking $a+b+c+d$ common from the first column, we get

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

$$=(a+b+c+d)\begin{vmatrix}1&b&c&d\\0&a-b&d-c&c-d\\0&d-b&a-c&b-d\\0&c-b&b-c&a-d\end{vmatrix}\text{ by } R_2\to R_2-R_1,\ R_3\to R_3-R_1,\ R_4\to R_4-R_1$$

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

$$=(a+b+c+d)\begin{vmatrix}a+c-b-d&0&c-d\\0&a+b-c-d&b-d\\a+c-b-d&a+b-c-d&a-d\end{vmatrix}$$

by $C_1\to C_1+C_3,\ C_2\to C_2+C_3$

$$=(a+b+c+d)\,(a+c-b-d)\,(a+b-c-d)\begin{vmatrix}1&0&c-d\\0&1&b-d\\1&1&a-d\end{vmatrix}$$

taking $a+c-b-d$ common from the first column and $a+b-c-d$ common from the second column

$$=(a+b+c+d)\,(a+c-b-d)\times(a+b-c-d)\begin{vmatrix}1&0&c-d\\0&1&b-d\\0&1&a-c\end{vmatrix}\text{ by } R_3\to R_3-R_1$$

$$=(a+b+c+d)\,(a+c-b-d)\,(a+b-c-d)\,(a+d-b-c).$$

Ex. 31. *Prove that*

$$\begin{vmatrix}0&x&y&z\\x&0&z&y\\y&z&0&x\\z&y&x&0\end{vmatrix}=-(x+y+z)\,(x+y-z)\,(-x+y+z)\times(x-y+z).$$

(Kerala 1970)

Hint. Proceed as in Ex. 30.

Ex. 32. *Prove that*

$$\Delta\equiv\begin{vmatrix}(b+c)^2&a^2&a^2\\b^2&(c+a)^2&b^2\\c^2&c^2&(a+b)^2\end{vmatrix}=2abc\,(a+b+c)^3.$$

(Rohilkhand 80; Agra 65; Meerut 89; Delhi 81; Poona 70)

Solution. Applying $C_1\to C_1-C_3,\ C_2\to C_2-C_3$, we get

$$\Delta=\begin{vmatrix}(b+c+a)\,(b+c-a)&0&a^2\\0&(c+a+b)\,(c+a-b)&b^2\\(c+a+b)\,(c-a-b)&(c+a+b)\,(c-a-b)&(a+b)^2\end{vmatrix}$$

$$=(a+b+c)^2\begin{vmatrix}b+c-a&0&a^2\\0&c+a-b&b^2\\c-a-b&c-a-b&(a+b)^2\end{vmatrix}$$

taking $a+b+c$ common from each of the columns C_1 and C_2

$$=(a+b+c)^2\begin{vmatrix}b+c-a&0&a^2\\0&c+a-b&b^2\\-2b&-2a&2ab\end{vmatrix}\text{ by } R_3\to R_3-R_2-R_1$$

$$=(a+b+c)^2\begin{vmatrix}b+c&\frac{a^2}{b}&a^2\\\frac{b^2}{a}&c+a&b^2\\0&0&2ab\end{vmatrix}\text{ by } C_1\to C_1+\frac{1}{a}C_3,\ C_2\to C_2+\frac{1}{b}C_3$$

$$=(a+b+c)^2\,2ab\begin{vmatrix}b+c&\frac{a^2}{b}\\\frac{b^2}{a}&c+a\end{vmatrix}$$

$$=(a+b+c)^2\,2ab\,\{(b+c)\,(c+a)-ab\}=2abc\,(a+b+c)^3.$$

Ex. 33. *Prove that*

$$\begin{vmatrix}1+a^2-b^2&2ab&-2b\\2ab&1-a^2+b^2&2a\\2b&-2a&1-a^2-b^2\end{vmatrix}=(1+a^2+b^2)^3.$$

(Meerut 1985)

Solution. Let us denote the given determinant by Δ. Applying $C_1 \to C_1 - b\,C_3$, we get

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

$$= \begin{vmatrix} 1+a^2+b^2 & 2ab & -2b \\ 0 & 1-a^2+b^2 & 2a \\ 0 & -2a\,(1+b^2) & 1-a^2+b^2 \end{vmatrix} \text{ by } R_3 \to R_3 - bR_1$$

$$=(1+a^2+b^2)\,[(1-a^2+b^2)^2+4a^2\,(1+b^2)]$$

$$=(1+a^2+b^2)\,[\{(1+b^2)-a^2\}^2+4a^2\,(1+b^2)]$$

$$=(1+a^2+b^2)\,[(1+b^2)^2-2a^2\,(1+b^2)+a^4+4a^2\,(1+b^2)]$$

$$=(1+a^2+b^2)\,[(1+b^2)^2+2a^2\,(1+b^2)+a^4]$$

$$=(1+a^2+b^2)\,[(1+b^2)+a^2]^2=(1+a^2+b^2)^3.$$

Ex. 34. *Prove that*

$$\begin{vmatrix} a^2 & bc & ac+c^2 \\ a+b & b^2 & ac \\ ab & b^2+bc & c^2 \end{vmatrix} = 4a^2b^2c^2.$$

Solution. Taking a, b, c common from the first, second and third columns respectively, the given determinant

$$\Delta = abc \begin{vmatrix} a & c & a+c \\ a+b & b & a \\ b & b+c & c \end{vmatrix}$$

$$= abc \begin{vmatrix} 0 & 2c & 2c \\ a+b & b & a \\ b & b+c & c \end{vmatrix}, \text{ applying } R_1 \to R_1+R_3-R_2$$

$$= abc \begin{vmatrix} 0 & 0 & 2c \\ a+b & b-a & a \\ b & b & c \end{vmatrix}, \text{ applying } C_2 \to C_2-C_3$$

$= abc \,.\, 2c\,[b\,(a+b)-b\,(b-a)]$, expanding the determinant along the first row

$= abc\,.\,2c\,.\,2ab = 4a^2b^2c^2$.

Ex. 35. *Prove that*

$$\begin{vmatrix} a & c & a+c \\ a+b & b & a \\ b & b+c & c \end{vmatrix} = 4abc.$$

Solution. Proceed as in Ex. 34 above.

Ex. 36. *Prove that*

$$\begin{vmatrix} 0 & -c & b & -l \\ c & 0 & -a & -m \\ -b & a & 0 & -n \\ x & y & z & 0 \end{vmatrix} = (al+bm+cn)\,(ax+by+cz).$$

(Meerut 1973, 81, 83S; Gorakhpur 84; Kanpur 79)

Solution. If a, b, c are all zero, the result is obviously true. So let one of a, b, c be not zero. Without loss of generality we can take $a \neq 0$. Then

$$\begin{vmatrix} 0 & -c & b & -l \\ c & 0 & -a & -m \\ -b & a & 0 & -n \\ x & y & z & 0 \end{vmatrix}$$

$$= \frac{1}{a}\begin{vmatrix} 0 & -ac & ab & -al \\ c & 0 & -a & -m \\ -b & a & 0 & -n \\ x & y & z & 0 \end{vmatrix},$$ multiplying the first row by a and also multiplying outside the determinant by $1/a$

$$= \frac{1}{a}\begin{vmatrix} 0 & 0 & 0 & -(al+bm+cn) \\ c & 0 & -a & -m \\ -b & a & 0 & -n \\ x & y & z & 0 \end{vmatrix},$$ applying $R_1 \to R_1 + bR_2 + cR_3$

$$= \frac{1}{a}\,(al+bm+cn)\begin{vmatrix} c & 0 & -a \\ -b & a & 0 \\ x & y & z \end{vmatrix},$$ expanding the determinant along the first row.

$= \frac{1}{a}\,(al+bm+cn)\,[c\,(az-0)-a\,(-by-ax)]$, expanding the determinant along the first row

$$= \frac{1}{a}\,(al+bm+cn)\,(acz+aby+a^2x)$$

$$= \frac{1}{a}\,(al+bm+cn).a\,(cz+by+ax) = (al+bm+cn)\,(ax+by+cz).$$

Ex. 37. *Prove that*

$$\begin{vmatrix} 0 & x & y & z \\ -x & 0 & r & q \\ -y & -r & 0 & p \\ -z & -q & -p & 0 \end{vmatrix} = (px-qy+rz)^2.$$

(Meerut 1981)

Solution. If p, q, r are all zero, the result is obviously true. So let one of p, q, r be not zero. Without loss of generality we can take $p \neq 0$. Then

$$\begin{vmatrix} 0 & x & y & z \\ -x & 0 & r & q \\ -y & -r & 0 & p \\ -z & -q & -p & 0 \end{vmatrix}$$

$$=\frac{1}{p}\begin{vmatrix} 0 & px & y & z \\ -x & 0 & r & q \\ -y & -pr & 0 & p \\ -z & -pq & -p & 0 \end{vmatrix}$$, multiplying the second column by p and also multiplying outside the determinant by $1/p$

$$=\frac{1}{p}\begin{vmatrix} 0 & px-qy+rz & y & z \\ -x & 0 & r & q \\ -y & 0 & 0 & p \\ -z & 0 & -p & 0 \end{vmatrix}$$, applying $C_2 \to C_2 - qC_3 + rC_4$

$$=-\frac{1}{p}(px-qy+rz)\begin{vmatrix} -x & r & q \\ -y & 0 & p \\ -z & -p & 0 \end{vmatrix}$$ expanding the determinant along the second column

$$=-\frac{1}{p}(px-qy+rz)\,[q\,(yp-0)-p\,(px+rz)],$$ expanding the determinant along the third column

$$=-\frac{1}{p}(px-qy+rz)\,(pqy-p^2x-prz)$$

$$=\left(-\frac{1}{p}\right)(px-qy+rz)\,(-p)\,(px-qy+rz)$$

$$=(px-qy+rz)\,(px-qy+rz)=(px-qy+rz)^2.$$

Ex. 38. *Show that*

$$\begin{vmatrix} -1 & 0 & 0 & a \\ 0 & -1 & 0 & b \\ 0 & 0 & -1 & c \\ x & y & z & -1 \end{vmatrix}=1-ax-by-cz.$$

(Lucknow 1981; Poona 72)

Solution. Applying $C_4 \to C_4 + aC_1 + bC_2 + cC_3$, the given determinant

$$\Delta=\begin{vmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ x & y & z & -1+ax+by+cz \end{vmatrix}$$

$$=(-1+ax+by+cz)\begin{vmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix}$$, expanding the determinant along the fourth column

$$=(-1+ax+by+cz).(-1)=1-ax-by-cz.$$

Ex. 39. *Solve the equation*

$$\begin{vmatrix} 15-x & 11 & 10 \\ 11-3x & 17 & 16 \\ 7-x & 14 & 13 \end{vmatrix}=0.$$

Solution. Applying $C_2 \to C_2 - C_3$, we get

$$\begin{vmatrix} 15-x & 1 & 10 \\ 11-3x & 1 & 16 \\ 7-x & 1 & 13 \end{vmatrix}=0.$$

Applying $R_2 \to R_2 - R_1$, $R_3 \to R_3 - R_1$, we get

$$\begin{vmatrix} 15-x & 1 & 10 \\ -4-2x & 0 & 6 \\ -8 & 0 & 3 \end{vmatrix}=0$$

or $$-1\begin{vmatrix} -4-2x & 6 \\ -8 & 3 \end{vmatrix}=0,$$

expanding along the second column

or $-12-6x+48=0$ or $x=6.$

Ex. 40. *If $a+b+c=0$, solve the equation*

$$\begin{vmatrix} a-x & c & b \\ c & b-x & a \\ b & a & c-x \end{vmatrix}=0.$$ **(Kanpur 1989)**

Solution. Applying $C_1 \to C_1 + C_2 + C_3$, we get

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

or $$(a+b+c-x)\begin{vmatrix} 1 & c & b \\ 1 & b-x & a \\ 1 & a & c-x \end{vmatrix}=0$$

or $$-x\begin{vmatrix} 1 & c & b \\ 0 & b-c-x & a-b \\ 0 & a-c & c-b-x \end{vmatrix}=0$$ by $R_2 \to R_2 - R_1$, $R_3 \to R_3 - R_1$

or $x\,[(b-c-x)\,(c-b-x)-(a-c)\,(a-b)]=0$

or $x\,(x^2-b^2-c^2+2bc-a^2+ab+ca-bc)=0$

or $x\,(x^2-a^2-b^2-c^2+ab+bc+ca)=0.$

$\therefore\ x=0$ or $x^2=a^2+b^2+c^2-(ab+bc+ca)$

$=(a^2+b^2+c^2)-\frac{1}{2}\{(a+b+c)^2-(a^2+b^2+c^2)\}$

$=\frac{3}{2}(a^2+b^2+c^2).$

$\therefore\ x=0$ or $x=\pm\sqrt{\{\frac{3}{2}(a^2+b^2+c^2)\}}.$

Ex. 41. *Solve the equation*

$$\begin{vmatrix} 3x-8 & 3 & 3 \\ 3 & 3x-8 & 3 \\ 3 & 3 & 3x-8 \end{vmatrix}=0.$$

(Meerut 1974; Delhi 81)

Solution. Applying $C_1 \to C_1 + C_2 + C_3$ and taking $3x-2$ common from C_1, the given equation becomes

$$(3x-2)\begin{vmatrix} 1 & 3 & 3 \\ 1 & 3x-8 & 3 \\ 1 & 3 & 3x-8 \end{vmatrix}=0.$$

Applying $R_2 \to R_2 - R_1$, $R_3 \to R_3 - R_1$, the above equation becomes

$$(3x-2)\begin{vmatrix} 1 & 3 & 3 \\ 0 & 3x-11 & 0 \\ 0 & 0 & 3x-11 \end{vmatrix}=0$$

or $(3x-2)(3x-11)^2=0$.

$\therefore\ x=2/3$ or $x=11/3,\ 11/3$.

Ex. 42. *Solve the equation*

$$\begin{vmatrix} x-2 & 2x-3 & 3x-4 \\ x-4 & 2x-9 & 3x-16 \\ x-8 & 2x-27 & 3x-64 \end{vmatrix}=0.$$

Solution. Applying $R_2 \to R_2 - R_1$, $R_3 \to R_3 - R_1$, the given equation becomes

$$\begin{vmatrix} x-2 & 2x-3 & 3x-4 \\ -2 & -6 & -12 \\ -6 & -24 & -60 \end{vmatrix}=0$$

or $$\begin{vmatrix} x-2 & 2x-3 & 3x-4 \\ 1 & 3 & 6 \\ 1 & 4 & 10 \end{vmatrix}=0.$$

Expanding the determinant along the first row, the above equation becomes

$(x-2).6-(2x-3).4+(3x-4).1=0$

or $6x-8x+3x-12+12-4=0$ or $x-4=0$.

$\therefore\ x=4.$

Ex. 43. *Show that* $x=2$ *is a root of the equation*

$$\begin{vmatrix} x & -6 & -1 \\ 2 & -3x & x-3 \\ -3 & 2x & x+2 \end{vmatrix}=0$$

and solve it completely. (Meerut 1987)

Solution. Let $\triangle$ denote the determinant on the L.H.S. of the given equation. Putting $x=2$ in $\triangle$, we have

$$\triangle=\begin{vmatrix} 2 & -6 & -1 \\ 2 & -6 & -1 \\ -3 & 4 & 4 \end{vmatrix}=0,$$ the first two rows being identical.

Thus $\triangle$ vanishes for $x=2$. Hence $x=2$ is a root of the equation $\triangle=0$.

Applying $R_1 \to R_1 - R_2$, and then taking $x-2$ common from R_1, the given equation becomes

$$(x-2)\begin{vmatrix} 1 & 3 & -1 \\ 2 & -3x & x-3 \\ -3 & 2x & x+2 \end{vmatrix}=0.$$

Applying $C_2 \to C_2 - 3C_1$ and $C_3 \to C_3 + C_1$, the above equation becomes

$$(x-2)\begin{vmatrix} 1 & 0 & 0 \\ 2 & -3x-6 & x-1 \\ -3 & 2x+9 & x-1 \end{vmatrix}=0$$

or $$(x-2)(x-1)\begin{vmatrix} -3x-6 & 1 \\ 2x+9 & 1 \end{vmatrix}=0$$

or $(x-2)(x-1)(-3x-6-2x-9)=0$

or $(x-2)(x-1)(-5x-15)=0$.

$\therefore\ x=1,\ 2,\ -3.$

§ 9. Product of two determinants of the same order.

Rule for the multiplication of two determinants of the third order.

Let $$\Delta_1=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \text{ and } \Delta_2=\begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix}$$

be two determinants of the third order.

Then

$$\Delta_1.\Delta_2=\begin{vmatrix} a_1\alpha_1+b_1\alpha_2+c_1\alpha_3 & a_1\beta_1+b_1\beta_2+c_1\beta_3 & a_1\gamma_1+b_1\gamma_2+c_1\gamma_3 \\ a_2\alpha_1+b_2\alpha_2+c_2\alpha_3 & a_2\beta_1+b_2\beta_2+c_2\beta_3 & a_2\gamma_1+b_2\gamma_2+c_2\gamma_3 \\ a_3\alpha_1+b_3\alpha_2+c_3\alpha_3 & a_3\beta_1+b_3\beta_2+c_3\beta_3 & a_3\gamma_1+b_3\gamma_2+c_3\gamma_3 \end{vmatrix}.$$

This is the *row-by-column multiplication rule* for writing down the product of two determinants of the third order in the form of a determinant of the third order. This rule is applicable for writing down the product of two determinants of the n^{th} order, where n is arbitrary. It should be noted that this multiplication rule is the same as the rule for the multiplication of two matrices.

Other ways of multiplying the determinants. The value of a determinant does not change by interchanging the rows and columns. Therefore while writing down the product of two determinants of the same order, we can also follow the row-by-row multiplication rule, or the column-by-row multiplication rule or column-by-column multiplication rule.

Solved Examples

Ex. 1. *Show that*

$$\begin{vmatrix} 0 & c & b \\ c & 0 & a \\ b & a & 0 \end{vmatrix}^2=\begin{vmatrix} b^2+c^2 & ab & ac \\ ab & c^2+a^2 & bc \\ ac & bc & a^2+b^2 \end{vmatrix}.$$

(Gorakhpur 1985; Meerut 87)

Solution. We have

$$\begin{vmatrix}0&c&b\\c&0&a\\b&a&0\end{vmatrix}^2=\begin{vmatrix}0&c&b\\c&0&a\\b&a&0\end{vmatrix}\times\begin{vmatrix}0&c&b\\c&0&a\\b&a&0\end{vmatrix}$$

$$=\begin{vmatrix}0+c^2+b^2&0+0+ab&0+ac+0\\0+0+ab&c^2+0+a^2&bc+0+0\\0+ac+0&bc+0+0&b^2+a^2+0\end{vmatrix}\quad\text{applying row by column rule of multiplication.}$$

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

Ex. 2. *Express*

$$\Delta=\begin{vmatrix}(a-x)^2&(b-x)^2&(c-x)^2\\(a-y)^2&(b-y)^2&(c-y)^2\\(a-z)^2&(b-z)^2&(c-z)^2\end{vmatrix}$$

as product of two determinants and find its value.

(Lucknow 1979; Rohilkhand 81; Gorakhpur 84)

Solution. We have

$$\Delta=\begin{vmatrix}a^2-2ax+x^2&b^2-2bx+x^2&c^2-2cx+x^2\\a^2-2ay+y^2&b^2-2by+y^2&c^2-2cy+y^2\\a^2-2az+z^2&b^2-2bz+z^2&c^2-2cz+z^2\end{vmatrix}.$$

Since we are to express Δ as the product of two determinants of the third order, therefore we have written each element of Δ in such a way that it is expressed as the sum of three terms. Now by trial and inspection, we observe that

$$\Delta=\begin{vmatrix}1&x&x^2\\1&y&y^2\\1&z&z^2\end{vmatrix}\times\begin{vmatrix}a^2&-2a&1\\b^2&-2b&1\\c^2&-2c&1\end{vmatrix}\quad\text{row-by-row multiplication}$$

$$=2\begin{vmatrix}1&x&x^2\\1&y&y^2\\1&z&z^2\end{vmatrix}\times\begin{vmatrix}1&a&a^2\\1&b&b^2\\1&c&c^2\end{vmatrix}$$

Now we can easily show that

$$\begin{vmatrix}1&x&x^2\\1&y&y^2\\1&z&z^2\end{vmatrix}=(x-y)(y-z)(z-x).$$

Hence $\Delta=(x-y)(y-z)(z-x)(a-b)(b-c)(c-a)$.

Note. Each element of Δ has been expressed as the sum of three terms. Also in the first column a^2 is common to the first term of each element, $-2a$ is common to the second term of each element and 1 is common to the third term of each element. Leaving aside these common factors we get one determinant. These common factors give us the first row of the second determinant. Similarly for other columns.

Ex. 3. *Express*

$$\begin{vmatrix}(1+ax)^2&(1+ay)^2&(1+az)^2\\(1+bx)^2&(1+by)^2&(1+bz)^2\\(1+cx)^2&(1+cy)^2&(1+cz)^2\end{vmatrix}$$

as product of two determinants. **(Rohilkhand 1979)**

Solution. The given determinant is

$$=\begin{vmatrix}1+2ax+a^2x^2&1+2ay+a^2y^2&1+2az+a^2z^2\\1+2bx+b^2x^2&1+2by+b^2y^2&1+2bz+b^2z^2\\1+2cx+c^2x^2&1+2cy+c^2y^2&1+2cz+c^2z^2\end{vmatrix}$$

$$=\begin{vmatrix}1&2a&a^2\\1&2b&b^2\\1&2c&c^2\end{vmatrix}\times\begin{vmatrix}1&x&x^2\\1&y&y^2\\1&z&z^2\end{vmatrix},\ \text{with the help of row-by-row multiplication rule.}$$

Ex. 4. *Prove that*

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

$$=(a^3+b^3+c^3-3abc)^2.$$

(Lucknow 1978; Gorakhpur 81; Kanpur 89)

Solution. We are to express Δ as the product of two determinants of the third order. Therefore we should write each element of Δ in such a way that it is expressed as the sum of three terms.

Thus, we write

$$\Delta=\begin{vmatrix}-a^2+cb+bc&-ab+ab+c^2&-ac+b^2+ca\\-ab+c^2+ba&-b^2+ac+ca&-bc+bc+a^2\\-ac+ca+b^2&-bc+a^2+bc&-c^2+ab+ab\end{vmatrix}.$$

By trial and inspection, we find that

$$\Delta=\begin{vmatrix}a&b&c\\b&c&a\\c&a&b\end{vmatrix}\times\begin{vmatrix}-a&c&b\\-b&a&c\\-c&b&a\end{vmatrix}\quad\text{(row-by-row multiplication)}$$

$$=\begin{vmatrix}a&b&c\\b&c&a\\c&a&b\end{vmatrix}\times\begin{vmatrix}a&b&c\\b&c&a\\c&a&b\end{vmatrix}=\begin{vmatrix}a&b&c\\b&c&a\\c&a&b\end{vmatrix}^2.$$

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

[on expanding along the first row]

$=3abc-a^3-b^3-c^3.$

$\therefore \quad \Delta=(a^3+b^3+c^3-3abc)^2.$

Ex. 5. *If A_1, B_1, C_1, etc. denote the cofactors of a_1, b_1, c_1 etc. in*

$$\Delta=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}, \text{ then show that}$$

$$\Delta^2=\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix}.$$

[Delhi 1981; Rohilkhand 81; Lucknow 85]

Solution. Let $\Delta'=\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix}.$

Then applying the row-by-row multiplication rule, we get

$$\Delta'\Delta=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix}$$

$$=\begin{vmatrix} a_1A_1+b_1B_1+c_1C_1 & a_1A_2+b_1B_2+c_1C_2 & a_1A_3+b_1B_3+c_1C_3 \\ a_2A_1+b_2B_1+c_2C_1 & a_2A_2+b_2B_2+c_2C_2 & a_2A_3+b_2B_3+c_2C_3 \\ a_3A_1+b_3B_1+c_3C_1 & a_3A_2+b_3B_2+c_3C_2 & a_3A_3+b_3B_3+c_3C_3 \end{vmatrix}$$

$$=\begin{vmatrix} \Delta & 0 & 0 \\ 0 & \Delta & 0 \\ 0 & 0 & \Delta \end{vmatrix}=\Delta^3.$$

$\therefore \quad \Delta'=\Delta^2.$

Ex. 6. *Prove that*

$$\begin{vmatrix} yz-x^2 & zx-y^2 & xy-z^2 \\ zx-y^2 & xy-z^2 & yz-x^2 \\ xy-z^2 & yz-x^2 & zx-y^2 \end{vmatrix}=\begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}^2$$

(Kanpur 1986; Meerut 91P)

Solution. Let $\Delta=\begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}.$

The cofactor of the element x of the first row of

$$\Delta=yz-x^2.$$

The cofactor of the element y of the first row of

$$\Delta=-(y^2-xz)=xz-y^2.$$

The cofactor of the element z of the first row of

$$\Delta=xy-z^2.$$

Similarly find the cofactors of all the elements of Δ. Then as in Ex. 5, we have

$$\Delta^2=\begin{vmatrix} yz-x^2 & zx-y^2 & xy-z^2 \\ zx-y^2 & xy-z^2 & yz-x^2 \\ xy-z^2 & yz-x^2 & zx-y^2 \end{vmatrix}, \text{ replacing each element of } \Delta \text{ by its cofactor.}$$

Ex. 7. *If* **A** *and* **B** *be two square matrices of the same order, then*

$$|\,\mathbf{AB}\,|=|\,\mathbf{A}\,|\;|\,\mathbf{B}\,|.$$

Solution. Let $\mathbf{A}=\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$

and $\mathbf{B}=\begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{bmatrix}$

be two square matrices of the third order. Then we have

$$\mathbf{AB}=\begin{bmatrix} a_1\alpha_1+b_1\alpha_2+c_1\alpha_3 & a_1\beta_1+b_1\beta_2+c_1\beta_3 & a_1\gamma_1+b_1\gamma_2+c_1\gamma_3 \\ a_2\alpha_1+b_2\alpha_2+c_2\alpha_3 & a_2\beta_1+b_2\beta_2+c_2\beta_3 & a_2\gamma_1+b_2\gamma_2+c_2\gamma_3 \\ a_3\alpha_1+b_3\alpha_2+c_3\alpha_3 & a_3\beta_1+b_3\beta_2+c_3\beta_3 & a_3\gamma_1+b_3\gamma_2+c_3\gamma_3 \end{bmatrix}.$$

From our row-by-column rule for the multiplication of two determinants of the third order, we at once see that

$$|\,\mathbf{AB}\,|=|\,\mathbf{A}\,|.|\,\mathbf{B}\,|.$$

Ex. 8. *If* $\mathbf{A}=\begin{bmatrix} 2 & 3 & 1 \\ 1 & 4 & 2 \\ 0 & 1 & 1 \end{bmatrix}$ *and* $B=\begin{bmatrix} 1 & 6 & 0 \\ 3 & 2 & 1 \\ 1 & 2 & 3 \end{bmatrix}$

verify that $|\,\mathbf{AB}\,|=|\,\mathbf{A}\,|.|\,\mathbf{B}\,|$.

Solution. Applying row-by-column rule for the multiplication of two matrices, we have

$$\mathbf{AB}=\begin{bmatrix} 2.1+3.3+1.1 & 2.6+3.2+1.2 & 2.0+3.1+1.3 \\ 1.1+4.3+2.1 & 1.6+4.2+2.2 & 1.0+4.1+2.3 \\ 0.1+1.3+1.1 & 0.6+1.2+1.2 & 0.0+1.1+1.3 \end{bmatrix}$$

$$=\begin{bmatrix} 12 & 20 & 6 \\ 15 & 18 & 10 \\ 4 & 4 & 4 \end{bmatrix}.$$

Now $|\,\mathbf{AB}\,|=\begin{vmatrix} 12 & 20 & 6 \\ 15 & 18 & 10 \\ 4 & 4 & 4 \end{vmatrix}=\begin{vmatrix} 12 & 8 & -6 \\ 15 & 3 & -5 \\ 4 & 0 & 0 \end{vmatrix}$

applying $C_2 \to C_2-C_1$, $C_3 \to C_3-C_1$

$$=4\begin{vmatrix} 8 & -6 \\ 3 & -5 \end{vmatrix}=4\,(-40+18)=-88.$$

Also $|\,A\,|=\begin{vmatrix} 2 & 3 & 1 \\ 1 & 4 & 2 \\ 0 & 1 & 1 \end{vmatrix}=2\,(4-2)-1\,(3-1)=4-\;=\ldots$

and $|\,B\,|=\begin{vmatrix} 1 & 6 & 0 \\ 3 & 2 & 1 \\ 1 & 2 & 3 \end{vmatrix}=\begin{vmatrix} 1 & 0 & 0 \\ 3 & -16 & 1 \\ 1 & -4 & 3 \end{vmatrix}$ by $C_2\to C_2-6C_1$

$$=-48+4=-44.$$

$\therefore\ |\,A\,|.|\,B\,|=(2)\,(-44)=-88=|\,AB\,|.$

Ex. 9. *If $u=ax+by+cz, v=ay+bz+cx$ and $w=az+bx+cy$, prove that*

$$\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}\times\begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}=u^3+v^3+w^3-3uvw.$$

Solution. Let us denote the determinant on the left hand side by Δ_1 and Δ_2 respectively. Applying row-by-column rule of multiplication, we get

$$\Delta_1\Delta_2=\begin{vmatrix} ax+by+cz & ay+bz+cx & az+bx+cy \\ bx+cy+az & by+cz+ax & bz+cx+ay \\ cx+ay+bz & cy+az+bx & cz+ax+by \end{vmatrix}$$

$$=\begin{vmatrix} u & v & w \\ w & u & v \\ v & w & u \end{vmatrix}$$

$$=u\,(u^2-vw)-w\,(uv-w^2)+v\,(v^2-uw)$$

$$=u^3+v^3+w^3-3uvw.$$

Ex. 10. *If ω is one of the imaginary cube roots of unity, prove that*

$$\begin{vmatrix} 1 & \omega & \omega^2 & \omega^3 \\ \omega & \omega^2 & \omega^3 & 1 \\ \omega^2 & \omega^3 & 1 & \omega \\ \omega^3 & 1 & \omega & \omega^2 \end{vmatrix}^2=\begin{vmatrix} 1 & 1 & -2 & 1 \\ 1 & 1 & 1 & -2 \\ -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \end{vmatrix}.$$

(Rohilkhand 1979, Kanpur 79)

Solution. We have $\begin{vmatrix} 1 & \omega & \omega^2 & \omega^3 \\ \omega & \omega^2 & \omega^3 & 1 \\ \omega^2 & \omega^3 & 1 & \omega \\ \omega^3 & 1 & \omega & \omega^2 \end{vmatrix}^2$

$$=\begin{vmatrix} 1 & \omega & \omega^2 & 1 \\ \omega & \omega^2 & 1 & 1 \\ \omega^2 & 1 & 1 & \omega \\ 1 & 1 & \omega & \omega^2 \end{vmatrix}\times\begin{vmatrix} 1 & \omega & \omega^2 & 1 \\ \omega & \omega^2 & 1 & 1 \\ \omega^2 & 1 & 1 & \omega \\ 1 & 1 & \omega & \omega^2 \end{vmatrix}$$

$$=\begin{vmatrix} 1+\omega^2+\omega^4+1 & \omega+\omega^3+\omega^2+1 & \omega^2+\omega+\omega^2+\omega & 1+\omega+\omega^3+\omega^2 \\ \omega+\omega^3+\omega^2+1 & \omega^2+\omega^4+1+1 & \omega^3+\omega^2+1+\omega & \omega+\omega^2+\omega+\omega^2 \\ \omega^2+\omega+\omega^2+\omega & \omega^3+\omega^2+1+\omega & \omega^4+1+1+\omega^2 & \omega^2+1+\omega+\omega^3 \\ 1+\omega+\omega^3+\omega^2 & \omega+\omega^2+\omega+\omega^3 & \omega^2+1+\omega+\omega^3 & 1+1+\omega^2+\omega^4 \end{vmatrix}$$

$$=\begin{vmatrix} 1 & 1 & -2 & 1 \\ 1 & 1 & 1 & -2 \\ -2 & 1 & 1 & 1 \\ 1 & -2 & 1 & 1 \end{vmatrix}\quad[\because\ 1+\omega+\omega^2=0,\ \omega^3=1,\ \omega^4=\omega^3.\omega=\omega]$$

Ex. 11. *Show that*

$$\begin{vmatrix} a+ib & c+id \\ -c+id & a-ib \end{vmatrix}\times\begin{vmatrix} \alpha+i\beta & \gamma+i\delta \\ -\gamma+i\delta & \alpha-i\beta \end{vmatrix}$$ *can be expressed as* $$\begin{vmatrix} A+iB & C+iD \\ -C+iD & A-iB \end{vmatrix}$$

(Meerut 1980)

Hence show that the product of two numbers each of which is a sum of four squares is itself a sum of four squares.

Solution. Let $\Delta_1=\begin{vmatrix} a+ib & c+id \\ -c+id & a-ib \end{vmatrix}$,

and $\Delta_2=\begin{vmatrix} \alpha+i\beta & \gamma+i\delta \\ -\gamma+i\delta & \alpha-i\beta \end{vmatrix}$.

Applying row-by-column rule of multiplication, we get

$$\Delta_1\Delta_2=\begin{vmatrix} a\alpha-b\beta-c\gamma-d\delta+i\,(a\beta+c\delta+b\alpha-d\gamma) & a\gamma+c\alpha-b\delta+d\beta+i\,(d\alpha-c\beta+b\gamma+a\delta) \\ -a\gamma-c\alpha+b\delta-d\beta+i\,(d\alpha-c\beta+b\gamma+a\delta) & a\alpha-b\beta-c\gamma-d\delta-i\,(a\beta+c\delta+b\alpha-d\gamma) \end{vmatrix}$$

$$=\begin{vmatrix} A+iB & C+iD \\ -C+iD & A-iB \end{vmatrix},\ \text{where}$$

$A=a\alpha-b\beta-c\gamma-d\delta,\ B=a\beta+c\delta+b\alpha-d\gamma$

$C=a\gamma+c\alpha-b\delta+d\beta,\ D=d\alpha-c\beta+b\gamma+a\delta.$

Expanding the determinants on both sides, we get

$$(a^2+b^2+c^2+d^2)\,(\alpha^2+\beta^2+\gamma^2+\delta^2)=A^2+B^2+C^2+D^2.$$

Ex. 12. *Find the value of*

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

Solution. Applying row-by-row multiplication rule for the product of two determinants, the given product of two determinants

$$=\begin{vmatrix} \lambda^3+a^2\lambda+b^2\lambda+c^2\lambda & 0 & 0 \\ 0 & \lambda^3+a^2\lambda+b^2\lambda+c^2\lambda & 0 \\ 0 & 0 & \lambda^3+a^2\lambda+b^2\lambda+c^2\lambda \end{vmatrix}$$

$$=(\lambda^3+a^2\lambda+b^2\lambda+c^2\lambda)^3=\lambda^3\,(\lambda^2+a^2+b^2+c^2)^3.$$

Ex. 13. *By squaring the determinant*

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ \alpha & \beta & \gamma & \delta \\ \alpha^2 & \beta^2 & \gamma^2 & \delta^2 \\ \alpha^3 & \beta^3 & \gamma^3 & \delta^3 \end{vmatrix}$$

show that

$$\begin{vmatrix} s_0 & s_1 & s_2 & s_3 \\ s_1 & s_2 & s_3 & s_4 \\ s_2 & s_3 & s_4 & s_5 \\ s_3 & s_4 & s_5 & s_6 \end{vmatrix} =[(\alpha-\beta)(\alpha-\gamma)(\alpha-\delta)(\beta-\gamma)(\beta-\delta)(\gamma-\delta)]^2,$$

where $s_r=\alpha^r+\beta^r+\gamma^r+\delta^r$.

[Kanpur 1985; Poona 70)

Solution. First apply row-by-row multiplication rule to get the square of the given determinant in the form given in the question.

Now find the value of the given determinant by first applying C_2-C_1, C_3-C_1, C_4-C_1 and then proceed as usual.

Ex. 14. *Let*

$$D=\begin{vmatrix} 2a_1b_1 & a_1b_2+a_2b_1 & a_1b_3+a_3b_1 \\ a_1b_2+a_2b_1 & 2a_2b_2 & a_2b_3+a_3b_2 \\ a_1b_3+a_3b_1 & a_3b_2+a_2b_3 & 2a_3b_3 \end{vmatrix}.$$

Express the determinant D as a product of two determinants. Hence or otherwise show that $D=0$. **(Kolhapur 1972)**

Solution. We have D

$$=\begin{vmatrix} a_1 & b_1 & 0 \\ a_2 & b_2 & 0 \\ a_3 & b_3 & 0 \end{vmatrix} \times \begin{vmatrix} b_1 & a_1 & 0 \\ b_2 & a_2 & 0 \\ b_3 & a_3 & 0 \end{vmatrix},$$ as can be seen by applying row-by-row multiplication rule.

Hence $D=0$.

§ 10. System of Non-homogeneous Linear Equations. (Cramer's Rule). The equation $ax+by+cz=0$ is a linear homogeneous equation in x, y and z. On the other hand the equation $ax+by+cz+d=0$ is a linear non-homogeneous equation in x, y and z.

Let the n linear simultaneous equations in n unknowns $x_1, x_2, x_3, \ldots, x_n$ be

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

Let $$\Delta=\begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} \neq 0.$$

Suppose $A_{11}, A_{12}, A_{13}, \ldots$ etc. denote the cofactors of $a_{11}, a_{12}, a_{13}, \ldots$ etc. in Δ. Then multiplying the given equations respectively by $A_{11}, A_{21}, A_{31}, \ldots, A_{n1}$ and adding, we get

$$x_1\,(a_{11}A_{11}+a_{21}A_{21}+\ldots+a_{n1}A_{n1})+x_2\,(0)+x_3\,(0)+\ldots+x_n\,(0)=b_1A_{11}+b_2A_{21}+\ldots+b_nA_{n1}$$

or $$x_1\,\Delta=\Delta_1,$$

where Δ_1 is the determinant obtained by replacing the elements in the first column of Δ by the elements $b_1, b_2, \ldots, b_n$.

Again multiplying the given equations respectively by $A_{12}, A_{22}, \ldots, A_{n2}$ and adding, we get

$$x_2\Delta=\Delta_2,$$

where Δ_2 is the determinant obtained by replacing the elements in the second column of Δ by the elements $b_1, b_2, \ldots, b_n$.

Similarly, we get

$$x_3\Delta=\Delta_3$$

$$\ldots\ldots\ldots\ldots$$

$$x_n\Delta=\Delta_n.$$

This method of solving n simultaneous equations in n unknowns is known as Cramer's Rule.

Thus by Cramer's rule, if $\Delta\neq 0$, we have

$$x_i = \frac{\Delta_i}{\Delta},\ i = 1, 2, \ldots, n,$$

where Δ_i is the determinant obtained by replacing the ith column in Δ by the elements $b_1, b_2, \ldots, b_n$.

Solved Examples

Ex. 1. *Solve the following system of linear equations with the help of Cramer's rule :*

$$x+2y+3z=6,$$
$$2x+4y+z=7,$$
$$2x+2y+9z=14. \qquad \textbf{(Gorakhpur 1984)}$$

Solution.

Let $$\Delta = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 2 & 9 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & -5 \\ 0 & -4 & 0 \end{vmatrix} \text{ by } R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1$$

$$= -20.$$

Thus $\Delta \neq 0$ and therefore the system has a unique solution given by $\frac{x}{\Delta_1} = \frac{y}{\Delta_2} = \frac{z}{\Delta_3} = \frac{1}{\Delta}$, *i.e.*, given by

$$\frac{x}{\begin{vmatrix} 6 & 2 & 3 \\ 7 & 4 & 1 \\ 14 & 2 & 9 \end{vmatrix}} = \frac{y}{\begin{vmatrix} 1 & 6 & 3 \\ 2 & 7 & 1 \\ 3 & 14 & 9 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 1 & 2 & 6 \\ 2 & 4 & 7 \\ 3 & 2 & 14 \end{vmatrix}} = \frac{1}{\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 1 \\ 3 & 2 & 9 \end{vmatrix}}.$$

Now $$\begin{vmatrix} 6 & 2 & 3 \\ 7 & 4 & 1 \\ 14 & 2 & 9 \end{vmatrix} = \begin{vmatrix} 6 & 2 & 3 \\ -5 & 0 & -5 \\ 8 & 0 & 6 \end{vmatrix} \text{ by } R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1$$

$$= -2\,(-30+40) = -20.$$

Again $$\begin{vmatrix} 1 & 6 & 3 \\ 2 & 7 & 1 \\ 3 & 14 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 6 & 3 \\ 0 & -5 & -5 \\ 0 & -4 & 0 \end{vmatrix} \quad R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1$$

$$= -20.$$

Also $$\begin{vmatrix} 1 & 2 & 6 \\ 2 & 4 & 7 \\ 3 & 2 & 14 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 6 \\ 0 & 0 & -5 \\ 0 & -4 & -4 \end{vmatrix} \quad R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1$$

$\therefore$ the solution is given by

$$\frac{x}{-20} = \frac{y}{-20} = \frac{z}{-20} = \frac{1}{-20}.$$

Hence $x=1,\ y=1,\ z=1.$

Ex. 2. *Solve the following equations by Cramer's rule :*

$$2x-y+3z=9,$$
$$x+y+z=6,$$
$$x-y+z=2 \qquad \textbf{(Meerut 1977)}$$

Solution.

Let $$\Delta = \begin{vmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 0 & -2 & 0 \end{vmatrix} \quad R_3 \to R_3 - R_2$$

$$= -(-2)\,(2-3) = -2.$$

Thus $\Delta \neq 0$ and therefore the system has a unique solution given by $\frac{x}{\Delta_1} = \frac{y}{\Delta_2} = \frac{z}{\Delta_3} = \frac{1}{\Delta}$, *i.e.*, given by

$$\frac{x}{\begin{vmatrix} 9 & -1 & 3 \\ 6 & 1 & 1 \\ 2 & -1 & 1 \end{vmatrix}} = \frac{y}{\begin{vmatrix} 2 & 9 & 3 \\ 1 & 6 & 1 \\ 1 & 2 & 1 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 2 & -1 & 9 \\ 1 & 1 & 6 \\ 1 & -1 & 2 \end{vmatrix}} = \frac{1}{\begin{vmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{vmatrix}}$$

or given by $\frac{x}{-2} = \frac{y}{-4} = \frac{z}{-6} = \frac{1}{-2}.$

Hence $x=1,\ y=2,\ z=3.$

Ex. 3. *Solve the following system of linear equations by Cramer's rule :*

$$x+y+z=7,$$
$$x+2y+3z=16,$$
$$x+3y+4z=22. \qquad \textbf{(Gorakhpur 1980)}$$

Solution. We have

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{vmatrix} = -1.$$

Thus $\Delta \neq 0$ and therefore the system has a unique solution given by $\frac{x}{\Delta_1} = \frac{y}{\Delta_2} = \frac{z}{\Delta_3} = \frac{1}{\Delta}$, *i.e.* given by

$$\frac{x}{\begin{vmatrix} 7 & 1 & 1 \\ 16 & 2 & 3 \\ 22 & 3 & 4 \end{vmatrix}} = \frac{y}{\begin{vmatrix} 1 & 7 & 1 \\ 1 & 16 & 3 \\ 1 & 22 & 4 \end{vmatrix}} = \frac{z}{\begin{vmatrix} 1 & 1 & 7 \\ 1 & 2 & 16 \\ 1 & 3 & 22 \end{vmatrix}} = \frac{1}{\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{vmatrix}}$$

or given by $\frac{x}{\cdots} = \frac{y}{\cdots} = \frac{z}{\cdots} = \frac{1}{\cdots}.$

Hence $x=1$, $y=3$, $z=3$.

Ex. 4. *If a, b, c are all different, solve the system of equations :*

$$\begin{aligned} x+y+z&=1,\\ ax+by+cz&=k,\\ a^2x+b^2y+c^2z&=k^2. \end{aligned}$$

(Delhi 19..)

Solution. Let $\Delta=\begin{vmatrix} 1 & 1 & 1\\ a & b & c\\ a^2 & b^2 & c^2 \end{vmatrix}$

$=(a-b)(b-c)(c-a).$

Since a, b, c are all different, therefore $\Delta\neq 0$.

Hence the given system has a unique solution given by

$\dfrac{x}{\Delta_1}=\dfrac{y}{\Delta_2}=\dfrac{z}{\Delta_3}=\dfrac{1}{\Delta}$, *i.e.*, given by

$$\frac{x}{\begin{vmatrix} 1 & 1 & 1\\ k & b & c\\ k^2 & b^2 & c^2 \end{vmatrix}}=\frac{y}{\begin{vmatrix} 1 & 1 & 1\\ a & k & c\\ a^2 & k^2 & c^2 \end{vmatrix}}=\frac{z}{\begin{vmatrix} 1 & 1 & 1\\ a & b & k\\ a^2 & b^2 & k^2 \end{vmatrix}}=\frac{1}{\begin{vmatrix} 1 & 1 & 1\\ a & b & c\\ a^2 & b^2 & c^2 \end{vmatrix}}.$$

Hence $x=\dfrac{(k-b)(b-c)(c-k)}{(a-b)(b-c)(c-a)}$, $y=\dfrac{(a-k)(k-c)(c-a)}{(a-b)(b-c)(c-a)}$,

$z=\dfrac{(a-b)(b-k)(k-a)}{(a-b)(b-c)(c-a)}$.

Exercises

1. Show that

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

2. Give the correct answer out of the following :

The value of the determinant

$$\begin{vmatrix} -3 & 1 & 1 & 1\\ 1 & -3 & 1 & 1\\ 1 & 1 & -3 & 1\\ 1 & 1 & 1 & -3 \end{vmatrix} \text{ is}$$

(*A*) -1, (*B*) 1, (*C*) 0, (*D*) 4. **(Meerut 1977)**

3. Prove that

$$\begin{vmatrix} 1+x & 2 & 3 & 4\\ 1 & 2+x & 3 & 4\\ 1 & 2 & 3+x & 4\\ 1 & 2 & 3 & 4+x \end{vmatrix}=x^3(x+10).$$

(Agra 1980)

4. Show that

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

(Meerut 1979)

5. Evaluate

$$\begin{vmatrix} a & 1 & 1 & 1\\ 1 & a & 1 & 1\\ 1 & 1 & a & 1\\ 1 & 1 & 1 & a \end{vmatrix}.$$

(Agra 1974)

6. Show that

$$\begin{vmatrix} x+a & b & c & d\\ a & x+b & c & d\\ a & b & x+c & d\\ a & b & c & x+d \end{vmatrix}=x^3(x+a+b+c+d).$$

(Meerut 1983; Poona 70)

7. Show that

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

(Rajasthan 1963)

8. Express

$$\begin{vmatrix} 0 & (\alpha-\beta)^2 & (\alpha-\gamma)^2\\ (\alpha-\beta)^2 & 0 & (\beta-\gamma)^2\\ (\alpha-\gamma)^2 & (\beta-\gamma)^2 & 0 \end{vmatrix}$$

as a product of two determinants of the third order and hence find its value. **(Poona 1970)**

9. Describe Cramer's rule of finding solutions of simultaneous equations in four unknowns.

Answers

2. (*C*). **5.** $(a+3)(a-1)^3$.

8. $(\alpha-\beta)^2(\beta-\gamma)^2(\gamma-\alpha)^2$.

3

Inverse of a Matrix

***§ 1. Adjoint of a square matrix. Definition.**

Let $A=[a_{ij}]_{n\times n}$ *be any* $n\times n$ *matrix. The transpose* B' *of the matrix*

$$B=[A_{ij}]_{n\times n},$$

where A_{ij} *denotes the cofactor of the element* a_{ij} *in the determinant* $|A|$, *is called the adjoint of the matrix A and is denoted by the symbol Adj A.* **(Meerut 1980, 82, 83; Ranchi 70; Poona 70; Rohilkhand 90, 91; Karnatak 68; Allahabad 79)**

Thus the adjoint of a matrix A is the transpose of the matrix formed by the cofactors of A *i.e.*, if

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & \dots & a_{nn} \end{bmatrix},$$

then Adj A = the transpose of the matrix $\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix}$

= the matrix $\begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix}.$

Note. Sometimes the adjoint of a matrix is also called the **adjugate** of the matrix.

****§ 2. Theorem.** *If* A *be any n-rowed square matrix, then*

$$(Adj\ A)\ A=A\ (Adj\ A)=|A|\ I_n,$$

where I_n *is the n-rowed unit matrix.*

(Meerut 1991; Agra 74; Delhi 81; Kanpur 84; Rohilkhand 90)

The theorem states that the matrices A and Adj A are commutative and that their product is a scalar matrix every diagonal element of which is $|A|$.

Proof. Let $A=[a_{ij}]_{n\times n}$ be any n-rowed square matrix.

Let $\quad$ Adj $A=[b_{ij}]_{n\times n}$.

Then $b_{ij}=A_{ji}=$ co-factor of a_{ji} in $|A|$. ...(1)

Since the matrices A and Adj A are both n-rowed square matrices, therefore both the products A (adj A) and (adj. A) A exist and are of the type $n\times n$.

Also, the $(i, j)^{th}$ element of A (adj A)

$$=\sum_{k=1}^{n} a_{ik}\, b_{kj} \quad \text{[by def. of product of two matrices]}$$

$$=\sum_{k=1}^{n} a_{ik}\, A_{jk} \quad \text{[from (1)]}$$

$=0$ or $|A|$ according as $i\neq j$ or $i=j$.

Hence the $(i, j)^{th}$ element of A (Adj. A) $=|A|$ if $i=j$ and $=0$ if $i\neq j$. In other words all the elements of A (Adj. A) along the principal diagonal are equal to $|A|$ and the non-diagonal elements are all equal to zero.

Therefore A (adj. A)

$$=\begin{bmatrix} |A| & 0 & 0 & \dots & 0 \\ 0 & |A| & 0 & \dots & 0 \\ 0 & 0 & |A| & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & |A| \end{bmatrix}=|A|\ I_n.$$

Similarly, the $(i, j)^{th}$ element of (Adj. A) A

$$=\sum_{k=1}^{n} b_{ik}\, a_{kj}=\sum_{k=1}^{n} A_{ki}\, a_{kj}$$

$=0$ or $|A|$ according as $i\neq j$ or $i=j$.

Therefore (Adj. A) $A=|A|\ I_n$. Hence the theorem.

Example 1. *If* $A=\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, *then find Adj. A.*

Solution. In $|A|$, the cofactor of α is δ and the cofactor of β is $-\gamma$. Also the cofactor of γ is $-\beta$ and the cofactor of δ is α. Therefore the matrix B formed of the cofactors of the elements of $|A|$ is

$$B=\begin{bmatrix} \delta & -\gamma \\ -\beta & \alpha \end{bmatrix}.$$

Existence of the Inverse. Theorem. *The necessary and sufficient condition for a square matrix* **A** *to possess the inverse is that* $|A| \neq 0$. (Meerut 1987; I.A.S. 73; Kanpur 81; Delhi 80; Rohilkhand 79, 81; Jodhpur 65; Gorakhpur 85)

Proof. The condition is necessary. Let **A** be an $n \times n$ matrix and let **B** be the inverse of **A**.

Then $$AB = I_n.$$

$\therefore$ $|AB| = |I_n| = 1.$

$\therefore$ $|A||B| = 1.$ $\quad [\because |AB| = |A||B|]$

$\therefore$ $|A|$ must be different from 0.

Conversely, the condition is also sufficient.

If $|A| \neq 0$, then let us define a matrix **B** by the relation

$$B = \frac{1}{|A|}(\text{adj. } A).$$

Then $$AB = A\left(\frac{1}{|A|}\text{ adj. } A\right)$$

$$= \frac{1}{|A|}(A \text{ adj. } A) = \frac{1}{|A|}|A| I_n = I_n.$$

Similarly $$BA = \left(\frac{1}{|A|}\text{ adj. } A\right)A = \frac{1}{|A|}(\text{adj. } A)\, A$$

$$= \frac{1}{|A|} \cdot |A| I_n = I_n.$$

Thus $AB = BA = I_n$.

Hence the matrix **A** is invertible and **B** is the inverse of **A**.

Important. *If* **A** *be an invertible matrix, then the inverse of* **A** *is* $\frac{1}{|A|}$ *Adj.* **A**. *It is usual to denote the inverse of* **A** *by* A^{-1}.

Non-singular and singular matrices.

Definition. *A square matrix* **A** *is said to be non-singular or singular according as* $|A| \neq 0$ *or* $|A| = 0$. (Meerut 1986)

Thus the necessary and sufficient condition for a matrix to be invertible is that it is non-singular.

§ 4. Reversal law for the inverse of a product.

Theorem 1. *If* **A, B** *be two n-rowed non-singular matrices, then* **AB** *is also non-singular and*

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

i.e. the inverse of a product is the product of the inverses taken in the reverse order. (Allahabad 1971; Poona 70; I.A.S. 69; Gorakhpur 71; Delhi 81; Rohilkhand 81; Sagar 66; Agra 69)

Proof. Let **A** and **B** be two n-rowed non-singular matrices.

We have $\quad |AB| = |A||B|$.

Since $\quad |A| \neq 0$ and $|B| \neq 0$,

therefore $|AB| \neq 0$. Hence the matrix **AB** is invertible.

Let us define a matrix **C** by the relation $C = B^{-1} A^{-1}$.

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

Also $(AB)\, C = (AB)(B^{-1} A^{-1}) = A(BB^{-1}) A^{-1} = AI_n A^{-1}$
$= AA^{-1}$ $\quad [\because AI_n = A]$
$= I_n$.

Thus $C(AB) = (AB)\, C = I_n$.

Hence $\quad C = B^{-1} A^{-1}$ is the inverse of **AB**.

***Theorem 2.** *If* **A** *be an* $n \times n$ *non-singular matrix, then*

$$(A')^{-1} = (A^{-1})'$$

i.e. the operations of transposing and inverting are commutative. (Madurai 1985; Meerut 74; Agra 76)

Proof. Since $|A'| = |A| \neq 0$, therefore the matrix **A′** is also non-singular.

We have the identity $AA^{-1} = A^{-1} A = I_n$.

Taking transposes of both sides and applying the reversal law for transposes, we get

$$(AA^{-1})' = (A^{-1} A)' = I_n'$$

or $\quad (A^{-1})' A' = A' (A^{-1})' = I_n. \quad [\because I_n' = I_n].$

This equality shows that $(A^{-1})'$ is the inverse of the matrix **A′**.

Hence $\quad (A')^{-1} = (A^{-1})'$

i.e. the inverse of the transpose of a matrix is the transpose of the inverse.

Theorem 3. *If* **A** *be an* $n \times n$ *non-singular matrix, then*

$$(A^{-1})^{\theta} = (A^{\theta})^{-1}.$$

Proof. Since $|A^{\theta}| = \overline{|A|} \neq 0$, therefore the matrix A^{θ} is also non-singular.

We have the identity $AA^{-1} = A^{-1} A = I_n$.

Taking the conjugate transposes of both sides, we get

$$(AA^{-1})^{\theta} = (A^{-1} A)^{\theta} = (I_n)^{\theta}.$$

Applying the reversal law for transposed conjugates, we have

$$(A^{-1})^\theta A^\theta = A^\theta (A^{-1})^\theta = I_n. \qquad [\because I_n^\theta = I_n]$$

This equality shows that $(A^{-1})^\theta$ is the inverse of the matrix A^θ.

Hence $$(A^{-1})^\theta = (A^\theta)^{-1}.$$

Solved Examples.

Ex. 1. *Find the adjoint of the matrix*

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

Solution. We have $|A| = \begin{vmatrix} -1 & -2 & 3 \\ -2 & 1 & 1 \\ 4 & -5 & 2 \end{vmatrix}$.

The cofactors of the elements of the first row of the determinant $|A|$ are $\begin{vmatrix} 1 & 1 \\ -5 & 2 \end{vmatrix}, -\begin{vmatrix} -2 & 1 \\ 4 & 2 \end{vmatrix}, \begin{vmatrix} -2 & 1 \\ 4 & -5 \end{vmatrix}$ *i.e.* are 7, 8, 6 respectively.

The cofactors of the elements of the second row of the determinant $|A|$ are $-\begin{vmatrix} -2 & 3 \\ -5 & 2 \end{vmatrix}, \begin{vmatrix} -1 & 3 \\ 4 & 2 \end{vmatrix}, -\begin{vmatrix} -1 & -2 \\ 4 & -5 \end{vmatrix}$, *i.e.* are $-11, -14, -13$ respectively.

The cofactors of the elements of the third row of the determinant $|A|$ are $\begin{vmatrix} -2 & 3 \\ 1 & 1 \end{vmatrix}, -\begin{vmatrix} -1 & 3 \\ -2 & 1 \end{vmatrix}, \begin{vmatrix} -1 & -2 \\ -2 & 1 \end{vmatrix}$, *i.e.* are $-5, -5, -5$ respectively.

Therefore the Adj. A = the transpose of the matrix **B** where

$$B = \begin{bmatrix} 7 & 8 & 6 \\ -11 & -14 & -13 \\ -5 & -5 & -5 \end{bmatrix}. \quad \therefore \text{Adj. } A = \begin{bmatrix} 7 & -11 & -5 \\ 8 & -14 & -5 \\ 6 & -13 & -5 \end{bmatrix}.$$

Ex. 2. *Find the adjoint of the matrix*

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

(Meerut 1969; Rohilkhand 81; Allahabad 71)

Solution. We have $|A| = \begin{vmatrix} 1 & 2 & 3 \\ 0 & 5 & 0 \\ 2 & 4 & 3 \end{vmatrix}$.

The cofactors of the elements of the first row of $|A|$ are $\begin{vmatrix} 5 & 0 \\ 4 & 3 \end{vmatrix}, -\begin{vmatrix} 0 & 0 \\ 2 & 3 \end{vmatrix}, \begin{vmatrix} 0 & 5 \\ 2 & 4 \end{vmatrix}$ *i.e.*, are 15, 0, -10 respectively.

The cofactors of the elements of the second row of $|A|$ are $-\begin{vmatrix} 2 & 3 \\ 4 & 3 \end{vmatrix}, \begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix}, -\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}$ *i.e.*, are 6, -3, 0 respectively.

The cofactors of the elements of the third row of $|A|$ are $\begin{vmatrix} 2 & 3 \\ 5 & 0 \end{vmatrix}, -\begin{vmatrix} 1 & 3 \\ 0 & 0 \end{vmatrix}, \begin{vmatrix} 1 & 2 \\ 0 & 5 \end{vmatrix}$ *i.e.*, are -15, 0, 5 respectively.

Therefore the adj. A = the transpose of the matrix **B** where

$$B = \begin{bmatrix} 15 & 0 & -10 \\ 6 & -3 & 0 \\ -15 & 0 & 5 \end{bmatrix}. \quad \therefore \text{adj. } A = \begin{bmatrix} 15 & 6 & -15 \\ 0 & -3 & 0 \\ -10 & 0 & 5 \end{bmatrix}$$

Ex. 3. *Find the adjoint of the matrix*

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

and verify the theorem $A\,(\text{adj } A) = (\text{adj } A)\, A = |A|\, I.$

(Agra 1968; Meerut 69)

Solution. We have $|A| = \begin{vmatrix} 1 & 2 \\ 3 & -5 \end{vmatrix} = 1\,(-5) - (3)\,(2)$
$= -5 - 6 = -11.$

The cofactors of the elements of the first row of $|A|$ are -5, -3 respectively.

The cofactors of the elements of the second row of $|A|$ are -2, 1 respectively.

Therefore adj. A = the transpose of the matrix **B** where

$$B = \begin{bmatrix} -5 & -3 \\ -2 & 1 \end{bmatrix}. \quad \therefore \text{adj. } A = \begin{bmatrix} -5 & -2 \\ -3 & 1 \end{bmatrix}.$$

Now
$$A\,(\text{adj } A) = \begin{bmatrix} 1 & 2 \\ 3 & -5 \end{bmatrix} \begin{bmatrix} -5 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -5-6 & -2+2 \\ -15+15 & -6-5 \end{bmatrix}$$
$$= \begin{bmatrix} -11 & 0 \\ 0 & -11 \end{bmatrix} = (-11) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = |A|\, I_2.$$

Also
$$(\text{adj } A)\, A = \begin{bmatrix} -5 & -2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & -5 \end{bmatrix} = \begin{bmatrix} -5-6 & -10+10 \\ -3+3 & -6-5 \end{bmatrix}$$
$$= \begin{bmatrix} -11 & 0 \\ 0 & -11 \end{bmatrix} = |A|\, I_2.$$

Hence $A\,(\text{adj } A) = (\text{adj } A)\, A = |A|\, I.$

Ex. 4. *Find the inverse of the matrix*

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}.$$

(Kanpur 1979; Meerut 87)

Solution. We have

$$|\mathbf{A}|=\begin{vmatrix}1&3&3\\1&4&3\\1&3&4\end{vmatrix}=\begin{vmatrix}1&3&3\\0&1&0\\0&0&1\end{vmatrix}\text{ applying } R_2\to R_2-R_1,\ R_3\to R_3-R_1$$

$=1$, on expanding the determinant along the first column.

Since $|\mathbf{A}|\neq 0$, therefore the matrix $\mathbf{A}$ is non-singular and possesses inverse.

Now the cofactors of the elements of the first row of the determinant $|\mathbf{A}|$ are

$\begin{vmatrix}4&3\\3&4\end{vmatrix}, -\begin{vmatrix}1&3\\1&4\end{vmatrix}, \begin{vmatrix}1&4\\1&3\end{vmatrix}$ *i.e.*, are $7, -1, -1$ respectively.

The cofactors of the elements of the second row of $|\mathbf{A}|$ are

$-\begin{vmatrix}3&3\\3&4\end{vmatrix}, \begin{vmatrix}1&3\\1&4\end{vmatrix}, -\begin{vmatrix}1&3\\1&3\end{vmatrix}$ *i.e.*, are $-3, 1, 0$ respectively.

The cofactors of the elements of the third row of $|\mathbf{A}|$ are

$\begin{vmatrix}3&3\\4&3\end{vmatrix}, -\begin{vmatrix}1&3\\1&3\end{vmatrix}, \begin{vmatrix}1&3\\1&4\end{vmatrix}$ *i.e.*, are $-3, 0, 1$ respectively.

Therefore the Adj. $\mathbf{A}$= the transpose of the matrix

$$\begin{bmatrix}7&-1&-1\\-3&1&0\\-3&0&1\end{bmatrix}$$

$$\therefore\quad \text{Adj }\mathbf{A}=\begin{bmatrix}7&-3&-3\\-1&1&0\\-1&0&1\end{bmatrix}.$$

Now $\mathbf{A}^{-1}=\dfrac{1}{|\mathbf{A}|}\text{Adj }\mathbf{A}=\begin{bmatrix}7&-3&-3\\-1&1&0\\-1&0&1\end{bmatrix}$, since $|\mathbf{A}|=1$.

Note. After finding the inverse of a matrix $\mathbf{A}$, we must check our answer by verifying the relation $\mathbf{AA}^{-1}=\mathbf{I}$.

Ex. 5. *Find the inverse of the matrix*

$$\mathbf{A}=\begin{bmatrix}0&1&2\\1&2&3\\3&1&1\end{bmatrix}.$$

(Vikram 1990; Kerala 85; Gorakhpur 81; Agra 83; Meerut 86; Delhi 88)

Solution. We have

$$|\mathbf{A}|=\begin{vmatrix}0&1&2\\1&2&3\\3&1&1\end{vmatrix}=\begin{vmatrix}0&1&0\\1&2&-1\\3&1&-1\end{vmatrix}\text{, applying } C_3\to C_3-2C_2$$

$=-1\begin{vmatrix}1&-1\\3&-1\end{vmatrix}$, expanding the determinant along the first row

$=-2.$

Since $|\mathbf{A}|\neq 0$, therefore $\mathbf{A}^{-1}$ exists.

Now the cofactors of the elements of the first row of $|\mathbf{A}|$ are

$\begin{vmatrix}2&3\\1&1\end{vmatrix}, -\begin{vmatrix}1&3\\3&1\end{vmatrix}, \begin{vmatrix}1&2\\3&1\end{vmatrix}$ *i.e.*, are $-1, 8, -5$ respectively.

The cofactors of the elements of the second row of $|\mathbf{A}|$ are

$-\begin{vmatrix}1&2\\1&1\end{vmatrix}, \begin{vmatrix}0&2\\3&1\end{vmatrix}, -\begin{vmatrix}0&1\\3&1\end{vmatrix}$ *i.e.*, are $1, -6, 3$ respectively.

The cofactors of the elements of the third row of $|\mathbf{A}|$ are

$\begin{vmatrix}1&2\\2&3\end{vmatrix}, -\begin{vmatrix}0&2\\1&3\end{vmatrix}, \begin{vmatrix}0&1\\1&2\end{vmatrix}$ *i.e.*, are $-1, 2, -1$ respectively.

Therefore the Adj. $\mathbf{A}$= the transpose of the matrix $\mathbf{B}$ where

$$\mathbf{B}=\begin{bmatrix}-1&8&-5\\1&-6&3\\-1&2&-1\end{bmatrix}.\quad\therefore\ \text{Adj. }\mathbf{A}=\begin{bmatrix}-1&1&-1\\8&-6&2\\5&3&-1\end{bmatrix}.$$

Now $\mathbf{A}^{-1}=\dfrac{1}{|\mathbf{A}|}$ Adj. $\mathbf{A}$ and here $|\mathbf{A}|=-2$.

$$\therefore\quad \mathbf{A}^{-1}=-\tfrac{1}{2}\begin{bmatrix}-1&1&-1\\8&-6&2\\-5&3&-1\end{bmatrix}=\begin{bmatrix}\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\-4&3&-1\\\frac{5}{2}&-\frac{3}{2}&\frac{1}{2}\end{bmatrix}.$$

Ex. 6. *Find the inverse of the matrix*

$$\mathbf{A}=\begin{bmatrix}\cos\alpha&-\sin\alpha&0\\\sin\alpha&\cos\alpha&0\\0&0&1\end{bmatrix}.$$

(Meerut 1983)

Solution. We have

$$|\mathbf{A}|=\begin{vmatrix}\cos\alpha&-\sin\alpha&0\\\sin\alpha&\cos\alpha&0\\0&0&1\end{vmatrix}$$

$=1\begin{vmatrix}\cos\alpha&-\sin\alpha\\\sin\alpha&\cos\alpha\end{vmatrix}$, on expanding the determinant along the third row

$=\cos^2\alpha+\sin^2\alpha=1.$

Since $|\mathbf{A}|\neq 0$, therefore $\mathbf{A}^{-1}$ exists.

Now the cofactors of the elements of the first row of $|\mathbf{A}|$ are

$\begin{vmatrix}\cos\alpha&0\\0&1\end{vmatrix}, -\begin{vmatrix}\sin\alpha&0\\0&1\end{vmatrix}, \begin{vmatrix}\sin\alpha&\cos\alpha\\0&0\end{vmatrix}$

i.e., are $\cos\alpha$, $-\sin\alpha$, 0 respectively.

The cofactors of the elements of the second row of $|\mathbf{A}|$ are

$$-\begin{vmatrix} -\sin\alpha & 0 \\ 0 & 1 \end{vmatrix}, \begin{vmatrix} \cos\alpha & 0 \\ 0 & 1 \end{vmatrix}, -\begin{vmatrix} \cos\alpha & -\sin\alpha \\ 0 & 0 \end{vmatrix}$$

i.e., are $\sin\alpha$, $\cos\alpha$, 0 respectively.

The cofactors of the elements of the third row of $|\mathbf{A}|$ are

$$\begin{vmatrix} -\sin\alpha & 0 \\ \cos\alpha & 0 \end{vmatrix}, -\begin{vmatrix} \cos\alpha & 0 \\ \sin\alpha & 0 \end{vmatrix}, \begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{vmatrix}$$

i.e., are 0, 0, 1 respectively.

Therefore the Adj. $\mathbf{A}$ = the transpose of the matrix

$$\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\therefore \quad \text{Adj. } \mathbf{A} = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now $\mathbf{A}^{-1} = \dfrac{1}{|\mathbf{A}|}$ Adj $\mathbf{A}$ and here $|\mathbf{A}| = 1$.

$$\therefore \quad \mathbf{A}^{-1} = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Ex. 7. *Find the inverse of the matrix*

$$\mathbf{A} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

Also verify your result. **(Meerut 1988)**

Solution We have

$$|\mathbf{A}| = \begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{vmatrix} = \cos^2\alpha + \sin^2\alpha = 1.$$

Since $|\mathbf{A}| \neq 0$, therefore $\mathbf{A}^{-1}$ exists. The cofactors of the elements of the first row of $|\mathbf{A}|$ are $\cos\alpha$, $-\sin\alpha$ respectively.

The cofactors of the elements of the second row of $|\mathbf{A}|$ are $-(-\sin\alpha)$, $\cos\alpha$ *i.e.*, are $\sin\alpha$, $\cos\alpha$ respectively.

Therefore the Adj $\mathbf{A}$ = the transpose of the matrix

$$\begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

$$\therefore \quad \text{Adj } \mathbf{A} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Now $\mathbf{A}^{-1} = \dfrac{1}{|\mathbf{A}|}$ Adj $\mathbf{A}$ and here $|\mathbf{A}| = 1$.

$$\therefore \quad \mathbf{A}^{-1} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}.$$

Verification. We have

$$\mathbf{AA}^{-1} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \times \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\alpha + \sin^2\alpha & \cos\alpha\sin\alpha - \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha - \cos\alpha\sin\alpha & \sin^2\alpha + \cos^2\alpha \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

Also $$\mathbf{A}^{-1}\mathbf{A} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \times \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\alpha + \sin^2\alpha & -\cos\alpha\sin\alpha + \sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha + \cos\alpha\sin\alpha & \sin^2\alpha + \cos^2\alpha \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2.$$

Thus $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}_2$.

Ex. 8. *Given that* $\mathbf{A} = \begin{bmatrix} 1 & -2 & -1 \\ 2 & 3 & 1 \\ 0 & 5 & -2 \end{bmatrix}$, *compute*

(i) det. $\mathbf{A}$, *(ii) Adj* $\mathbf{A}$, *(iii)* $\mathbf{A}^{-1}$. **(Meerut 1980)**

Solution. We have

$$|\mathbf{A}| = \begin{vmatrix} 1 & -2 & -1 \\ 2 & 3 & 1 \\ 0 & 5 & -2 \end{vmatrix} = \begin{vmatrix} 1 & -2 & -1 \\ 0 & 7 & 3 \\ 0 & 5 & -2 \end{vmatrix}, \quad \text{applying } R_2 \to R_2 - 2R_1$$

$$= 1\,(-14-15) = -29.$$

Since $|\mathbf{A}| \neq 0$, therefore $\mathbf{A}^{-1}$ exists.

Now the cofactors of the elements of the first row of $|\mathbf{A}|$ are

$$\begin{vmatrix} 3 & 1 \\ 5 & -2 \end{vmatrix}, -\begin{vmatrix} 2 & 1 \\ 0 & -2 \end{vmatrix}, \begin{vmatrix} 2 & 3 \\ 0 & 5 \end{vmatrix}$$ *i.e.* are -11, 4, 10 respectively.

The cofactors of the elements of the second row of $|\mathbf{A}|$ are

$$-\begin{vmatrix} -2 & -1 \\ 5 & -2 \end{vmatrix}, \begin{vmatrix} 1 & -1 \\ 0 & -2 \end{vmatrix}, -\begin{vmatrix} 1 & -2 \\ 0 & 5 \end{vmatrix}$$ *i.e.* are -9, -2, -5 respectively.

The cofactors of the elements of the third row of $|\mathbf{A}|$ are

$$\begin{vmatrix} -2 & -1 \\ 3 & 1 \end{vmatrix}, -\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}, \begin{vmatrix} 1 & -2 \\ 2 & 3 \end{vmatrix}$$ *i.e.* are 1, -3, 7 respectively.

Therefore Adj $\mathbf{A}$ = the transpose of the matrix $\mathbf{B}$ formed by replacing each element in $\mathbf{A}$ by its cofactor.

Now $B=\begin{bmatrix}-11 & 4 & 10\\ -9 & -2 & -5\\ 1 & -3 & 7\end{bmatrix}$.

$\therefore$ Adj $A=\begin{bmatrix}-11 & -9 & 1\\ 4 & -2 & -3\\ 10 & -5 & 7\end{bmatrix}$.

Now $A^{-1}=\dfrac{1}{|A|}$ Adj $A=-\dfrac{1}{29}\begin{bmatrix}-11 & -9 & 1\\ 4 & -2 & -3\\ 10 & -5 & 7\end{bmatrix}$.

Ex. 9. *Find the inverse of the matrix*

$$S=\begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{bmatrix}$$

and show that SAS^{-1} *is a diagonal matrix where*

$$A=\tfrac{1}{2}\begin{bmatrix}b+c & c-a & b-a\\ c-b & c+a & a-b\\ b-c & a-c & a+b\end{bmatrix}.$$

(Allahabad 1965; Agra 77; Lucknow 69)

Solution. We have

$$|S|=\begin{vmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{vmatrix}=-1\,(0-1)+1\,(1-0)=2.$$

The cofactors of the elements of the first row of $|S|$ are $-1, 1, 1$ respectively. Those of the second row are $1, -1, 1$ and those of the third row are $1, 1, -1$. Therefore

$$\text{Adj } S=\begin{bmatrix}-1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\end{bmatrix}.$$

Now $S^{-1}=\dfrac{1}{|S|}$ Adj $S=\tfrac{1}{2}\begin{bmatrix}-1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\end{bmatrix}$.

We have $SA=\begin{bmatrix}0 & 1 & 1\\ 1 & 0 & 1\\ 1 & 1 & 0\end{bmatrix}\times\tfrac{1}{2}\begin{bmatrix}b+c & c-a & b-a\\ c-b & c+a & a-b\\ b-c & a-c & a+b\end{bmatrix}$

$=\tfrac{1}{2}\begin{bmatrix}0 & 2a & 2a\\ 2b & 0 & 2b\\ 2c & 2c & 0\end{bmatrix}=\begin{bmatrix}0 & a & a\\ b & 0 & b\\ c & c & 0\end{bmatrix}$.

$\therefore$ $SAS^{-1}=\begin{bmatrix}0 & a & a\\ b & 0 & b\\ c & c & 0\end{bmatrix}\times\tfrac{1}{2}\begin{bmatrix}-1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\end{bmatrix}$

$=\tfrac{1}{2}\begin{bmatrix}2a & 0 & 0\\ 0 & 2b & 0\\ 0 & 0 & 2c\end{bmatrix}=\begin{bmatrix}a & 0 & 0\\ 0 & b & 0\\ 0 & 0 & c\end{bmatrix}$.

Therefore SAS^{-1} is a diagonal matrix.

Ex. 10. *If* **O** *be a zero matrix of order n, show that*

adj. $O=O$.

Solution. The $(i, j)^{th}$ element of adj. O

$=$the cofactor of the $(j, i)^{th}$ element of $|O|$.

But each element of $|O|$ is equal to zero. Therefore the cofactor of each element of $|O|$ is zero. Thus each element of adj O is also equal to zero. Hence adj $O=O$.

Ex. 11. *If* I_n *be a unit matrix of order n, show that*

adj $I_n=I_n$.

Solution. The $(i, j)^{th}$ element of adj I_n

$=$the cofactor of the $(j, i)^{th}$ element of $|I_n|$.

But in $|I_n|$, the cofactors of all the elements which lie along the principal diagonal are equal to 1 and the cofactors of all other elements are equal to zero.

Therefore all the elements along the principal diagonal of adj I_n are equal to 1 and all other elements are equal to zero.

$\therefore$ adj. $I_n=I_n$.

Ex 12. *If* **A** *be a square matrix, then show that*

adj $A'=(adj\ A)'$.

Solution. Let A be a square matrix of order n. Then both adj A' and (adj A)$'$ are square matrices of order n.

Now the $(i, j)^{th}$ element of (adj A)$'$

$=$the $(j, i)^{th}$ element of adj A

$=$the cofactor of the $(i, j)^{th}$ element in $|A|$

$=$the cofactor of the $(j, i)^{th}$ element in $|A'|$

$=$the $(i, j)^{th}$ element of adj. A'.

Hence (adj A)$'$=adj A'.

Ex. 13. *If* **A** *is a symmetric matrix, then prove that adj* **A** *is also symmetric.* **(Punjab Hons. 1971)**

Solution. Let A be a symmetric matrix. Then $A'=A$.

Now (adj A)$'$=adj A'

$=$adj A $\quad [\because A'=A]$.

Since (adj A)$'$=adj A, therefore adj A is a symmetric matrix.

Ex. 14. *If the non-singular matrix A is symmetric, then prove that A^{-1} is also symmetric.*

Solution. Let A be a non-singular symmetric matrix. Then A^{-1} exists and $A'=A$.

We have $(A^{-1})'=(A')^{-1}$

$=A^{-1}$ $[\because A'=A]$.

Since $(A^{-1})'=A^{-1}$, therefore A^{-1} is symmetric.

Ex. 15. *Verify that the adjoint of a diagonal matrix of order 3 is a diagonal matrix.*

Solution. Let $A=\begin{bmatrix} l & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & n \end{bmatrix}$ be a diagonal matrix of order 3.

We have $|A|=\begin{vmatrix} l & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & n \end{vmatrix}$.

The cofactors of all the elements of $|A|$ which do not lie along the principal diagonal are equal to zero.

Also the cofactors of the diagonal elements l, m, n are

$$\begin{vmatrix} m & 0 \\ 0 & n \end{vmatrix}, \begin{vmatrix} l & 0 \\ 0 & n \end{vmatrix}, \begin{vmatrix} l & 0 \\ 0 & m \end{vmatrix},$$

i.e. are mn, ln, lm respectively.

Therefore the adj A = the transpose of the matrix B where

$$B=\begin{bmatrix} mn & 0 & 0 \\ 0 & ln & 0 \\ 0 & 0 & lm \end{bmatrix}. \quad \therefore \text{ adj } A=\begin{bmatrix} mn & 0 & 0 \\ 0 & ln & 0 \\ 0 & 0 & lm \end{bmatrix},$$

which is also a diagonal matrix.

Ex. 16. (*i*) *Show that if A is a non-singular matrix, then* $det\,(A^{-1})=(det\,A)^{-1}$ **(I.A.S. 1973)**

(*ii*) *If B is non-singular, prove that the matrices A and $B^{-1}AB$ have the same determinant, A and B being both square matrices of order n.* **(I.A.S. 1972)**

Solution (i) Since A is a non-singular matrix, therefore $\det A \neq 0$ and A^{-1} exists.

Now $AA^{-1}=I \Rightarrow \det (AA^{-1})=\det I \Rightarrow (\det A)(\det A^{-1})=1$

$\Rightarrow (\det A^{-1})=1/(\det A) \Rightarrow \det (A^{-1})=(\det A)^{-1}$.

(ii) We have $\det (B^{-1}AB)=\det (B^{-1})(\det A)(\det B)$

$=(\det B^{-1})(\det B)(\det A)=(\det B^{-1}B)(\det A)$

$=(\det I)(\det A)=1\,(\det A)=\det A$.

Ex. 17. *If the matrices A and B commute, then A^{-1} and B^{-1} also commute.*

Solution. Since A and B commute, therefore $AB=BA$.

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

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

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

Thus A^{-1} and B^{-1} also commute.

Ex. 18. *If A, B, C be three matrices conformable for multiplication, then* $(ABC)^{-1}=C^{-1}B^{-1}A^{-1}$.

Solution. We have $(ABC)^{-1}=\{A\,(BC)\}^{-1}=(BC)^{-1}A^{-1}$

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

Ex. 19 (a). *Prove that if a matrix A is non-singular, then* $AB=AC$ *implies* $B=C$, *where B and C are square matrices of the same order as A.*

Solution. Since the matrix A is non-singular, therefore A^{-1} exists.

Hence $AB=AC \Rightarrow A^{-1}(AB)=A^{-1}(AC)$

$\Rightarrow (A^{-1}A)B=(A^{-1}A)C \Rightarrow I_n B=I_n C \Rightarrow B=C$.

Ex. 19 (b). *Given* $AB=AC$, *does it follow that* $B=C$? *Can you provide a counter example?* **(I.C.S. 1988)**

Solution. If the matrix A is non-singular, then $AB=AC$ implies $B=C$. But if the matrix A is singular, then $AB=AC$ does not necessarily imply that $B=C$. The following example will make it clear.

Let $A=\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $C=\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$.

We have $AB=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}=AC$, though $B \neq C$.

Ex. 20. *If the product of two non-zero square matrices is a zero matrix, show that both of them must be singular matrices.* **(Delhi 1959; Sagar 66)**

Solution. Let A and B be two non-zero square matrices each of the type $n \times n$. It is given that AB = a null matrix *i.e.*, $AB=O$.

Let $|B|$ be not equal to zero. Then B^{-1} exists. So post-multiplying both sides of $AB=O$ by B^{-1}, we get

$ABB^{-1}=O$ or $AI_n=O$ or $A=O$.

But A is not a zero matrix. Hence $|B|$ must be equal to zero.

Now suppose $|A|$ is not equal to zero. Then A^{-1} exists. So pre-multiplying both sides of $AB=O$ by A^{-1}, we get

$$A^{-1}AB=O \text{ or } I_nB=O \text{ or } B=O.$$

But B is not a null matrix. Hence $|A|$ must be equal to zero.

Ex. 21. *If $|A|=0$, prove that $|adj\ A|=0$.*

Solution. Let A be a square matrix of the type $n\times n$. Then

$$A\,(\text{adj } A)=|A|\,I_n.$$

Since $|A|=0$, therefore we get

$$A\,(\text{adj } A)=\text{a null matrix.}$$

If A is a null matrix, adj A is also a null matrix and so

$$|A|=|\text{adj } A|=0.$$

If A is not a null matrix, then

A adj A=a null matrix holds either if adj A is a null matrix or if adj A is a singular matrix. In either case $|\text{adj } A|=0$.

Hence the result.

Ex. 22. *If A be an $n\times n$ matrix, prove that*

$$|adj\ A|=|A|^{n-1}. \quad \textbf{(Rohilkhand 1991; Agra 88)}$$

Solution. We have $A.\text{Adj } A=|A|\,I_n$.

$\therefore$ $|A \text{ adj } A|=||A|\,I_n|$.

$\therefore$ $|A|.|\text{adj } A|=|A|^n\,|I_n|$.

[Since $|AB|=|A|\,|B|$ and $|kA|=k^n\,|A|$]

$\therefore$ $|A|\,|\text{adj } A|=|A|^n$ $\quad [\because\ |I_n|=1]$

$\therefore$ $|\text{adj } A|=|A|^{n-1}$, if $|A|\neq 0$.

If $|A|=0$, then $|\text{adj } A|=0$. So in this case also

$$|\text{adj } A|=|A|^{n-1}$$

Hence the result.

Remark. From the above example we conclude that if $|A|\neq 0$, then $|\text{adj } A|\neq 0$. *Thus if A is a non-singular matrix, then adj A is also non-singular.* **(Allahabad 1979)**

Ex. 23. *If A is a non-singular matrix, then show that*

$$adj\ adj\ A=|A|^{n-2}A.$$

Solution. We have

$$A\,(\text{adj } A)=|A|\,I_n. \quad \ldots(1)$$

If we take adj A in place of A, then (1) gives

$$(\text{adj } A)\,(\text{adj adj } A)=|\text{adj } A|\,I_n$$

or $\quad (\text{adj } A)\,(\text{adj adj } A)=|A|^{n-1}\,I_n \quad [\because\ |\text{adj } A|=|A|^{n-1}]$

Pre-multiplying both sides of this last relation by A, we get

$$A\,\{(\text{adj } A)\,(\text{adj adj } A)\}=A\,\{|A|^{n-1}\,I_n\}$$

or $\quad (A \text{ adj } A)\,(\text{adj adj } A)=|A|^{n-1}\,(AI_n)$

[$\because$ matrix multiplication is associative]

or $\quad (|A|\,I_n)\,(\text{adj adj } A)=|A|^{n-1}A \quad [\because\ AI_n=A]$

or $\quad |A|\,(I_n \text{ adj adj } A)=|A|^{n-1}A$

or $\quad |A|\,\text{adj adj } A=|A|^{n-1}A. \quad \ldots(2)$

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

So cancelling $|A|$ from both sides of (2), we get

$$\text{adj adj } A=|A|^{n-2}A.$$

Ex. 24. *Let A and B be two square matrices of order n. If $AB=I$, then prove that $BA=I$.*

Solution. We have $AB=I$.

$\therefore$ $|AB|=|I|=1$. $\quad [\because\ |I|=1]$

$\therefore$ $|A|\,|B|=1$. $\quad [\because\ |AB|=|A|.|B|]$

$\therefore$ $|A|\neq 0$. $\quad \therefore\ A^{-1}$ exists.

Now $AB=I$

$\Rightarrow A^{-1}(AB)=A^{-1}I \Rightarrow (A^{-1}A)\,B=A^{-1}$

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

$\Rightarrow BA=A^{-1}A \Rightarrow BA=I.$

Ex. 25. *Are the following statements true ? Give reasons in favour of your answers.*

(i) *A, B are n-rowed square matrices such that $AB=O$ and B is non-singular. Then $A=O$.*

(ii) *Only a square, non-singular matrix possesses inverse which is unique.* **(Gujrat 1970)**

Solution. (i) Yes.

We have $AB=O$

$\Rightarrow (AB)\,B^{-1}=O\,B^{-1}$

$\Rightarrow A\,(BB^{-1})=O \Rightarrow AI=O$

$\Rightarrow A=O.$

(ii) Yes. For complete solution refer § 3 of this chapter.

Ex. 26. *If A and B are square matrices of the same order, then*

$$adj\ (AB)=adj\ B.\ adj\ A.$$

(Rohilkhand 1977, Indore 72, Nagarjuna 78)

Solution. We have

$$AB \text{ adj } (AB)=|AB|\,I_n=(\text{adj } AB)\,AB. \quad \ldots(1)$$

Also $\quad AB\,(\text{adj } B.\text{adj } A)=A\,(B \text{ adj } B)\,\text{adj } A$

$=A\,|B|\,I_n \text{ adj } A=|B|\,(A \text{ adj } A)$

$=|B|\,|A|\,I_n=|BA|\,I_n$

$=|AB|\,I_n. \quad \ldots(2)$

Similarly, we have

$(\text{adj } B \text{ adj } A)\,AB=\text{adj } B\,[(\text{adj } A)\,A]\,B$

$=\text{adj } B.|A|\,I_n\,B=|A|.(\text{adj } B)\,B$

$=|A|.|B|\,I_n=|AB|\,I_n. \quad \ldots(3)$

From (1), (2) and (3), the required result follows, provided **AB** is non-singular.

Note. The result adj (**AB**)$=$adj **B** adj **A** holds good even if **A** or **B** is singular. However the proof given above is applicable only if **A** and **B** are non-singular.

Ex. 27. *Find the value of adj* ($\mathbf{P}^{-1}$) *in terms of* **P** *where* **P** *is a non-singular matrix and hence show that*

$$\text{adj}\,(\mathbf{Q}^{-1}\,\mathbf{B}\mathbf{P}^{-1})=\mathbf{PAQ},$$

given that adj $\mathbf{B}=\mathbf{A}$ *and* $|\mathbf{P}|=|\mathbf{Q}|=1$. **(Kanpur 1980)**

Solution. Since **P** is non-singular, therefore it is invertible.

$$\therefore\quad \mathbf{P}^{-1}\mathbf{P}=\mathbf{I}=\mathbf{P}\mathbf{P}^{-1}.$$

Therefore adj $(\mathbf{P}^{-1}\mathbf{P})=$adj $\mathbf{I}=$adj $(\mathbf{P}\mathbf{P}^{-1})$

or $\quad(\text{adj}\,\mathbf{P})\,(\text{adj}\,\mathbf{P}^{-1})=\mathbf{I}=(\text{adj}\,\mathbf{P}^{-1})\,(\text{adj}\,\mathbf{P}).\quad\ldots(1)$

Note that if **S** and **T** are two square matrices of order n, then it can be shown that adj $(\mathbf{ST})=(\text{adj}\,\mathbf{T})\,(\text{adj}\,\mathbf{S})$. Also adj $\mathbf{I}=\mathbf{I}$.

From the relation (1), we see that

$$(\text{adj}\,\mathbf{P}^{-1})=(\text{adj}\,\mathbf{P})^{-1}.$$

Now $\mathbf{P}^{-1}=\dfrac{1}{|\mathbf{P}|}$ adj $\mathbf{P}=$adj $\mathbf{P}$, if $|\mathbf{P}|=1$. Similarly if $|\mathbf{Q}|=1$, then $\mathbf{Q}^{-1}=$adj $\mathbf{Q}$.

We have adj $(\mathbf{Q}^{-1}\,\mathbf{B}\mathbf{P}^{-1})=(\text{adj}\,\mathbf{P}^{-1})\,(\text{adj}\,\mathbf{B})\,(\text{adj}\,\mathbf{Q}^{-1})$
$=(\text{adj}\,\mathbf{P})^{-1}\,\mathbf{A}\,(\text{adj}\,\mathbf{Q})^{-1}=(\mathbf{P}^{-1})^{-1}\,\mathbf{A}\,(\mathbf{Q}^{-1})^{-1}=\mathbf{PAQ}.$

§ 5. Use of the inverse of a matrix to find the solution of a system of linear equations. **(Andhra 1990)**

Consider a system of n linear equations in n unknowns

$$x_1, x_2, \ldots, x_n$$

i.e., $\quad a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n=b_i,\ i=1, 2, \ldots, n.$

These equations can be written in the form of a single matrix equation $\quad\mathbf{AX}=\mathbf{B},\quad\ldots(1)$

where

$$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \ldots & \ldots & \ldots & \ldots\\ \ldots & \ldots & \ldots & \ldots\\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{bmatrix}_{n\times n},\quad \mathbf{X}=\begin{bmatrix} x_1\\ x_2\\ \ldots\\ \ldots\\ x_n\end{bmatrix}_{n\times 1},\quad \mathbf{B}=\begin{bmatrix} b_1\\ b_2\\ \ldots\\ \ldots\\ b_n\end{bmatrix}$$

Suppose **A** is a non-singular matrix

i.e., $\quad|\mathbf{A}|\neq 0.$

Then $\mathbf{A}^{-1}$ exists. Therefore pre-multiplying (1), by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{-1}\,(\mathbf{AX})=\mathbf{A}^{-1}\,\mathbf{B}$$

or $\quad(\mathbf{A}^{-1}\,\mathbf{A})\,\mathbf{X}=\mathbf{A}^{-1}\,\mathbf{B}$

or $\quad\mathbf{I}_n\mathbf{X}=\mathbf{A}^{-1}\,\mathbf{B}$

or $\quad\mathbf{X}=\mathbf{A}^{-1}\,\mathbf{B},$

which gives us the solution of the given equations. Also this solution will be unique as shown below.

Suppose $\mathbf{X}_1$ and $\mathbf{X}_2$ are two solutions of $\mathbf{AX}=\mathbf{B}$.

Then $\quad\mathbf{AX}_1=\mathbf{B}$ and $\mathbf{AX}_2=\mathbf{B}$.

$$\begin{aligned}\therefore\quad \mathbf{AX}_1=\mathbf{AX}_2 &\Rightarrow \mathbf{A}^{-1}\,(\mathbf{AX}_1)=\mathbf{A}^{-1}\,(\mathbf{AX}_2)\\ &\Rightarrow (\mathbf{A}^{-1}\,\mathbf{A})\,\mathbf{X}_1=(\mathbf{A}^{-1}\,\mathbf{A})\,\mathbf{X}_2\\ &\Rightarrow \mathbf{I}_n\mathbf{X}_1=\mathbf{I}_n\mathbf{X}_2 \Rightarrow \mathbf{X}_1=\mathbf{X}_2.\end{aligned}$$

Hence the solution is unique.

Ex. 1. *Write down in matrix form the system of equations*

$$\begin{aligned}2x-y+3z&=9\\ x+y+z&=6\\ x-y+z&=2\end{aligned}$$

and find $\mathbf{A}^{-1}$, *if*

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

and hence solve the given equations. **(Meerut 1980; Kanpur 86)**

Solution. The given system of equations can be written in matrix form as,

$$\mathbf{AX}=\mathbf{B},\quad\ldots(1)$$

where $\quad\mathbf{A}=\begin{bmatrix}2 & -1 & 3\\ 1 & 1 & 1\\ 1 & -1 & 1\end{bmatrix},\ \mathbf{X}=\begin{bmatrix}x\\ y\\ z\end{bmatrix},\ \mathbf{B}=\begin{bmatrix}9\\ 6\\ 2\end{bmatrix}.$

We have $|\mathbf{A}|=2\,(1+1)-1\,(-1+3)+1\,(-1-3)$
$=4-2-4=-2.$

Therefore **A** is non-singular and thus $\mathbf{A}^{-1}$ exists. Let us now find $\mathbf{A}^{-1}$

The cofactors of the elements of the first row of $|\mathbf{A}|$ are

$\begin{vmatrix}1 & 1\\ -1 & 1\end{vmatrix},\ -\begin{vmatrix}1 & 1\\ 1 & 1\end{vmatrix},\ \begin{vmatrix}1 & 1\\ 1 & -1\end{vmatrix}$ *i.e.*, are 2, 0, -2 respectively

The cofactors of the elements of the second row of $|\mathbf{A}|$ are

$-\begin{vmatrix}-1 & 3\\ -1 & 1\end{vmatrix},\ \begin{vmatrix}2 & 3\\ 1 & 1\end{vmatrix},\ -\begin{vmatrix}2 & -1\\ 1 & -1\end{vmatrix}$ *i.e.*, are -2, -1, 1 respectively.

The cofactors of the elements of the third row of $|\mathbf{A}|$ are

$\begin{vmatrix}-1 & 3\\ 1 & 1\end{vmatrix},\ -\begin{vmatrix}2 & 3\\ 1 & 1\end{vmatrix},\ \begin{vmatrix}2 & -1\\ 1 & 1\end{vmatrix}$ *i.e.*, are -4, 1, 3, respectively.

$\therefore$ Adj A = the transpose of the matrix **B** where

$$\mathbf{B}=\begin{bmatrix} 2 & 0 & -2 \\ -2 & -1 & 1 \\ -4 & 1 & 3 \end{bmatrix}. \quad \therefore \quad \text{Adj } \mathbf{A}=\begin{bmatrix} 2 & -2 & -4 \\ 0 & -1 & 1 \\ -2 & 1 & 3 \end{bmatrix}.$$

$$\therefore \quad \mathbf{A}^{-1}=\frac{1}{|\mathbf{A}|} \text{ adj } \mathbf{A}$$

$$=-\tfrac{1}{2}\begin{bmatrix} 2 & -2 & -4 \\ 0 & -1 & 1 \\ -2 & 1 & 3 \end{bmatrix}=\begin{bmatrix} -1 & 1 & 2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}.$$

Now pre-multiplying (1) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{-1}(\mathbf{AX})=\mathbf{A}^{-1}\mathbf{B}$$

or $$(\mathbf{A}^{-1}\mathbf{A})\mathbf{X}=\mathbf{A}^{-1}\mathbf{B}$$

or $$\mathbf{I}_3\mathbf{X}=\mathbf{A}^{-1}\mathbf{B}$$

or $$\mathbf{X}=\mathbf{A}^{-1}\mathbf{B}.$$

We have $$\mathbf{A}^{-1}\mathbf{B}=\begin{bmatrix} -1 & 1 & 2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}\begin{bmatrix} 9 \\ 6 \\ 2 \end{bmatrix}$$

$$=\begin{bmatrix} -1.9+1.6+2.2 \\ 0.9+\frac{1}{2}.6-\frac{1}{2}.2 \\ 1.9-\frac{1}{2}.6-\frac{3}{2}.2 \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

$$\therefore \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

i.e., $x=1$, $y=2$, $z=3$ is the required solution.

Exercises

Find the adjoints of the following matrices :

1. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 3 & -1 & 4 \end{bmatrix}$. **2.** $\begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{bmatrix}$. **(Agra 1979)**

3. Find the inverse of the matrix

$$\mathbf{A}=\begin{bmatrix} a+ib & c+id \\ -c+id & a-ib \end{bmatrix} \text{ if } a^2+b^2+c^2+d^2=1.$$

(Gorakhpur 1964)

4. Find the inverse of the matrix

$$\mathbf{A}=\begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}.$$ **(Marathwada 1971)**

Find the inverse of each of the following matrices :

5. $\begin{bmatrix} 1 & 2 & 5 \\ 2 & 3 & 1 \\ -1 & 1 & 1 \end{bmatrix}$. **(Delhi 1970)**

6. $\begin{bmatrix} 4 & -5 & 6 \\ -1 & 2 & 3 \\ -2 & 4 & 7 \end{bmatrix}$. **(Rajasthan 1970)**

7. $\begin{bmatrix} 2 & 3 & 4 \\ 4 & 3 & 1 \\ 1 & 2 & 4 \end{bmatrix}$. **(Delhi 1981)**

8. $\begin{bmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & -2 \end{bmatrix}$. **(Meerut 1983)**

9. $\begin{bmatrix} 14 & 3 & -2 \\ 6 & 8 & -1 \\ 0 & 2 & -7 \end{bmatrix}$. **(Rajasthan 1969)**

10. $\begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$. **(Kanpur 1969; Meerut 77)**

11. $\begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix}$. **(Kanpur 1981)**

12. $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$. **(Meerut 1979)**

13. $\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}$. **(Meerut 1973)**

14. $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$. **(Meerut 1974)**

15. $\begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix}$.

16. $\begin{bmatrix} 2 & -1 & 4 \\ -3 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix}$. **(Delhi 1970)**

17. Show that

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ -2 & -4 & -5 \end{bmatrix} \text{ is the inverse of } \begin{bmatrix} 3 & -2 & -1 \\ -4 & 1 & -1 \\ 2 & 0 & 1 \end{bmatrix}.$$

(Agra 1972)

18. If $\mathbf{A}^T$ denotes the transpose of a matrix **A**, and

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

find $(\mathbf{A}^T)^{-1}$. **(Meerut 1976)**

19. Let $\mathbf{A}=\begin{bmatrix} -2 & 1 & 3 \\ 0 & -1 & 1 \\ 1 & 2 & 0 \end{bmatrix}$. Show that **A** is non-singular.

Determine adj **A** and $\mathbf{A}^{-1}$. **(I. A. S. 1973)**

20. Find the inverse of the matrix A, where

$$A=\begin{bmatrix}3 & -3 & 4\\ 2 & -3 & 4\\ 0 & -1 & 1\end{bmatrix}. \quad \text{Verify that } A^3=A^{-1}.$$ **(Poona 1970)**

21. Find the inverse of the matrix A and hence solve the equation AX=Y, where

$$A=\begin{bmatrix}1 & 0 & 1\\ -2 & 1 & 0\\ 0 & -1 & 1\end{bmatrix},\ X=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix},\ Y=\begin{bmatrix}-1\\ 1\\ -5\end{bmatrix}.$$

(Kerala 1970)

22. Prove that if A is idempotent and $A \neq I$, then A is singular.

23. If A is a square matrix of order n, prove that

$$|\text{Adj}\ (\text{Adj}\ A)|=|A|^{(n-1)^2}$$ **(Poona 1970)**

24. If A be a square matrix, then show that

$$\text{Adj}\ A^\theta=(\text{Adj}\ A)^\theta.$$

25. If A is a Hermitian matrix, then prove that Adj A is also Hermitian.

Answers

1. $\begin{bmatrix}8 & -5 & -2\\ -4 & -3 & 1\\ -7 & 3 & -1\end{bmatrix}$. **2.** $\begin{bmatrix}1 & 4 & -2\\ -2 & -5 & 4\\ 1 & -2 & 1\end{bmatrix}$.

3. $\begin{bmatrix}a-ib & -c-id\\ c-id & a+ib\end{bmatrix}$. **4.** $\begin{bmatrix}3 & -2\\ -7 & 5\end{bmatrix}$.

5. $\dfrac{1}{21}\begin{bmatrix}2 & 3 & -13\\ -3 & 6 & 9\\ 5 & -3 & -1\end{bmatrix}$. **6.** $\dfrac{1}{3}\begin{bmatrix}2 & 59 & -27\\ 1 & 40 & -18\\ 0 & -6 & 3\end{bmatrix}$.

7. $-\dfrac{1}{5}\begin{bmatrix}10 & -4 & -9\\ -15 & 4 & 14\\ 5 & -1 & -6\end{bmatrix}$. **8.** $-\dfrac{1}{11}\begin{bmatrix}-22 & -4 & -1\\ -55 & -11 & 0\\ 0 & 1 & 3\end{bmatrix}$.

9. $-\dfrac{1}{654}\begin{bmatrix}-54 & 17 & 13\\ 42 & -98 & 2\\ 12 & -28 & 94\end{bmatrix}$. **10.** $\dfrac{1}{5}\begin{bmatrix}2 & 2 & -3\\ -3 & 2 & 2\\ 2 & -3 & 2\end{bmatrix}$.

11. $-\dfrac{1}{4}\begin{bmatrix}1 & -3 & 4\\ -3 & 1 & 0\\ 1 & 1 & -4\end{bmatrix}$. **12.** $\begin{bmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ -1 & 0 & 1\end{bmatrix}$.

13. $\dfrac{1}{10}\begin{bmatrix}-5 & 2 & -16\\ 0 & 2 & 4\\ 5 & 0 & 10\end{bmatrix}$. **14.** $-\dfrac{1}{3}\begin{bmatrix}6 & -5 & 1\\ -15 & 8 & -1\\ 6 & -3 & 0\end{bmatrix}$.

15. $\begin{bmatrix}-9 & 8 & -5\\ 8 & 7 & 4\end{bmatrix}$ **16.** $-\dfrac{1}{}\begin{bmatrix}-1 & 6 & -1\\ 5 & 8 & -14\end{bmatrix}$.

18. $\begin{bmatrix}-9 & -8 & -2\\ 8 & 7 & 2\\ -5 & -4 & -1\end{bmatrix}$.

19. Adj. $A=\begin{bmatrix}-2 & 6 & 4\\ 1 & -3 & 2\\ 1 & 5 & 2\end{bmatrix}$, $A^{-1}=\dfrac{1}{8}$ Adj A.

21. $\dfrac{1}{3}\begin{bmatrix}1 & -1 & -1\\ 2 & 1 & -2\\ 2 & 1 & 1\end{bmatrix}$; $x_1=1,\ x_2=3,\ x_3=-2$.

27. (i) Yes; (ii) Yes.

§ 6. Orthogonal and Unitary Matrices.

Orthogonal Matrix. Definition. *A square matrix A is said to be orthogonal if* $A'A=I$. **(Rohilkhand 1990)**

If A is an orthogonal matrix, then $A'A=I$

$\Rightarrow |A'A|=|I| \Rightarrow |A'|\,.\,|A|=1$ $\quad [\because \det\ (AB)=(\det A).\ (\det B)]$

$\Rightarrow |A|.|A|=1$ $\quad [\because |A'|=|A|]$

$\Rightarrow |A|^2=1 \Rightarrow |A|=\pm 1 \Rightarrow |A| \neq 0$

$\Rightarrow$ A is invertible.

Also then $A'A=I \Rightarrow A'=A^{-1}$

which in turn implies $AA'=I$.

Thus A is an orthogonal matrix if and only if

$$A'A=I=AA'.$$

Theorem. *If A, B be n-rowed orthogonal matrices, AB and BA are also orthogonal matrices.* **(Sagar 1968)**

Proof. Since A and B are both n-rowed square matrices, therefore AB is also an n-rowed square matrix.

Since $|AB|=|A|\ |B|$ and $|A| \neq 0$, also $|B| \neq 0$, therefore $|AB| \neq 0$. Hence AB is a non-singular matrix.

Now $\quad (AB)'=B'A'$.

$\therefore\ (AB)'\ (AB)=(B'A')\ (AB)$

$=B'\ (A'A)\ B$

$=B'IB$ $\quad [\because A'A=I]$

$=B'B$

$=I.$ $\quad [\because B'B=I]$

$\therefore$ AB is orthogonal. Similarly we can prove that BA is also orthogonal.

Unitary Matrix. Definition. A square matrix A is said to be

Since $|A^\theta| = \overline{|A|}$ and $|A^\theta A| = |A^\theta|\,|A|$,

therefore if $A^\theta A = I$, we have $|A|\,\overline{|A|} = 1$.

Thus the determinant of a unitary matrix is of unit modulus.

For a matrix to be unitary, it must be non-singular.

Hence $A^\theta A = I$ implies $AA^\theta = I$.

Theorem. *If* **A**, **B** *be n-rowed unitary matrices,* **AB** *and* **BA** *are also unitary matrices.*

Proof. Since **A** and **B** are both n-rowed square matrices, therefore **AB** is also an n-rowed square matrix.

Since $|AB| = |A|\,|B|$ and $|A| \neq 0$, also $|B| \neq 0$, therefore $|AB| \neq 0$. Hence **AB** is non-singular matrix.

Now $(AB)^\theta = B^\theta A^\theta$.

$$\begin{aligned}
\therefore\ (AB)^\theta (AB) &= (B^\theta A^\theta)(AB) \\
&= B^\theta (A^\theta A) B \\
&= B^\theta I B \qquad [\because\ A^\theta A = I] \\
&= B^\theta B \\
&= I. \qquad [\because\ B^\theta B = I]
\end{aligned}$$

$\therefore$ **AB** is unitary. Similarly we can prove that **BA** is also unitary.

Solved Examples

Ex. 1. *Verify that the matrix*

$$\tfrac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix} \textit{ is orthogonal.}$$

(Sagar 1968)

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

Then $A' = \tfrac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 2 & -2 & -1 \end{bmatrix}$.

We have
$$AA' = \tfrac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 2 & -2 & -1 \end{bmatrix}$$
$$= \tfrac{1}{9}\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

Hence the matrix **A** is orthogonal.

Ex. 2. *Show that the matrix*

$$\tfrac{1}{3}\begin{bmatrix} -1 & 2 & 2 \\ 2 & -1 & 1 \\ 2 & 2 & -1 \end{bmatrix} \textit{ is orthogonal.}$$

(Lucknow 1984)

Ex. 3. *Show that the matrix*

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \textit{ is orthogonal.}$$

(Madras 1983)

Solution. Let $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$.

Then $A' = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

We have
$$AA' = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Hence the matrix **A** is orthogonal.

Ex. 5. *If* (l_r, m_r, n_r), $r = 1, 2, 3$ *be the direction cosines of three mutually perpendicular lines referred to an orthogonal cartesian co-ordinate system, then prove that*

$$\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} \textit{ is an orthogonal matrix.}$$

Solution. Let $A = \begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix}$.

Then $A' = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{bmatrix}$.

We have

$$AA' = \begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix}\begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{bmatrix}$$
$$= \begin{bmatrix} l_1^2+m_1^2+n_1^2 & l_1l_2+m_1m_2+n_1n_2 & l_1l_3+m_1m_3+n_1n_3 \\ l_2l_1+m_2m_1+n_2n_1 & l_2^2+m_2^2+n_2^2 & l_2l_3+m_2m_3+n_2n_3 \\ l_3l_1+m_3m_1+n_3n_1 & l_3l_2+m_3m_2+n_3n_2 & l_3^2+m_3^2+n_3^2 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3. \quad \left[\begin{array}{l} \because\ l_1^2+m_1^2+n_1^2 = 1 \text{ etc.} \\ \text{and } l_1l_2+m_1m_2+n_1n_2 = 0 \text{ etc.} \end{array}\right]$$

Hence the matrix **A** is orthogonal.

Ex. 5. *Show that every 2-rowed real orthogonal matrix is of any one of the forms*

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ or } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

Solution. Let $A=\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$

be any 2-rowed real orthogonal matrix.

We have $A'=\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$

Therefore $A'A=\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}$

$$=\begin{bmatrix} a_1^2+a_2^2 & a_1b_1+a_2b_2 \\ a_1b_1+a_2b_2 & b_1^2+b_2^2 \end{bmatrix}$$

$$=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=I_2.$$

Comparing these, we get

$$a_1^2+a_2^2=1,\ b_1^2+b_2^2=1,\ a_1b_1+a_2b_2=0. \quad \ldots(1)$$

Since a_1, a_2, b_1, b_2 are to be all real, therefore the numerical value of each of them cannot exceed unity. Hence there exist real angle θ and ϕ such that

$$a_1=\cos\theta,\ b_1=\cos\phi,$$

so that
$$a_2=\pm\sin\theta,\ b_2=\pm\sin\phi. \quad \ldots(2)$$

The last of the equations (1), then gives

$$\cos(\phi-\theta)=0 \text{ or } \cos(\phi+\theta)=0$$

according as we take the same or different signs in (2). Considering all the possibilities for the values of a_1, a_2, b_1, b_2 we obtain the following four possible orthogonal matrices :

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \begin{bmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{bmatrix},$$

$$\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}, \begin{bmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

Changing θ to $-\theta$, we see that the first and second matrices respectively coincide with the fourth and third so that we have only two families of orthogonal matrices of order 2 given by

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ or } \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix};$$

θ being the parameter.

Ex. 6. *Show that the Pauli's spin matrices*

$$\sigma_x=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ \sigma_y=\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \sigma_z=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Solution. (i) We have $\sigma_x=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$;

$$\therefore\ (\sigma_x)^\theta=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

$$\therefore\ (\sigma_x)^\theta\,\sigma_x=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=I_2.$$

Hence σ_x is unitary.

(ii) $(\sigma_y)^\theta=\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$.

$$\therefore\ (\sigma_y)^\theta\sigma_y=\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}=\begin{bmatrix} -i^2 & 0 \\ 0 & -i^2 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=I_2.$$

Hence σ_y is unitary.

(iii) $(\sigma_z)^\theta=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

$$\therefore\ (\sigma_z)^\theta\,(\sigma_z)=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=I_2.$$

Hence σ_z is unitary.

Ex. 7. *Prove that the matrix*

$$\begin{bmatrix} \frac{1+i}{2} & \frac{-1+i}{2} \\ \frac{1+i}{2} & \frac{1-i}{2} \end{bmatrix} \textit{ is unitary.}$$

Solution. Let us denote the given matrix by A. Then

$$A'=\begin{bmatrix} \frac{1+i}{2} & \frac{1+i}{2} \\ \frac{-1+i}{2} & \frac{1-i}{2} \end{bmatrix}$$

$$\therefore\ A^\theta=(\overline{A'})=\begin{bmatrix} \frac{1-i}{2} & \frac{1-i}{2} \\ \frac{-1-i}{2} & \frac{1+i}{2} \end{bmatrix}.$$

$$\text{Now } A^\theta A=\begin{bmatrix} \frac{1-i}{2} & \frac{1-i}{2} \\ \frac{-1-i}{2} & \frac{1+i}{2} \end{bmatrix}\begin{bmatrix} \frac{1+i}{2} & \frac{-1+i}{2} \\ \frac{1+i}{2} & \frac{1-i}{2} \end{bmatrix}$$

$$=\begin{bmatrix} \frac{1}{4}(1-i^2)+\frac{1}{4}(1-i^2) & -\frac{1}{4}(1-i)^2+\frac{1}{4}(1-i)^2 \\ -\frac{1}{4}(1+i)^2+\frac{1}{4}(1+i)^2 & \frac{1}{4}(1-i^2)+\frac{1}{4}(1-i^2) \end{bmatrix}$$

$$=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I.$$

$\therefore$ A is unitary.

Ex. 8. *Verify that the matrix*

$$\begin{bmatrix}0 & 0 & 0 & -1\\ -1 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\\ 0 & 0 & -1 & 0\end{bmatrix} \textit{ is orthogonal.}$$

(Lucknow 1971)

Ex. 9. *Show that if* A *is an orthogonal matrix, then* A′ *and* A^{-1} *are also orthogonal.* **(Lucknow 1979; Sagar 64)**

Solution. A is orthogonal $\Rightarrow A'A=I$

$\Rightarrow (A'A)'=I' \Rightarrow A'(A')'=I$

$\Rightarrow A'$ is orthogonal.

Again A is orthogonal $\Rightarrow A'A=I \Rightarrow (A'A)^{-1}=I^{-1}$

$\Rightarrow A^{-1}(A')^{-1}=I$

$\Rightarrow A^{-1}(A^{-1})'=I \quad [\because (A')^{-1}=(A^{-1})']$

$\Rightarrow A^{-1}$ is orthogonal.

Note. A unitary matrix over the field of real numbers is orthogonal *i.e.* a real unitary matrix is an orthogonal matrix.

§ 7. Partitioning of Matrices. (Sagar 1965)

A matrix may be subdivided into sub-matrices by drawing lines parallel to its rows and columns. It is sometimes found very useful to consider these sub-matrices as the elements of the original matrix. Thus a matrix may be regarded as constituted of elements which are themselves matrices of different sizes. Consider the matrix

$$A=\left[\begin{array}{ccc:cc}1 & 2 & 2 & 4 & 5\\ 2 & 3 & 9 & 5 & 1\\ \hdashline 3 & 4 & 11 & 8 & 7\\ 0 & 1 & 13 & 5 & 4\end{array}\right].$$

Let $A_{11}=\begin{bmatrix}1 & 2 & 2\\ 2 & 3 & 9\end{bmatrix}$, $A_{12}=\begin{bmatrix}4 & 5\\ 5 & 1\end{bmatrix}$,

$A_{21}=\begin{bmatrix}3 & 4 & 11\\ 0 & 1 & 13\end{bmatrix}$, $A_{22}=\begin{bmatrix}8 & 7\\ 5 & 4\end{bmatrix}$.

Then we may write

$$A=\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}.$$

Thus the matrix A has been partitioned. The dotted lines indicate the partitions. The elements $A_{11}, A_{12}, A_{21}, A_{22}$ are themselves matrices which are the sub-matrices of A.

A matrix may be partitioned in several ways. Two of the more useful partitionings of a matrix are obtained by considering it as constituted by its rows and columns as sub-matrix elements. Thus if A be an $m\times n$ matrix, we may write

$$A=\begin{bmatrix}R_1\\ R_2\\ \vdots\\ R_m\end{bmatrix}, \quad A=[C_1, C_2, \ldots, C_n].$$

If $A=\begin{bmatrix}P & Q\\ R & S\end{bmatrix}$, then it can be easily verified that

$$A'=\begin{bmatrix}P' & R'\\ Q' & S'\end{bmatrix}.$$

Matrices partitioned identically for addition. Two matrices A and B of the same size are said to have been partitioned *identically* if when expressed as matrices of matrices, they are of the same size and the corresponding elements are also matrices of the same size.

For example, the matrices

$$A=\left[\begin{array}{cc:c}0 & 1 & 2\\ \hdashline 2 & 3 & 4\\ 4 & 5 & 6\\ \hdashline 5 & 6 & 7\end{array}\right] \text{ and } B=\left[\begin{array}{cc:c}0 & 7 & 9\\ \hdashline 4 & 8 & 8\\ 18 & 11 & 10\\ \hdashline 10 & 13 & 12\end{array}\right]$$

are identically partitioned.

If A and B be two matrices of the same size identically partitioned as

$$A=\begin{bmatrix}A_{11} & A_{12} & \ldots A_{1s}\\ A_{21} & A_{22} & \ldots A_{2s}\\ \ldots & \ldots & \ldots\\ A_{t1} & A_{t2} & \ldots A_{ts}\end{bmatrix}, \quad B=\begin{bmatrix}B_{11} & B_{12} & \ldots B_{1s}\\ B_{21} & B_{22} & \ldots B_{2s}\\ \ldots & \ldots & \ldots\\ B_{t1} & B_{t2} & \ldots B_{ts}\end{bmatrix}$$

then it can be easily verified that

$$A+B=\begin{bmatrix}A_{11}+B_{11} & A_{12}+B_{12} & \ldots & A_{1s}+B_{1s}\\ A_{21}+B_{21} & A_{22}+B_{22} & \ldots & A_{2s}+B_{2s}\\ \ldots & \ldots & \ldots & \ldots\\ A_{t1}+B_{t1} & A_{t2}+B_{t2} & \ldots & A_{ts}+B_{ts}\end{bmatrix}.$$

Thus if two matrices A and B of the same size are partitioned identically, it is possible to add the two matrices in usual manner, as if the sub-matrices are the elements.

Matrices partitioned conformably for multiplication.

Let A and B be $m \times n$ and $n \times p$ matrices respectively so that the product AB exists. Let the matrix A be partitioned in any arbitrary manner. Let the matrix B be now partitioned in such a manner that the partitioning lines drawn parallel to the rows of B are in the same relative position as the partitioning lines drawn parallel to the columns of A. Such a partition is possible as the matrix A has the same number of columns as is the number of rows in the matrix B.

The matrices A and B, partitioned in the manner described above, are said to be conformably partitioned for multiplication, for, with such partitionings, it is possible to multiply the two matrices in the usual manner, as if the sub-matrices are the elements. For example

$$\mathbf{A}=\left[\begin{array}{ccc|c|c} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 4 & 5 & 1 \\ \hline 3 & 4 & 5 & 1 & 0 \\ 1 & 0 & -1 & 8 & 3 \end{array}\right],$$

$$\mathbf{B}=\left[\begin{array}{ccc|cc|c} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 0 \\ 2 & 3 & 4 & 5 & 0 & 1 \\ \hline 3 & 4 & 4 & 0 & 1 & 2 \\ \hline 4 & 5 & 8 & 1 & 2 & 3 \end{array}\right]$$

are partitioned conformably to multiplication.

It must be noted that the partitioning lines drawn parallel to the rows of A have no connection with the partitioning lines drawn parallel to the columns of B.

Let A and B be $m \times n$ and $n \times p$ matrices respectively partitioned conformably for multiplication. Let

$$\mathbf{A}=\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \dots & \mathbf{A}_{1s} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \dots & \mathbf{A}_{2s} \\ \dots & \dots & \dots & \dots \\ \mathbf{A}_{q1} & \mathbf{A}_{q2} & \dots & \mathbf{A}_{qs} \end{bmatrix}, \quad \mathbf{B}=\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \dots & \mathbf{B}_{1t} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \dots & \mathbf{B}_{2t} \\ \dots & \dots & \dots & \dots \\ \mathbf{B}_{s1} & \mathbf{B}_{s2} & \dots & \mathbf{B}_{st} \end{bmatrix}.$$

Partitioning lines drawn parallel to the columns of A are in the same relative position as the partitioning lines drawn parallel to the rows of B. Therefore the number of columns in the matrix A is equal to the number of rows in the matrix B. Also it can be easily seen that

(*i*) $\mathbf{A}_{ik}\,\mathbf{B}_{kj}$ exists for all permissible values of *i*, *j* and *k*.

(*ii*) $\mathbf{A}_{i1}\,\mathbf{B}_{1j}+\mathbf{A}_{i2}\mathbf{B}_{2j}+\dots+\mathbf{A}_{is}\mathbf{B}_{sj}$ exists for all values of *i* and *j*.

(*iii*) If the product AB=C be partitioned according to the row partition of A and the column partition of B so that

$$\mathbf{C}=[\mathbf{C}_{ij}],\ i=1, 2, 3, \dots, q,$$
$$j=1, 2, 3, \dots, t,$$

then $$\mathbf{C}_{ij}=\mathbf{A}_{i1}\,\mathbf{B}_{1j}+\mathbf{A}_{i2}\,\mathbf{B}_{2j}+\dots+\mathbf{A}_{is}\,\mathbf{B}_{sj}.$$

Note. **A useful way of partitioning two matrices A and B conformable to multiplication is to write them as**

$$\mathbf{A}=\begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \dots \\ \mathbf{R}_m \end{bmatrix}, \quad \mathbf{B}=[\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_p],$$

where $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m$ are the row vectors of A and $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_p$ are the column vectors of B. Then

$$\mathbf{AB}=\begin{bmatrix} \mathbf{R}_1\mathbf{C}_1 & \mathbf{R}_1\mathbf{C}_2 & \mathbf{R}_1\mathbf{C}_3 & \dots & \mathbf{R}_1\mathbf{C}_p \\ \mathbf{R}_2\mathbf{C}_1 & \mathbf{R}_2\mathbf{C}_2 & \mathbf{R}_2\mathbf{C}_3 & \dots & \mathbf{R}_2\mathbf{C}_p \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{R}_m\mathbf{C}_1 & \mathbf{R}_m\mathbf{C}_2 & \mathbf{R}_m\mathbf{C}_3 & \dots & \mathbf{R}_m\mathbf{C}_p \end{bmatrix}.$$

Solved Examples

Ex. 1. ***If* P, Q *are non-singular matrices, show that if***

$$\mathbf{A}=\begin{bmatrix} \mathbf{P} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix}, \text{ then } \mathbf{A}^{-1}=\begin{bmatrix} \mathbf{P}^{-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q}^{-1} \end{bmatrix}.$$

(Punjab Hon's 1971)

Solution. **Let the inverse of**

$$\mathbf{A}=\begin{bmatrix} \mathbf{P} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix}$$

partitioned conformably to pre-multiplication by A, be denoted by

$$\begin{bmatrix} \mathbf{M} & \mathbf{R} \\ \mathbf{N} & \mathbf{S} \end{bmatrix}.$$

Then $\begin{bmatrix} P & O \\ O & Q \end{bmatrix}\begin{bmatrix} M & R \\ N & S \end{bmatrix}=\begin{bmatrix} I & O \\ O & I \end{bmatrix}$.

So that

$$PM+ON=I \quad i.e., \quad PM=I,$$
$$PR+OS=O \quad i.e., \quad PR=O,$$
$$OM+QN=O, \quad i.e., \quad QN=O,$$
and
$$OR+QS=I \quad i.e., \quad QS=I.$$

Since P is non-singular and $PR=O$, therefore $R=O$.

Also P is non-singular and $PM=I$, therefore $M=P^{-1}$.

Similarly Q is non-singular. Therefore $QN=O$ implies that $N=O$ and $QS=I$ implies that $S=Q^{-1}$.

Hence $A^{-1}=\begin{bmatrix} P^{-1} & O \\ O & Q^{-1} \end{bmatrix}$.

Ex. 2. ***Find the inverse of*** $\begin{bmatrix} A & B \\ C & O \end{bmatrix}$, *where* B, C *are non-singular.*

Solution. Let the inverse of

$$M=\begin{bmatrix} A & B \\ C & O \end{bmatrix}$$

partitioned conformably to pre-multiplication by M, be denoted by

$$\begin{bmatrix} P & R \\ Q & S \end{bmatrix}.$$

Then $\begin{bmatrix} A & B \\ C & O \end{bmatrix}\begin{bmatrix} P & R \\ Q & S \end{bmatrix}=\begin{bmatrix} I & O \\ O & I \end{bmatrix}$.

So that

$$AP+BQ=I, \quad \ldots(i)$$
$$AR+BS=O, \quad \ldots(ii)$$
$$CP+OQ=O, \quad \ldots(iii)$$
and
$$CR+OS=I. \quad \ldots(iv)$$

Since C is non-singular, therefore $CP=O$ implies that $P=O$ and $CR=I$ implies that $R=C^{-1}$.

When $P=O$, from the equation (i), we get $BQ=I$.

But B is non-singular. Therefore $Q=B^{-1}$.

Now from (ii), we get $BS=-AR$.

Pre-multiplication with B^{-1} gives

$$B^{-1}BS=-B^{-1}AR$$
i.e.,
$$IS=-B^{-1}AR$$
i.e.,
$$S=-B^{-1}AC^{-1}, \text{ since } R=C^{-1}.$$

Hence $\begin{bmatrix} A & B \\ C & O \end{bmatrix}^{-1}=\begin{bmatrix} O & C^{-1} \\ B^{-1} & -B^{-1}AC^{-1} \end{bmatrix}$.

Ex. 3. *If* A, B, C *are non-singular, but not necessarily of the same size, show that*

$$\begin{bmatrix} A & H & G \\ O & B & F \\ O & O & C \end{bmatrix}^{-1}=\begin{bmatrix} A^{-1} & -A^{-1}HB^{-1} & A^{-1}HB^{-1}FC^{-1}-A^{-1}GC^{-1} \\ O & B^{-1} & -B^{-1}FC^{-1} \\ O & O & C^{-1} \end{bmatrix}$$

Solution. Let the inverse of

$$M=\begin{bmatrix} A & H & G \\ O & B & F \\ O & O & C \end{bmatrix}$$

partitioned conformable to pre-multiplication by M, be denoted by

$$\begin{bmatrix} P & S & V \\ Q & T & W \\ R & U & X \end{bmatrix}.$$

Then
$$\begin{bmatrix} A & H & G \\ O & B & F \\ O & O & C \end{bmatrix}\begin{bmatrix} P & S & V \\ Q & T & W \\ R & U & X \end{bmatrix}=\begin{bmatrix} I & O & O \\ O & I & O \\ O & O & I \end{bmatrix}.$$

So that

$$AP+HQ+CR=I \quad \ldots(i)$$
$$AS+HT+GU=O \quad \ldots(ii)$$
$$AV+HW+GX=O \quad \ldots(iii)$$
$$BQ+FR=O \quad \ldots(iv)$$
$$BT+FU=I \quad \ldots(v)$$
$$BW+FX=O \quad \ldots(vi)$$
$$CR=O \quad \ldots(vii)$$
$$CU=O \quad \ldots(viii)$$
$$CX=I \quad \ldots(ix)$$

Since C is non-singular, therefore from (vii) and (viii), we get $R=O$ and $U=O$ and from (ix), we get $X=C^{-1}$.

Putting the value of X in (vi), we get $BW+FC^{-1}=O$

i.e. $\quad BW=-FC^{-1}$.

Since B is non-singular, therefore, we get $W=-B^{-1}FC^{-1}$.

Putting the value of U in (v), we get $BT=I$. Since B is non-singular, therefore this relation gives $T=B^{-1}$.

Putting the value of R in (iv), we get $BQ=O$. Since B is non-singular, therefore this relation gives $Q=O$.

Putting the values of X and W in (iii), we get

$$AV-HB^{-1}FC^{-1}+GC^{-1}=O.$$
$$\therefore \quad AV=HB^{-1}FC^{-1}-GC^{-1}.$$

Since A is non-singular, therefore this relation gives

$$V=A^{-1}HB^{-1}FC^{-1}-A^{-1}GC^{-1}.$$

Putting the values of U and T in (ii), we get

$$AS+HB^{-1}=O \quad i.e. \quad AS=-HB^{-1}$$

$$i.e. \quad S=-A^{-1}HB^{-1}.$$

Finally putting the values of R and Q in (i), we get

$$AP=I \text{ which gives } P=A^{-1}.$$

Hence

$$\begin{bmatrix} A & H & G \\ O & B & F \\ O & O & C \end{bmatrix}^{-1}$$

$$=\begin{bmatrix} A^{-1} & -A^{-1}HB^{-1} & A^{-1}HB^{-1}FC^{-1}-A^{-1}GC^{-1} \\ O & B^{-1} & -B^{-1}FC^{-1} \\ O & O & C^{-1} \end{bmatrix}.$$

---

4 Rank of a Matrix

§ 1. Submatrix of a matrix. Suppose A is any matrix of the type $m \times n$. Then a matrix obtained by leaving some rows and columns from A is called a **submatrix** of A. In particular the matrix A itself is a submatrix of A because it is obtained from A by leaving no rows or columns.

Minors of a Matrix. We know that every square matrix possesses a determinant. If A be an $m \times n$ matrix, then the determinant of every square sub-matrix of A is called a **minor** of the matrix A. If we leave $m-p$ rows and $n-p$ columns from A, we shall get a square submatrix of A of order p. The determinant of this square submatrix is called a p-rowed minor of A.

For example, let

$$A=\begin{bmatrix} 2 & 4 & 1 & 9 & 1 \\ 0 & 5 & 2 & 5 & 2 \\ 1 & 9 & 7 & 3 & 4 \\ 3 & -2 & 8 & 1 & 8 \end{bmatrix}_{4\times 5}.$$

In a determinant the number of rows is equal to the number of columns. Therefore there can be no 5-rowed minor of A.

If we leave any column from A, we shall get a square submatrix of A of order 4. Thus

$$\begin{vmatrix} 2 & 4 & 1 & 9 \\ 0 & 5 & 2 & 5 \\ 1 & 9 & 7 & 3 \\ 3 & -2 & 8 & 1 \end{vmatrix}, \begin{vmatrix} 2 & 4 & 9 & 1 \\ 0 & 5 & 5 & 2 \\ 1 & 9 & 3 & 4 \\ 3 & -2 & 1 & 8 \end{vmatrix}, \text{ etc.}$$

are 4-rowed minors of A.

If we leave two columns and one row from A, we shall get a square submatrix of A of order 3. Thus

$$\begin{vmatrix} 2 & 4 & 1 \\ 0 & 5 & 2 \\ 1 & 9 & 7 \end{vmatrix}, \begin{vmatrix} 4 & 1 & 9 \\ 5 & 2 & 5 \\ 9 & 7 & 3 \end{vmatrix}, \begin{vmatrix} 5 & 2 & 5 \\ 9 & 7 & 3 \\ -2 & 8 & 1 \end{vmatrix}, \text{ etc.}$$

are 3-rowed minors of A.

The numbers 2, 4, 1, 9, 1, 0, 5 etc. are all 1-rowed minors of A. We can also write 2-rowed minors of A.

§ 2. Rank of a matrix. Definition.

(Meerut 1988; Kerala 70, 71; Kanpur 87; Poona 70; Allahabad 76; Gujrat 70; Delhi 81; Jiwaji 69)

A number r is said to be the rank of a matrix A if it possesses the following two properties :

(i) There is at least one square submatrix of A of order r whose determinant is not equal to zero.

(ii) If the matrix A contains any square submatrix of order $r+1$, then the determinant of every square submatrix of A of order $r+1$ should be zero.

In short the rank of a matrix is the order of any highest order non-vanishing minor of th matrix.

We shall denote the rank of a matrix A by the symbol $\rho(A)$.

It is obvious that the rank r of an $(m \times n)$ matrix can at most be equal to the smaller of the numbers m and n, but it may be less.

If there is a matrix A which has at least one non-zero minor of order n and there is no minor of A of order $n+1$, then the rank of A is n. **Thus the rank of every non-singular matrix of order n is n.** The rank of a square matrix A of order n can be less than n if and only if A is singular *i.e.*, $|A|=0$.

Note 1. Since the rank of every non-zero matrix is $\geqslant 1$, *we agree to assign the rank, zero, to every null matrix.*

(Allahabad 1979)

Note 2. Every $(r+1)$-rowed minor of a matrix can be expressed as the sum of its r-rowed minors. Therefore if all the r-rowed minors of a matrix are equal to zero, then obviously all its $(r+1)$-rowed minors will also be equal to zero.

Important. The following two simple results will help us very much in finding the rank of a matrix.

(i) The rank of a matrix is $\leqslant r$, if all $(r+1)$-rowed minors of the matrix vanish.

(ii) The rank of a matrix is $\geqslant r$, if there is at least one r-rowed minor of the matrix which is not equal to zero.

Examples.

(a) Let $A = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ be a unit matrix of order 3.

We have $|A|=1$. Therefore A is a non-singular matrix. Hence rank $A=3$. In particular, the rank of a unit matrix of order n is n.

(b) Let $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Since A is a null matrix, therefore rank $A=0$.

(c) Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 0 & 2 & 2 \end{bmatrix}$.

We have $|A| = 1(6-8)-2(4-6) = 2 \neq 0$. Thus A is a non-singular matrix. Therefore rank $A=3$.

(d) Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$.

We have $|A| = 1(24-25)-2(18-20)+3(15-16)=0$.

Therefore the rank of A is less than 3. Now there is at least one minor of A of order 2, namely $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ which is not equal to zero. Hence rank $A=2$.

(e) Let $A = \begin{bmatrix} 3 & 1 & 2 \\ 6 & 2 & 4 \\ 3 & 1 & 2 \end{bmatrix}$.

We have $|A|=0$, since the first two columns are identical. Also each 2-rowed minor of A is equal to zero. But A is not a null matrix. Hence rank $A=1$.

(f) Let $A = \begin{bmatrix} 2 & 4 & 3 & 2 \\ 3 & 5 & 1 & 4 \end{bmatrix}$. Here we see that there is at least one minor of A of order 2 *i.e.*, $\begin{vmatrix} 2 & 4 \\ 3 & 5 \end{vmatrix}$ which is not equal to zero. Also there is no minor of A of order greater that 2. Hence rank of $A=2$.

Echelon form of a matrix. Definition. *A matrix A is said to be in Echelon form if :*

(i) Every row of A which has all its entries 0 occurs below every row which has a non-zero entry.

(ii) The first non-zero entry in each non-zero row is equal to 1.

(iii) The number of zeros before the first non-zero element in a row is less than the number of such zeros in the next row.

Some authors do not require condition (ii).

Important result. *The rank of a matrix in* **Echelon form** *is equal to the number of non-zero rows of the matrix.*

Example. *Find the rank of the matrix*

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

(Meerut 1984)

Solution. The matrix A has one zero row. We see that it occurs below every non-zero row.

Further the number of zeros before the first non-zero element in the first row is one. The number of zeros before the first non-zero element in the second row is two. Thus the number of zeros before the first non zero element in any row is less than the number of such zeros in the next row.

Thus the matrix A is in Echelon form.

$\therefore$ rank A = the number of non-zero rows of A = 2.

§ 3. Theorem. *The rank of the transpose of a matrix is the same as that of the original matrix.*

(Rohilkhand 1976; Meerut 67; Poona 72)

Proof. Let $A=[a_{ij}]_{m\times n}$ be a matrix of the type $m\times n$ and A′ be the transpose of the matrix A. Then $A'=[b_{ji}]_{n\times m}$, where $b_{ji}=a_{ij}$. Suppose rank of A is r. Then there is at least one square sub-matrix of A of order r whose determinant is not equal to zero. Let R be a square sub-matrix of A of order r such that $|R|\neq 0$. If R′ is the transpose of the matrix R, then obviously R′ is a sub-matrix of A′.

Since the value of a determinant does not change by interchanging the rows and columns, therefore $|R'|=|R|\neq 0$.

Hence rank $A'\geqslant r$.

On the other hand, if A′ contains a square sub-matrix S of order $(r+1)$, then corresponding to S, S′ is a sub-matrix of A of order $(r+1)$. But the rank of A is r. Therefore $|S|=|S'|=0$. Thus A′ cannot contain an $(r+1)$-rowed square sub-matrix with non-zero determinant. Hence rank $A'\leqslant r$.

Now rank $A'\geqslant r$ and rank $A'\leqslant r$ implies rank $A'=r$.

Solved Examples

Ex. 1. *Find the rank of each of the following matrices* :

(*i*) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$, (*ii*) $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{bmatrix}$.

Solution. (i) Let $A=\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$.

We have $|A|=1\,(2-0)-2\,(4-0)+3\,(2-0)$, expanding along the first row

$$=2-8+6=0.$$

But there is at least one minor of order 2 of the matrix A, namely $\begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix}$ which is not equal to zero. Hence the rank A = 2.

(ii) Let $A=\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{bmatrix}$.

Here there is at least one minor of order 2 of the matrix A, namely $\begin{vmatrix} 1 & 3 \\ 2 & 5 \end{vmatrix}$ which is not equal to 0. Also there is no minor of the matrix A of order greater than 2. Hence rank A = 2.

***Ex. 2.** *Show that the rank of a matrix every element of which is unity, is* 1.

Solution. Let A denote a matrix every element of which is unity. All the 2-rowed minors of A obviously vanish. But A is a non-zero matrix. Hence rank A = 1.

Ex. 3. **A** *is a non-zero column and* **B** *a non-zero row matrtx, show that rank* (**AB**) = 1.

Solution. Let $A=\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \cdots \\ a_{m1} \end{bmatrix}$ and $B=[b_{11}\ b_{12}\ b_{13}\ldots b_{1n}]$

be two non-zero column and row matrices respectively.

We have $AB=\begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{11}b_{13} & \ldots a_{11}b_{1n} \\ a_{21}b_{11} & a_{21}b_{12} & a_{21}b_{13} & \ldots a_{21}b_{1n} \\ \cdots & \cdots & \cdots & \cdots\ \cdots \\ a_{m1}b_{11} & a_{m1}b_{12} & a_{m1}b_{13} & \ldots a_{m1}b_{1n} \end{bmatrix}$.

Since A and B are non-zero matrices, therefore the matrix AB will also be non-zero. The matrix AB will have at least one non-zero element obtained by multiplying corresponding non-zero elements of A and B.

All the two-rowed minors of A obviously vanish. But A is a non-zero matrix. Hence rank $A=1$.

Ex. 4. *A is an n-rowed square matrix of rank $(n-1)$, show that Adj $A \neq O$.*

Solution. A is an n-rowed square matrix of rank $(n-1)$. Therefore at least one $(n-1)$-rowed minor of the matrix A is not equal to zero. Now every $(n-1)$-rowed minor of the matrix A is equal in magnitude to the cofactor of some element in $|A|$. Thus at least one element in $|A|$ has its cofactor not equal to zero. Therefore at least one element of the matrix Adj A is not equal to zero.

Hence Adj $A \neq O$.

Ex. 5. *Show that the rank of a matrix is $\geqslant$ the rank of every sub-matrix thereof.*

Solution. Let A_1 be a sub-matrix of the matrix A. Let r be the rank of the matrix A_1. Then the matrix A_1 must have at least one r-rowed minor not equal to zero. Non every r-rowed minor of the matrix A_1 will also be an r-rowed minor of the matrix A. Therefore the matrix A will have at least one r-rowed minor not equal to zero. Hence the rank of the matrix A will be $\geqslant r$.

Ex. 6. *Prove that the points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ are collinear if and only if the rank of the matrix*

$$\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$$

is less than 3.

Solution. Suppose the points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ are collinear and they lie on the line whose equation is

$$ax+by+c=0.$$

Then

$$ax_1+by_1+c=0, \qquad \ldots(\text{i})$$

$$ax_2+by_2+c=0, \qquad \ldots(\text{ii})$$

$$ax_3+by_3+c=0. \qquad \ldots(\text{iii})$$

Eliminating a, b and c between (i), (ii) and (iii), we get

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}=0.$$

Thus the rank of the matrix

$$A=\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$$

is less than 3.

Conversely, if the rank of the matrix A is less than 3, then

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}=0.$$

Therefore the area of the triangle whose vertices are (x_1, y_1), (x_2, y_2), (x_3, y_3) is equal to zero. Hence the three points are collinear.

Ex. 7. *If* $U=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, *find the ranks of U and U^2.*

Solution. The matrix U is in Echelon form.

$\therefore$ rank U = the number of non-zero rows of U = 3.

Now
$$U^2=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}\times\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix U^2 is also in Echelon form.

$\therefore$ The rank U^2 = the number of non-zero rows of $U^2=2$.

Ex. 8. *Show that the rank of each of*

$$\begin{bmatrix} A_1 & O \\ A_2 & O \end{bmatrix}, \begin{bmatrix} A_1 & B_1 \\ O & O \end{bmatrix},$$

is at most r ; A_1 being an $r \times r$ matrix.

Solution. (i) Let $M=\begin{bmatrix} A_1 & O \\ A_2 & O \end{bmatrix}$.

Since A_1 is an $r \times r$ matrix, therefore the matrix A_2 will also have r columns. Now every $(r+1)$-rowed square sub-matrix of the matrix M has at least one column of zeros. Therefore all minors of order $(r+1)$ of the matrix M are zero. Hence the rank of the matrix M is $\leqslant r$.

(ii) Let $M=\begin{bmatrix} A_1 & B_1 \\ O & O \end{bmatrix}$.

Since A_1 is an $r \times r$ matrix, therefore the matrix B_1 has also r rows. Now every $(r+1)$-rowed square sub-matrix of the matrix M has at least one row of zeros. Therefore all minors of order $(r+1)$ of the matrix M are zero. Hence the rank of the matrix M is $\leqslant r$.

Ex. 9. *Show that the rank of a matrix does not alter on affixing any number of additional rows or columns of zeros.*

Solution. Let **A** be a matrix of rank r. Let **M** be the matrix obtained from the matrix **A** by affixing some additional rows and columns of zeros.

Let $\mathbf{M}=\begin{bmatrix}\mathbf{A} & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}$.

Now every $(r+1)$-rowed minor of the matrix **M** will either also be a minor of the matrix **A** or it will have at least one row or one column of zeros. Since the matrix **A** is of rank r, therefore every $(r+1)$-rowed minor of the matrix **A** (if there is any) will be equal to zero. Thus every $(r+1)$-rowed minor of the matrix **M** is also equal to zero. Since the matrix **A** has at least one minor of order r not equal to zero, therefore at least one r-rowed minor of the matrix **M** will also be not equal to zero. Hence the rank of the matrix **M** is also equal to r.

Ex. 10. *Show that no skew-symmetric matrix can be of rank* 1.

Solution. Let $\mathbf{A}=\begin{bmatrix}0 & h & g & l\\ -h & 0 & f & m\\ -g & -f & 0 & n\\ -l & -m & -n & 0\end{bmatrix}$

be an 4×4 skew-symmetric matrix.

If h, g, l, m, n are all equal to zero, the matrix **A** will be of rank zero. If at least one of these elements, say, g is not equal to zero, then at least one 2-rowed minor of the matrix **A** *i.e.* the minor $\begin{vmatrix}0 & g\\ -g & 0\end{vmatrix}$ is not equal to zero as its value is g^2 which is not equal to zero. Therefore the rank of the matrix **A** is $\geqslant 2$.

Thus in either case the rank of the matrix **A** is not equal to one.

The method of proof can be given in the case of a skew-symmetric matrix of any order.

§ 4. Elementary Operations or Elementary transformations of a matrix. **(Sagar 1966; Karnatak 68; Poona 70; Gujrat 71)**

An elementary transformation (or an *E*-transformation) is an operation of any one of the following types :

1. *The interchange of any two rows (or columns).*

2. *The multiplication of the elements of any row (or column) by any non-zero number.*

3. *The addition to the elements of any other row (or column) the corresponding elements of any other row (or column) multiplied by any number.*

An elementary transformation is called a **row transformation** or a **column transformation** according as it applies to rows or columns.

§ 5. Symbols to be employed for the elementary transformations.

The following notation will be used to denote the six elementary transformations :

1. The interchange of i^{th} and j^{th} rows will be denoted by $R_i \leftrightarrow R_j$.

2. The multiplication of the i^{th} row by a non-zero number k will be denoted by $R_i \to kR_i$.

3. The addition of k times the j^{th} row to the i^{th} row will be denoted by $R_i \to R_i + kR_j$.

The corresponding column transformations will be denoted by writing C, in place of R *i.e.*, by $C_i \leftrightarrow C_j$, $C_i \to kC_i$, $C_i \to C_i + kC_j$ respectively.

Important. It is quite obvious that if a matrix **B** is obtained from **A** by an elementary transformation, **A** can also be obtained from **B** by an elementary transformation of the same type.

For example, let

$$\mathbf{A}=\begin{bmatrix}1 & 4 & 2 & 9\\ 2 & 5 & 1 & 3\\ 3 & 7 & 8 & 4\end{bmatrix}_{3\times 4}.$$

The elementary transformation $R_2 \to R_2 + 2R_3$ transforms **A** into a matrix **B**, where

$$\mathbf{B}=\begin{bmatrix}1 & 4 & 2 & 9\\ 8 & 19 & 17 & 11\\ 3 & 7 & 8 & 4\end{bmatrix}_{3\times 4}.$$

Now if we apply the elementary transformation $R_2 \to R_2 - 2R_3$ to the matrix **B**, we see that the matrix **B** transforms to the matrix **A**.

Again suppose we apply the elementary transformation $R_2 \to 3R_2$ to the matrix **A**.

Then **A** transforms into a matrix **C**, where

$$\mathbf{C}=\begin{bmatrix}1 & 4 & 2 & 9\\ 6 & 15 & 3 & 9\\ 3 & 7 & 8 & 4\end{bmatrix}$$

Now if we apply the elementary transformation $R_2\to\frac{1}{3}R_2$ to the matrix **C**, we see that the matrix **C** transforms back to the matrix **A**.

§ 6. Elementary matrices. **(Nagarjuna 1974; Meerut 89)**

Definition. *A matrix obtained from a unit matrix by a single elementary transformation is called an elementary matrix (or E-matrix). For example,*

$$\begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix},\begin{bmatrix}1 & 0 & 0\\ 0 & 4 & 0\\ 0 & 0 & 1\end{bmatrix},\begin{bmatrix}1 & 2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

are the elementary matrices obtained from $\mathbf{I}_3$ by subjecting it to the elementary operations $C_1\leftrightarrow C_3$, $R_2\to 4R_2$, $R_1\to R_1+2R_2$ respectively.

It may be worthwhile to note that an *E*-matrix can be obtained from **I** by subjecting it to a row transformation or a column transformation. We shall use the following symbols to denote elementary matrices of different types :

(i) $\mathbf{E}_{ij}$ will denote the *E*-matrix obtained by interchanging the i^{th} and j^{th} rows of a unit matrix. The students can easily see that the matrices obtained by interchanging the i^{th} and j^{th} rows or the i^{th} and j^{th} columns of a unit matrix are the same. Therefore $\mathbf{E}_{ij}$ will also denote the elementary matrix obtained by interchanging the i^{th} and j^{th} columns of a unit matrix.

(ii) $\mathbf{E}_i(k)$ will denote the *E*-matrix obtained by multiplying the i^{th} row of a unit matrix by a non-zero number k. It can be easily seen that the matrices obtained by multiplying the i^{th} row or the i^{th} column of a unit matrix by k are the same. Therefore $\mathbf{E}_i(k)$ will also denote the elementary matrix obtained by multiplying the i^{th} column of a unit matrix by a non-zero number k.

(iii) $\mathbf{E}_{ij}(m)$ will denote the elementary matrix, obtained by adding to the elements of the i^{th} row of a unit matrix, the products by any number m of the corresponding elements of the j^{th} row. It may be easily seen that the *E*-matrix $\mathbf{E}_{ij}(m)$ can also be obtained by adding to the elements of the j^{th} column of a unit matrix, the products by m of the corresponding elements of the i^{th} column.

It can be easily seen that

$$|\mathbf{E}_{ij}|=-1,\ |\mathbf{E}_i(k)|=k\neq 0,\ |\mathbf{E}_{ij}(m)|=1.$$

Thus all the elementary matrices are non-singular.

Therefore each elementary matrix possesses inverse.

Now it is very interesting to note that the elementary transformations of a matrix can also be obtained by algebraic operations on the same by the corresponding elementary matrices. In this connection we have the following theorem.

§ 7. Theorem. *Every elementary row (column) transformation of a matrix can be obtained by pre-multiplication (post-multiplication) with corresponding elementary matrix.*

(Nagarjuna 1977; Allahabad 65; Kanpur 80; Gujrat 71)

We shall first prove that every elementary row transformation of a product **AB** of two matrices **A** and **B** can be obtained by subjecting the pre-factor **A** to the same elementary row transformation. Similarly every elementary column transformation of a product **AB** of two matrices **A** and **B** can be obtained by subjecting the postfactor **B** to the same elementary column transformation.

Let $\mathbf{A}=[a_{ij}]$ and $\mathbf{B}=[b_{jk}]$ be two $m\times n$ and $n\times p$ matrices respectively so that the product **AB** is defined.

Let $\mathbf{R}_1, \mathbf{R}_2, \mathbf{R}_3, \ldots, \mathbf{R}_m$ denote the row vectors of the matrix **A** and $\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_p$ denote the column vectors of the matrix **B**. We can then write,

$$\mathbf{A}=\begin{bmatrix}\mathbf{R}_1\\ \mathbf{R}_2\\ \mathbf{R}_3\\ \cdots\\ \mathbf{R}_m\end{bmatrix},\ \mathbf{B}=[\mathbf{C}_1\ \mathbf{C}_2\ \mathbf{C}_3\ldots\mathbf{C}_p]$$

$$\therefore\quad \mathbf{AB}=\begin{bmatrix}\mathbf{R}_1\mathbf{C}_1 & \mathbf{R}_1\mathbf{C}_2 & \mathbf{R}_1\mathbf{C}_3\ldots\mathbf{R}_1\mathbf{C}_p\\ \mathbf{R}_2\mathbf{C}_1 & \mathbf{R}_2\mathbf{C}_2 & \mathbf{R}_2\mathbf{C}_3\ldots\mathbf{R}_2\mathbf{C}_p\\ \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots\\ \mathbf{R}_m\mathbf{C}_1 & \mathbf{R}_m\mathbf{C}_2 & \mathbf{R}_m\mathbf{C}_3\ldots\mathbf{R}_m\mathbf{C}_p\end{bmatrix}.$$

Now if σ denotes any elementary row transformation, it is quite obvious from the above representation that $(\sigma\mathbf{A})\,\mathbf{B}=\sigma\,(\mathbf{AB})$. For example, if σ denotes the elementary row transformation $R_1\leftrightarrow R_2$, it is quite obvious that $(\sigma\mathbf{A})\,\mathbf{B}=\sigma\,(\mathbf{AB})$.

Similarly it is quite obvious that if the columns $C_1, C_2, \ldots, C_p$ of **B** be subjected to any elementary column transformation, the columns of **AB** are also subjected to the same elementary transformation. Hence the result.

Now to prove our main theorem, if **A** be an $m \times n$ matrix, we can write $A = I_m A$.

If σ denotes any elementary row transformation, we have

$$\sigma A = \sigma (I_m A) = (\sigma I_m) A = EA,$$

where **E** is the *E*-matrix corresponding to the same row transformation σ.

Similarly, we can write $A = AI_n$.

If σ denotes any elementary column transformation, we have

$$\sigma (A) = \sigma (A I_n) = A \sigma (I_n) = AE_1,$$

where E_1 is the *E*-matrix correspondingto the same column transformation σ.

Example. Let $A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 7 & 1 \\ 3 & 8 & 4 \end{bmatrix}$

The *E*-transformation $R_1 \to R_1 + 2R_3$ transforms **A** into **B**, where

$$B = \begin{bmatrix} 7 & 20 & 10 \\ 2 & 7 & 1 \\ 3 & 8 & 4 \end{bmatrix}$$

Also if we apply the row transformation $R_1 \to R_1 + 2R_3$ to the unit matrix,

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

then the *E*-matrix **E** thus obtained is

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now $EA = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 2 \\ 2 & 7 & 1 \\ 3 & 8 & 4 \end{bmatrix}$

$$= \begin{bmatrix} 1.1+0.2+2.3 & 1.4+0.7+2.8 & 1.2+0.1+2.4 \\ 0.1+1.2+0.3 & 0.4+1.7+0.8 & 0.2+1.1+0.4 \\ 0.1+0.2+1.3 & 0.4+0.7+1.8 & 0.2+0.1+1.4 \end{bmatrix}$$

$$= \begin{bmatrix} 7 & 20 & 10 \\ 2 & 7 & 1 \\ 3 & 8 & 4 \end{bmatrix} = B$$

Similarly we can see that a column transformation of **A** can be affected by post-multiplying **A** with the corresponding elementary matrix.

Non-Singularity and inverses of the Elementary matrices.
(Magadh 1969)

(*i*) *The elementary matrix corresponding to the E-operation $R_i \leftrightarrow R_j$ is its own inverse.*

Let E_{ij} denote the elementary matrix obtained by interchanging the i^{th} and j^{th} rows of a unit matrix.

The interchange of the i^{th} and j^{th} rows of E_{ij} will transform E_{ij} to the unit matrix. But every elementary row transformation of a matrix can be brought about by pre-multiplication with the corresponding elementary matrix. Therefore the row transformation which changes E_{ij} to **I** can be affected on pre-multiplication by E_{ij}.

Thus $E_{ij} E_{ij} = I$ or $(E_{ij})^{-1} = E_{ij}$.

Hence E_{ij} is its own inverse.

Similarly, *we can show that the elementary matrix corresponding to the E-operation $C_i \leftrightarrow C_j$ is its own inverse.*

(*ii*) *The inverse of the E-matrix corresponding to the E-operation $R_i \to kR_i$, $(k \neq 0)$, is the E-matrix corresponding to the E-operation $R_i \to k^{-1} R_i$.*

Let $E_i(k)$ denote the elementary matrix obtained by multiplying the elements of the i^{th} row of a unit matrix **I** by a non-zero number k.

The operation of the multiplication of the i^{th} row of $E_i(k)$, by k^{-1} will transform $E_i(k)$ to the unit matrix **I**. This row transformation of $E_i(k)$ can be effected on pre-multiplication by the corresponding elementary matrix $E_i(k^{-1})$.

Thus $E_i(k^{-1}) E_i(k) = I$ or $\{E_i(k)\}^{-1} = E_i(k^{-1})$.

Similarly, *we can show that the inverse of the E-matrix corresponding to the E-operation $C_i \to kC_i$, $k \neq 0$, is the E-matrix corresponding to the E-operation $C_i \to k^{-1} C_i$.*

(*iii*) *The inverse of the E-matrix corresponding to the E-operation $R_i \to R_i + kR_j$ is the E-matrix corresponding to the E-operation $R_i \to R_i - kR_j$.*

Let $E_{ij}(k)$ denote the elementary matrix obtained by adding to the elements of the i^{th} row of a unit matrix **I**, the products by any number k of the corresponding elements of the j^{th} row of **I**.

If we add to the elements of the j^{th} row of $E_{ij}(k)$, the products by $-k$ of the corresponding elements of its j^{th} row, then this row operation will transform $E_{ij}(k)$ to the unit matrix **I**. Now this row transformation of $E_{ij}(k)$ can be effected on pre-multiplication by the corresponding elementary matrix $E_{ij}(-k)$.

Therefore $\quad E_{ij}(-k)\, E_{ij}(k)=I$

or $\quad \{E_{ij}(k)\}^{-1}=E_{ij}(-k)$.

Similarly, *we can show that the inverse of the E-matrix corresponding to the E-operation $C_i \to C_i + kC_j$ is the E-matrix corresponding to the E-operation $C_i \to C_i - kC_j$*

From the above theorem, we thus conclude that the **inverse of an elementary matrix is also an elementary matrix of the same type.**

*****§ 8. Invariance of rank under elementary transformations.**

Theorem. *Elementary transformations do not change the rank of a matrix.* **(Allahabad 1979; Nagarjuna 77; Andhra 74)**

Proof. Let $A=[a_{ij}]$ be an $m \times n$ matrix of rank r. We shall prove the theorem in three stages.

Case I. *Interchange of a pair of rows does not change the rank.* **(Meerut 1972, 73, 74, 82; Delhi 79)**

Let **B** be the matrix obtained from the matrix **A** by the *E*-transformation $R_p \leftrightarrow R_q$ and let s be the rank of the matrix **B**.

Let B_0 be any $(r+1)$-rowed square sub-matrix of **B**. The $(r+1)$ rows of the sub-matrix B_0 of **B** are also the rows of some uniquely determined sub-matrix A_0 of **A**. The identical rows of A_0 and B_0 may occur in the same relative positions or in different relative positions. Since the interchange of two rows of a determinant changes only the sign, we have

$$|B_0|=|A_0| \text{ or } |B_0|=-|A_0|.$$

The matrix **A** is of rank r. Therefore every $(r+1)$-rowed minor of **A** vanishes *i.e.,* $|A_0|=0$. Hence $|B_0|=0$. Thus we see that every $(r+1)$-rowed minor of **A** also vanishes. Therefore, s (the rank of **B**) cannot exceed r (the rank of **A**).

$$\therefore \quad s \leqslant r.$$

Again, as **A** can also be obtained from **B** by an interchange of rows we have $r \leqslant s$.

Hence $\quad r=s$.

Case II. *Multiplication of the elements of a row by a non-zero number does not change the rank.* **(Meerut 1975, 76, 80, 81)**

Let **B** be the matrix obtained from the matrix **A** by the *E*-transformation $R_p \to kR_p$, $(k \neq 0)$, and let s be the rank of the matrix **B**.

Now if $|B_0|$ be any $(r+1)$-rowed minor of **B**, there exists a uniquely determined minor $|A_0|$ of **A** such that

$|B_0|=|A_0|$ (this happens if the p^{th} row of **B** is one of those rows which are struck off to obtain B_0 from **B**),

or $|B_0|=k|A_0|$ (this happens when p^{th} row is retained while obtaining B_0 from **B**).

Since the matrix **A** is of rank r, therefore every $(r+1)$-rowed minor of **A** vanishes *i.e.,* $|A_0|=0$. Hence $|B_0|=0$. Thus we see that every $(r+1)$-rowed minor of **B** also vanishes. Therefore, s (the rank of **B**) cannot exceed r (the rank of **A**).

$$\therefore \quad s \leqslant r.$$

Also, since **A** can be obtained from **B** by *E*-transformation of the same type *i.e.,* $R_p \to (1/k)\, R_p$, therefore, by interchanging the roles of **A** and **B** we find that

$$r \leqslant s.$$

Thus $\quad r=s$.

Case III. *Addition to the elements of a row the products by any number k of the corresponding elements of any other row does not change the rank.* **(Meerut 1981; Delhi 81)**

Let **B** be the matrix obtained from the matrix **A** by the *E*-transformation $R_p \to R_p + kR_q$, and let s be the rank of the matrix **B**.

Let B_0 be any $(r+1)$-rowed square sub-matrix of **B** and A_0 be the correspondingly placed sub-matrix of **A**.

The transformation $R_p \to R_p + kR_q$ has changed only the p^{th} row of the matrix **A**. Also the value of the determinant does not change if we add to the elements of any row the corresponding elements of any other row multiplied by some number k.

Therefore, if no row of the sub-matrix A_0 is part of the p^{th} row of **A**, or if two rows of A_0 are parts of the p^{th} and q^{th} rows of **A**, then

$$|B_0|=|A_0|.$$

Since the rank of **A** is r, therefore $|A_0|=0$, and consequently $|B_0|=0$.

Again, if a row of A_0 is a part of the p^{th} row of **A**, but no row is a part of the q^{th} row, then

$$|B_0|=|A_0|+k|C_0|,$$

where C_0 is an $(r+1)$-rowed square matrix which can be obtained from A_0 by replacing the elements of A_0 in the row which corresponds to the p^{th} row of A by the corresponding elements in the q^{th} row of A. Obviously all the $r+1$ rows of the matrix C_0 are exactly the same as the rows of some $(r+1)$-rowed square sub-matrix of A, though arranged in some different order. Therefore $|C_0|$ is ± 1 times some $(r+1)$-rowed minor of A. Since the rank of A is r, therefore, every $(r+1)$-rowed minor of A is zero, so that $|A_0|=0$, $|C_0|=0$, and consequently $|B_0|=0$.

Thus we see that every $(r+1)$-rowed minor of B also vanishes. Hence, s (the rank of B) cannot exceed r (the rank of A).

$$\therefore \quad s \leqslant r$$

Also, since A can be obtained from B by an E-transformation of the same type *i.e.*, $R_p \to R_p - kR_q$, therefore, interchanging the roles of A and B, we have $r \leqslant s$.

Thus $\qquad r=s.$

We have thus shown that rank of a matrix remains unaltered by an E-row transformation. Therefore we can also say that the rank of a matrix remains unaltered by a series of elementary row transformations.

Similarly we can show that the rank of a matrix remains unaltered by a series of elementary column transformations.

Finally, we conclude that the rank of a matrix remains unaltered by a finite chain of elementary operations.

Corollary. We have already proved that every elementary row (column) transformation of a matrix can be affected by pre-multiplication (post multiplication) with the corresponding elementary matrix. Combining this theorem with the theorem just established, we conclude the following important result :

The pre-multiplication or post-multiplication by an elementary matrix, and as such by any series of elementary matrices, does not alter the rank of a matrix.

§ 5. Reduction to Normal Form.

Theorem. *Every $m \times n$ matrix of rank r can be reduced to the form* $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ *by a finite chain of E-operations, where* I_r *is the r-rowed unit matrix.* **(Nagarjuna 1980; Delhi 80; Banaras 68; Sagar 66; Poona 70; Gujrat 70; Punjab Hons. 71)**

Proof. Let $A=[a_{ij}]$ be an $m \times n$ matrix of rank r. If A is a zero matrix, then r is equal to zero and we have nothing to prove. So let us take A as a non-zero matrix.

Since A is a non-zero matrix, therefore A has at least one element different from zero, say $a_{pq}=k\neq 0$.

By interchanging the p^{th} row with the first row and the q^{th} column with the first column respectively, we obtain a matrix B whose leading element is equal to k which is not equal to zero.

Multiplying the elements of the first row of the matrix B by $\frac{1}{k}$, we obtain a matrix C whose leading element is equal to unity.

$$\text{Let } C=\begin{bmatrix} 1 & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & c_{m3} & \dots & c_{mn} \end{bmatrix}.$$

Subtracting suitable multiples of the first column of C from the remaining columns, and suitable multiples of the first row from the remaining rows, we obtain a matrix D in which all elements of the first row and first column except the leading elements are equal to zero.

$$\text{Let } D=\begin{bmatrix} 1 & 0 & 0 & 0\dots 0 \\ 0 & & & \\ 0 & & A_1 & \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{bmatrix},$$

where A_1 is an $(m-1)\times(n-1)$ matrix.

If now, A_1 be a non-zero matrix, we can deal with it as we did with A. If the elementary operations applied to A_1 for this purpose be applied to D, they will not affect the first row and the first column of D. Continuing this process, we shall finally obtain a matrix M, such that

$$M=\begin{bmatrix} I_k & O \\ O & O \end{bmatrix}.$$

The matrix M has the rank k. Since the matrix M has been obtained from the matrix A by elementary transformations and elementary transformations do not alter the rank, therefore we must have $k=r$.

Hence every $m \times n$ matrix of rank r can be reduced to the form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ by a finite chain of elementary transformations.

Note. The above form is usually called the **first canonical form or normal** form of a matrix.

Corollary 1. *The rank of an $m\times n$ matrix* A *is r if and only if (iff) it can be reduced to the form* $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ *by a finite chain of E-operations.*

The condition is necessary. The proof has been given in the above theorem.

The condition is also sufficient. The matrix A has been transformed into the form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ by elementary transformations which do not alter the rank of the matrix. Since the rank of the matrix $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ is r, therefore the rank of the matrix A must also be r.

Corollary 2. *If* A *be an $m\times n$ matrix of rank r, there exist non-singular matrices* P *and* Q *such that*

$$PAQ=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}.$$ **(Andhra 1981; Rohilkhand 90)**

Proof. If A be an $m\times n$ matrix of rank r, it can be transformed into the form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ by elementary operations. Since E-row (column) operations are equivalent to pre-(post)-multiplication by the corresponding elementary matrices, we have the following result :

If A be an $m\times n$ matrix of rank r, there exist E-matrices $P_1, P_2, \ldots, P_s, Q_1, Q_2, \ldots, Q_t$ such that

$$P_s P_{s-1} \ldots P_1 AQ_1 Q_2 \ldots Q_t=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}.$$

Now each elementary matrix is non-singular and the product of non-singular matrices is also non-singular. Therefore if $P=P_s P_{s-1} \ldots P_2 P_1$ and $Q=Q_1 Q_2 \ldots Q_t$, then P and Q are non-singular matrices. Hence

$$PAQ=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}.$$

§ 10. Equivalence of Matrices. Definition. *If* B *be an $m\times n$ matrix obtained from an $m\times n$ matrix* A *by finite number of elementary transformations of* A, *then* A *is called equivalent to* B. Symbolically, we write A$\sim$B, which is read as '*A is equivalent to* B'.

(Agra 1974; Madras 80; Andhra 74)

The following three properties of the relation '$\sim$' **in the set** of all $m\times n$ matrices are quite obvious :

(i) **Reflexivity.** If A is any $m\times n$ matrix, then A$\sim$A. Obviously A can be obtained from A by the elementary transformation

$$R_i\to kR_i, \text{ where } k=1.$$

(ii) **Symmetry.** If A$\sim$B, then B$\sim$A. If B **can be obtained** from A by a finite number of elementary **transformations of A,** then A can also be obtained from B by a finite number of elementray transformations of B.

(iii) **Transitivity.** If A$\sim$B, B$\sim$C, then A$\sim$C.

If B can be obtained from A by a series of elementary transformations of A and C can be obtained from B by a series of elementary transformations of B, then C can also be obtained from A by a series of elementary transformations of A.

Therefore the relation '$\sim$' in the set of all $m\times n$ matrices is **an equivalence relation.**

Solved Examples

Ex. 1. *If* A *and* B *be two equivalent matrices, then show that rank* A$=$*rank* B. **(Meerut 1983; Agra 74)**

Solution. If A$\sim$B, then B can be obtained from A by a finite number of elementary transformations of A. Now the elementary transformations do not change the rank of a matrix.

$\therefore$ If A$\sim$B, then rank A$=$rank B.

Ex. 2. *Show that if* A *and* B *are equivalent matrices, then there exist non-singular matrices* P *and* Q *such that*

$$B=PAQ.$$ **(Nagarjuna 1980)**

Solution. If A$\sim$B, then B can be obtained from A by a finite number of elementary transformations of A. But elementary row (column) transformations are equivalent to pre-(post) multiplication by the corresponding elementary matrices. Therefore if A$\sim$B, there exist E-matrices $P_1, P_2, \ldots, P_s, Q_1, Q_2, \ldots, Q_t$, such that

$$P_s P_{s-1} \ldots P_1 AQ_1Q_2 \ldots Q_t=B.$$

Now each elementary matrix is non-singular and the product of non-singular matrices is also non-singular.

Therefore if $P = P_s P_{s-1} \ldots P_2 P_1$ and $Q = Q_1 Q_2 \ldots Q_t$, then there exist non-singular matrices P and Q, such that

$$PAQ = B.$$

Ex. 3. *Show that if two matrices A and B have the same size and the same rank, they are equivalent.*

(Nagarjuna 1977; Meerut 83)

Solution. Let A and B be two $m \times n$ matrices of the same rank r. Then by § 9, we have

$$A \sim \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \text{ and also } B \sim \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}.$$

By the symmetry of the equivalence relation,

$$B \sim \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \text{ implies } \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \sim B.$$

Now by the transitivity of the equivalence relation

$$A \sim \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \text{ and } \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \sim B \text{ implies } A \sim B.$$

Ex. 4. *Show that the order in which any elementary row and any elementary column transformations can be performed is immaterial.*

Solution. Let A be any $m \times n$ matrix. Let E_1 and E_2 be the elementary matrices corresponding to the row and column transformations of A. Then by the associative law for the multiplication of matrices, we have

$$E_1 (A E_2) = (E_1 A) E_2.$$

Hence the result follows.

Ex. 5. (*i*) *Use elementary transformations to reduce the following matrix* A *to triangular form and hence find rank* A :

$$A = \begin{bmatrix} 5 & 3 & 14 & 4 \\ 0 & 1 & 2 & 1 \\ 1 & -1 & 2 & 0 \end{bmatrix}.$$

(Meerut 1980, 83)

(*ii*) *Find the rank of the matrix*

$$\begin{bmatrix} 8 & 1 & 3 & 6 \\ 0 & 3 & 2 & 2 \\ -8 & -1 & -3 & 4 \end{bmatrix}$$

(Meerut 1984)

Solution. (i) We have the matrix

$$A \sim \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 5 & 3 & 14 & 4 \end{bmatrix} \text{ by } R_1 \leftrightarrow R_3$$

$$\sim \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 8 & 4 & 4 \end{bmatrix} \text{ by } R_3 \to R_3 - 5R_1$$

$$\sim \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & -12 & -4 \end{bmatrix} \text{ by } R_3 \to R_3 - 8R_2.$$

The last equivalent matrix is in Echelon form (or in triangular form). The number of non-zero rows in this matrix is 3. Therefore its rank is 3. Hence rank A=3.

(ii) Let us denote the given matrix by A. To find the rank of A, we shall reduce it to echelon form. Performing the column operation $C_1 \to \frac{1}{8} C_1$, we get

$$A \sim \begin{bmatrix} 1 & 1 & 3 & 6 \\ 0 & 3 & 2 & 2 \\ -1 & -1 & -3 & 4 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 3 & 6 \\ 0 & 3 & 2 & 2 \\ 0 & 0 & 0 & 10 \end{bmatrix} \text{ by } R_3 \to R_3 + R_1.$$

The last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 3. Therefore its rank is 3. Hence rank A=3.

Ex. 6. *Find the rank of the matrix*

$$A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 6 & 8 & 7 & 5 \end{bmatrix}$$

(Sagar 1966)

Solution. We have the matrix

$$A \sim \begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_4 \to R_4 - R_3 - R_2 - R_1$$

or

$$A \sim \begin{bmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & -3 & 2 \\ 0 & -4 & -8 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1.$$

Now *E*-transformations do not change the rank of a matrix. We see that the determinant of the last equivalent matrix is zero. But the leading minor of the third order of this matrix

i.e. $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & -3 \\ 0 & -4 & -8 \end{vmatrix} = -12$ *i.e.* $\neq 0$. Therefore the rank of this matrix is 3. Hence rank A=3.

Important Note. To determine the rank of a matrix, we can reduce it to Echelon form or to normal form. But sometimes we are given such matrices that if we carefully study their rows and columns, then we shall find that some rows or columns are linearly dependent on some of the others. These can be reduced to zeros by E-row or column transformations. Then we try to find some non-vanishing determinant of the highest order in the equivalent matrix. The order of this determinant will determine the rank of the given matrix.

Ex. 7. *Is the matrix* $\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{bmatrix}$ *equivalent to* $\mathbf{I}_3$?

Solution. Let $\mathbf{A}=\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{bmatrix}$. We have

$$|\mathbf{A}|=\begin{vmatrix} 1 & 2 & 1 \\ -1 & 0 & 2 \\ 2 & 1 & -3 \end{vmatrix}=\begin{vmatrix} 1 & 2 & 1 \\ 0 & 2 & 3 \\ 0 & -3 & -5 \end{vmatrix} \quad R_2+R_1,\ R_3-2R_1$$

$$=-10+9=-1 \text{ i.e., } \neq 0.$$

Thus the matrix A is non-singular. Hence it is of rank 3.

Then rank of $\mathbf{I}_3$ is also 3. Since A and $\mathbf{I}_3$ are matrices of the same size and have the same rank, therefore $\mathbf{A} \sim \mathbf{I}_3$.

***Ex. 8.** *Reduce the matrix*

$$\mathbf{A}=\begin{bmatrix} 1 & 2 & 1 & 0 \\ -2 & 4 & 3 & 0 \\ 1 & 0 & 2 & -8 \end{bmatrix} \text{ to canonical (normal) form.}$$

Solution. The given matrix $\mathbf{A} \sim$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 8 & 5 & 0 \\ 1 & -2 & 1 & -8 \end{bmatrix} \text{ performing the } E\text{-column transformations } C_2 \to C_2-2C_1,\ C_3 \to C_3-C_1.$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 8 & 5 & 0 \\ 0 & -2 & 1 & -8 \end{bmatrix} \text{ performing the } E\text{-row transformations } R_2 \to R_2+2R_1,\ R_3 \to R_3-R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & -\frac{1}{4} & 1 & -8 \end{bmatrix} \text{ performing the } E\text{-column transformation } C_2 \to \tfrac{1}{8} C_2.$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{9}{4} & -8 \end{bmatrix} \text{ by } C_3 \to C_3-5C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 9 & -32 \end{bmatrix} \text{ by } R_3 \to 4R_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 9 & -32 \end{bmatrix} \text{ by } R_3 \to R_3+R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -32 \end{bmatrix} \text{ by } C_3 \to \tfrac{1}{9}C_3$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ by } C_4 \to C_4+32C_3.$$

Thus the matrix A is equivalent to the matrix $[\mathbf{I}_3 \quad \mathbf{O}]$. Hence A is of rank 3.

Ex. 9. *Reduce the matrix,*

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

to the normal form $\begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$ *and hence determine its rank.*

(Delhi 1980; Meerut 89)

Solution. We have the matrix $\mathbf{A} \sim$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 4 & 5 & -8 & 14 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{bmatrix} \text{ by } C_2 \to C_2+C_1,\ C_3 \to C_3-2C_1,\ C_4 \to C_4+3C_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & -8 & 14 \\ 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 \end{bmatrix} \text{ by } R_2 \to R_2-4R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 3 & 0 & 4 \\ 0 & 5 & -8 & 14 \end{bmatrix} \text{ by } R_2 \leftrightarrow R_4$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & -2 \\ 0 & 5 & -8 & 4 \end{bmatrix} \text{ by } C_4 \to C_4-2C_2$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&-2\\0&0&-8&4\end{bmatrix} \text{ by } R_3\to R_3-3R_2,\ R_4\to R_4-5R_2$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-2&0\\0&0&4&-8\end{bmatrix} \text{ by } C_3\leftrightarrow C_4$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&-2&1\end{bmatrix} \text{ by } C_3\to-\tfrac{1}{2}C_3,\ C_4\to-\tfrac{1}{8}C_4$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix} \text{ by } R_4\to R_4+2R_3$$

$\sim I_4$. Hence the matrix A is of rank 4.

Ex. 10. *Find the ranks of* A, B, A+B, AB *and* BA *where*

$$A=\begin{bmatrix}1&1&-1\\2&-3&4\\3&-2&3\end{bmatrix},\ B=\begin{bmatrix}-1&-2&-1\\6&12&6\\5&10&5\end{bmatrix}.$$

(Karnatak 1969)

Solution. (i) We have the matrix A$\sim$

$$\begin{bmatrix}1&0&0\\2&-5&6\\3&-5&6\end{bmatrix} \text{ by } C_2\to C_2-C_1,\ C_3\to C_3+C_1$$

$$\sim\begin{bmatrix}1&0&0\\0&-5&6\\0&-5&6\end{bmatrix} \text{ by } R_2\to R_2-2R_1,\ R_3\to R_3-3R_1$$

$$\sim\begin{bmatrix}1&0&0\\0&1&1\\0&1&1\end{bmatrix} \text{ by } C_2\to-\tfrac{1}{5}C_2,\ C_3\to\tfrac{1}{6}C_3$$

$$\sim\begin{bmatrix}1&0&0\\0&1&0\\0&1&0\end{bmatrix} \text{ by } C_3\to C_3-C_2$$

$$\sim\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix} \text{ by } R_3\to R_3-R_2.$$

Hence the matrix A is of rank 2.

(ii) We have the matrix B$\sim$

$$\begin{bmatrix}-1&-2&-1\\1&2&1\\1&2&1\end{bmatrix} \text{ by } R_2\to\tfrac{1}{6}R_2,\ R_3\to\tfrac{1}{5}R_3$$

$$\sim\begin{bmatrix}-1&-1&-1\\1&1&1\\1&1&1\end{bmatrix} \text{ by } C_2\to\tfrac{1}{2}C_2$$

$$\sim\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix} \text{ by } R_1\to-R_1$$

$$\sim\begin{bmatrix}1&0&0\\1&0&0\\1&0&0\end{bmatrix} \text{ by } C_2\to C_2-C_1,\ C_3\to C_3-C_1$$

$$\sim\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix} \text{ by } R_2\to R_2-R_1,\ R_3\to R_3-R_1.$$

Hence the matrix B is of rank 1.

(iii) We have A+B

$$=\begin{bmatrix}1&1&-1\\2&-3&4\\3&-2&3\end{bmatrix}+\begin{bmatrix}-1&-2&-1\\6&12&6\\5&10&5\end{bmatrix}=\begin{bmatrix}0&-1&-2\\8&9&10\\8&8&8\end{bmatrix}$$

$$\sim\begin{bmatrix}0&-1&-2\\8&9&10\\1&1&1\end{bmatrix} \text{ by } R_3\to\tfrac{1}{8}R_3$$

$$\sim\begin{bmatrix}1&1&1\\8&9&10\\0&-1&-2\end{bmatrix} \text{ by } R_1\leftrightarrow R_3$$

$$\sim\begin{bmatrix}1&0&0\\8&1&2\\0&-1&-2\end{bmatrix} \text{ by } C_2\to C_2-C_1,\ C_3\to C_3-C_1$$

$$\sim\begin{bmatrix}1&0&0\\0&1&2\\0&-1&-2\end{bmatrix} \text{ by } R_2\to R_2-8R_1$$

$$\sim\begin{bmatrix}1&0&0\\0&1&0\\0&-1&0\end{bmatrix} \text{ by } C_3\to C_3-2C_2$$

$$\sim\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix} \text{ by } R_3 \to R_3 + R_2.$$

Hence the rank of $\mathbf{A}+\mathbf{B}$ is 2.

(iv) We have

$$\mathbf{AB}=\begin{bmatrix}1 & 1 & -1\\2 & -3 & 4\\3 & -2 & 3\end{bmatrix}\times\begin{bmatrix}-1 & -2 & -1\\6 & 12 & 6\\5 & 10 & 5\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix} \text{ by the row into column rule of multiplication.}$$

Hence the rank of the matrix **AB** is zero.

(v) We have

$$\mathbf{BA}=\begin{bmatrix}-1 & -2 & -1\\6 & 12 & 6\\5 & 10 & 5\end{bmatrix}\times\begin{bmatrix}1 & 1 & -1\\2 & -3 & 4\\3 & -2 & 3\end{bmatrix}$$

$$=\begin{bmatrix}-8 & 7 & -10\\48 & -42 & 60\\40 & -35 & 50\end{bmatrix}.$$

Now $$\mathbf{BA}\sim\begin{bmatrix}1 & 1 & 1\\-6 & -6 & -6\\-5 & -5 & -5\end{bmatrix} \text{ by } C_1\to-\tfrac{1}{8}C_1,\ C_2\to\tfrac{1}{7}C_2,\ C_3\to-\tfrac{1}{10}C_3.$$

$$\sim\begin{bmatrix}1 & 1 & 1\\1 & 1 & 1\\1 & 1 & 1\end{bmatrix} \text{ by } R_2\to-\tfrac{1}{6}R_2,\ R_3\to-\tfrac{1}{5}R_3.$$

Therefore the matrix **BA** is of rank 1.

Ex. 11. *Find the rank of the matrix*

$$A=\begin{bmatrix}6 & 1 & 3 & 8\\4 & 2 & 6 & -1\\10 & 3 & 9 & 7\\16 & 4 & 12 & 15\end{bmatrix}.$$

(Meerut 1983; Kolhapur 72; Vikram 66; Sagar 68)

Solution Sometimes to determine the rank of a matrix we need not reduce it to its normal form. Certain rows or columns can easily be seen to be linearly dependent on some of the others and hence they can be reduced to zeros by *E*-row or column transformations. Then we try to find some non-vanishing determinant of the highest order in the matrix, the order of which determines the rank.

We have the matrix

$$A\sim\begin{bmatrix}6 & 1 & 3 & 8\\4 & 2 & 6 & -1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix} \text{ by } R_3\to R_3-R_2-R_1,\ R_4\to R_4-R_3-R_1.$$

Since $\begin{vmatrix}6 & 1\\4 & 2\end{vmatrix}=8\neq 0$, therefore Rank $(A)=2$.

Ex. 12. *Find the rank of the matrix*

$$A=\begin{bmatrix}1 & 3 & 4 & 3\\3 & 9 & 12 & 9\\-1 & -3 & -4 & -3\end{bmatrix}.$$

(Gorakhpur 1965; Agra 72)

Solution. We have the matrix

$$A\sim\begin{bmatrix}1 & 3 & 4 & 3\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix} \text{ by } R_2\to R_2-3R_1,\ R_3\to R_3+R_1.$$

All the 2-rowed minors of this matrix obviously vanish. But this is a non-zero matrix. Hence rank $(A)=1$.

Note. The students should remember that *E*-transformations do not alter the rank of a matrix.

Ex. 13. *Find the rank of the matrix*

$$A=\begin{bmatrix}4 & 2 & 1 & 3\\6 & 3 & 4 & 7\\2 & 1 & 0 & 1\end{bmatrix}.$$

Solution. We have

$$A\sim\begin{bmatrix}4 & 2 & 1 & 0\\6 & 3 & 4 & 0\\2 & 1 & 0 & 0\end{bmatrix} \text{ by } C_4\to C_4-C_2-C_3$$

$$\sim\begin{bmatrix}0 & 0 & 1 & 0\\-10 & -5 & 4 & 0\\2 & 1 & 0 & 0\end{bmatrix} \text{ by } C_2\to C_2-2C_3,\ C_1\to C_1-4C_3.$$

We see that each minor of order 3 in the last equivalent matrix is equal to zero. But there is a minor of order 2 *i.e.*, $\begin{vmatrix}-5 & 4\\1 & 0\end{vmatrix}$ which is equal to -4 *i.e.*, $\neq 0$. Hence rank $A=2$.

Ex. 14. *Find the rank of the matrix*

$$A=\begin{bmatrix}2 & 3 & -1 & -1\\1 & -1 & -2 & -4\\3 & 1 & 3 & -2\\6 & 3 & 0 & -7\end{bmatrix}.$$

(Nagarjuna 1980; Gorakhpur 85; Andhra 74; Meerut 88)

Solution. We have

$$A \sim \begin{bmatrix} 2 & 3 & -1 & -1 \\ 1 & -1 & -2 & -4 \\ 3 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_4 \to R_4 - R_3 - R_2 - R_1.$$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 2 & 3 & -1 & -1 \\ 3 & 1 & 3 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_1 \leftrightarrow R_2.$$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 0 & 5 & 3 & 7 \\ 0 & 4 & 9 & 10 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - 3R_1.$$

We see that the determinant of the last equivalent matrix is zero. But there is one minor of order 3 *i.e.*,

$$\begin{vmatrix} 1 & -1 & -2 \\ 0 & 5 & 3 \\ 0 & 4 & 9 \end{vmatrix} \text{ which is equal to } 45 - 12 = 33 \text{ i.e., } \neq 0.$$

Therefore the rank of the matrix A is 3.

Ex. 15. *Find the rank of the matrix*

$$A = \begin{bmatrix} 1 & a & b & 0 \\ 0 & c & d & 1 \\ 1 & a & b & 0 \\ 0 & c & d & 1 \end{bmatrix}.$$ **(Agra M.Sc. 1970)**

Solution. We have the matrix

$$A \sim \begin{bmatrix} 1 & a & b & 0 \\ 0 & c & d & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \text{ by } R_3 \to R_3 - R_1,\ R_4 \to R_4 - R_2.$$

This matrix is in Echelon form.

$\therefore$ rank A = the number of non-zero rows in this matrix = 2.

Ex. 16. *Find the rank of the matrix*

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 6 & 10 \end{bmatrix}.$$ **(Agra M.Sc. 1968)**

Solution. We have the matrix

$$A \sim \begin{bmatrix} 1 & 0 & 3 \\ 2 & 0 & 7 \\ 3 & 0 & 10 \end{bmatrix} \text{ by } C_2 \to C_2 - 2C_1.$$

The determinant of this matrix is zero. But it has a non-zero minor of order 2 namely $\begin{vmatrix} 1 & 3 \\ 2 & 7 \end{vmatrix} = 7 - 6 = 1$.

$\therefore$ rank A = 2.

Ex. 17. *Determine the rank of the following matrices :*

(*i*) $\begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 9 \\ 1 & 3 & 4 & 1 \end{bmatrix}$. **(Agra 1970)**

(*ii*) $\begin{bmatrix} 1 & 2 & -1 & 4 \\ 2 & 4 & 3 & 5 \\ -1 & -2 & 6 & -7 \end{bmatrix}$. **(Sagar 1964)**

Solution. (i) Let us denote the given matrix by A. Performing the elementary row operations $R_2 \to R_2 - 3R_1$, $R_3 \to R_3 - R_1$, we see that

$$A \sim \begin{bmatrix} 1 & 3 & 4 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 3 & 4 & 3 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_2 \leftrightarrow R_3.$$

We see that the last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 2. Hence its rank is 2. Therefore rank A = 2.

(ii) Let us denote the given matrix by A. Performing the elementary operations $R_2 \to R_2 - 2R_1$, $R_3 \to R_3 + R_1$, we see that

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

$$\sim \begin{bmatrix} 1 & 2 & -1 & 4 \\ 0 & 0 & 5 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_3 \to R_3 - R_2.$$

The last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 2. Hence its rank is 2. Therefore rank A=2.

Ex. 18. *Determine the rank of the following matrices :*

(*i*) $\begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 3 & 4 & 1 \\ 2 & 3 & 7 & 5 \\ 2 & 5 & 11 & 6 \end{bmatrix}$ **(Kranatak 1971; Kanpur 79)**

(*ii*) $\begin{bmatrix} -2 & -1 & -3 & -1 \\ 1 & 2 & 3 & -1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & -1 \end{bmatrix}$ **(Meerut 1985; Delhi 81)**

Solution. (i) Let us denote the given matrix by A. Performing the elementary operations $R_3 \to R_3 - R_1$, $R_4 \to R_4 - R_1$, we see that

$$A \sim \begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 3 & 4 & 1 \\ 0 & 4 & 4 & 1 \\ 0 & 6 & 8 & 2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 3 & 4 & 1 \\ 0 & 4 & 4 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_4 \to R_4 - 2R_2$$

$$\sim \begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 12 & 16 & 4 \\ 0 & 12 & 12 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_2 \to 4R_2,\ R_3 \to 3R_3$$

$$\sim \begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 12 & 16 & 4 \\ 0 & 0 & -4 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_3 \to R_3 - R_2.$$

The last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 3. Therefore its rank is 3. Hence rank A=3.

(ii) Let us denote the given matrix by A. Performing the elementary operation $R_1 \leftrightarrow R_2$, we see that

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

$$\sim \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 3 & 3 & -3 \\ 0 & -2 & -2 & 2 \\ 0 & 1 & 1 & -1 \end{bmatrix} \text{ by } R_2 \to R_2 + 2R_1,\ R_3 \to R_3 - R_1$$

$$\sim \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \end{bmatrix} \text{ by } R_2 \to \tfrac{1}{3}R_2,\ R_3 \to -\tfrac{1}{2}R_3$$

$$\sim \begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_3 \to R_3 - R_2,\ R_4 \to R_4 - R_2.$$

The last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 2. Therefore rank A=2.

Ex. 19. *Determine the rank of the following matrices :*

(*i*) $\begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 3 & 2 & 2 \\ 2 & 4 & 3 & 4 \\ 3 & 7 & 4 & 6 \end{bmatrix}$ (*ii*) $\begin{bmatrix} 3 & -2 & 0 & -1 \\ 0 & 2 & 2 & 1 \\ 1 & -2 & -3 & 2 \\ 0 & 1 & 2 & 1 \end{bmatrix}$.

(Indore 70; Kanpur 79) **(Gujrat 71; Kanpur 79)**

Solution. (i) Let us denote the given matrix by A. Performing the elementary operations $R_2 \to R_2 - R_1$, $R_3 \to R_3 - 2R_1$, $R_4 \to R_4 - 3R_1$, we see that

$$A \sim \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_4 \to R_4 - R_2.$$

The last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 3. Therefore its rank is 3. Hence rank A=3.

(ii) Let us denote the given matrix by A. Performing the elementary row operation $R_1 \leftrightarrow R_3$, we see that

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

$$\sim \begin{bmatrix} 1 & -2 & -3 & 2 \\ 0 & 2 & 2 & 1 \\ 0 & 4 & 9 & -7 \\ 0 & 1 & 2 & 1 \end{bmatrix} \text{ by } R_3 \to R_3 - 3R_1$$

$$\sim \begin{bmatrix} 1 & -2 & -3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 4 & 9 & -7 \\ 0 & 2 & 2 & 1 \end{bmatrix} \text{ by } R_2 \leftrightarrow R_4$$

$$\sim \begin{bmatrix} 1 & -2 & -3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -11 \\ 0 & 0 & -2 & -1 \end{bmatrix} \text{ by } R_3 \to R_3 - 4R_2,\ R_4 \to R_4 - 2R_2$$

$$\sim \begin{bmatrix} 1 & -2 & -3 & 2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -11 \\ 0 & 0 & 0 & -23 \end{bmatrix} \text{ by } R_4 \to R_4 + 2R_3.$$

The last equivalent matrix is in Echelon form. The number of non-zero rows in this matrix is 4. Therefore its rank is 4. Hence rank $A=4$.

Ex. 20. *Are the following pairs of matrices equivalent?*

(*i*) $$\begin{bmatrix} 2 & -1 & 3 & 4 \\ 0 & 3 & 4 & 1 \\ 2 & 3 & 7 & 5 \\ 2 & 5 & 11 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -5 & 6 \\ 3 & -2 & 1 & 2 \\ 5 & -2 & -9 & 14 \\ 4 & -2 & -4 & 8 \end{bmatrix}$$

(*ii*) $$\begin{bmatrix} 4 & 0 & 2 \\ 3 & 1 & 0 \\ 5 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 3 & 9 & 0 & 2 \\ 7 & -2 & 0 & 1 \\ 8 & 1 & 1 & 5 \end{bmatrix}.$$

Solution. (i) The two matrices are of the same size. If they have the same rank, then they are equivalent otherwise not. It can be seen that the rank of the first matrix is 4 and that of the second matrix is 2. Hence they are not equivalent.

(ii) The two matrices are not of the same size. Therefore they cannot be equivalent.

Ex. 21. *Find the rank of the matrix*

$$A = \begin{bmatrix} 2 & -2 & 0 & 6 \\ 4 & 2 & 0 & 2 \\ 1 & -1 & 0 & 3 \\ 1 & -2 & 1 & 2 \end{bmatrix}$$

by reducing it to normal form. **(Karnatak 1969)**

Solution. Performing the operation $R_1 \to \frac{1}{2} R_1$, $R_2 \to \frac{1}{2} R_2$, we see that

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

$$\sim \begin{bmatrix} 1 & -1 & 0 & 3 \\ 0 & 3 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix} \text{ by } R_2 \to R_2 - 2R_1,\ R_3 \to R_3 - R_1,\ R_4 \to R_4 - R_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \end{bmatrix} \text{ by } C_2 \to C_2 + C_1,\ C_4 \to C_4 - 3C_1$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -5 \end{bmatrix} \text{ by } R_2 \leftrightarrow R_4$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -5 \end{bmatrix} \text{ by } R_2 \to (-1)\, R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & -8 \end{bmatrix} \text{ by } R_4 \to R_4 - 3R_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & -8 \end{bmatrix} \text{ by } C_3 \to C_3 + C_2,\ C_4 \to C_4 - C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & -8 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ by } R_4 \leftrightarrow R_3$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&1\\0&0&0&0\end{bmatrix} \text{ by } C_3\to\tfrac{1}{3}C_3,\ C_4\to-\tfrac{1}{8}C_4,$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&0\end{bmatrix} \text{ by } C_4\to C_4-C_3$$

which is the normal form $\begin{bmatrix}\mathbf{I}_3 & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}$. Hence rank $\mathbf{A}=3$.

Ex. 22. *Find two non-singular matrices* **P** *and* **Q** *such that* **PAQ** *is in the normal form where*

$$\mathbf{A}=\begin{bmatrix}1&1&1\\1&-1&-1\\3&1&1\end{bmatrix}.$$

Also find the rank of the matrix **A**. **(Meerut 1991; Kanpur 87)**

Solution. We write $\mathbf{A}=\mathbf{I}_3\mathbf{A}\mathbf{I}_3$ *i.e.*

$$\begin{bmatrix}1&1&1\\1&-1&-1\\3&1&1\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\mathbf{A}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

Now we go on applying E-operations on the matrix **A** (the left hand member of the above equation) until it is reduced to the normal form. Every E-row operation will also be applied to the pre-factor $\mathbf{I}_3$ (or its transform) of the product on the right hand member of the above equation and every E-column operation to the post factor $\mathbf{I}_3$ (or its transform).

Performing $R_2\to R_2-R_1$, $R_3\to R_3-3R_1$, we get

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

Performing $C_2\to C_2-C_1$, $C_3\to C_3-C_1$, we get

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

Performing $R_2\to-\tfrac{1}{2}R_2$, we get

$$\begin{bmatrix}1&0&0\\0&1&1\\0&-2&-2\end{bmatrix}=\begin{bmatrix}1&0&0\\\tfrac{1}{2}&-\tfrac{1}{2}&0\\-3&0&1\end{bmatrix}\mathbf{A}\begin{bmatrix}1&-1&-1\\0&1&0\\0&0&1\end{bmatrix}$$

Performing $R_3\to R_3+2R_2$, we get

$$\begin{bmatrix}1&0&0\\0&1&1\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\\tfrac{1}{2}&-\tfrac{1}{2}&0\\-2&-1&1\end{bmatrix}\mathbf{A}\begin{bmatrix}1&-1&-1\\0&1&0\\0&0&1\end{bmatrix}.$$

Performing $C_3\to C_3-C_2$, we have

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}=\begin{bmatrix}1&0&0\\\tfrac{1}{2}&-\tfrac{1}{2}&0\\-2&-1&1\end{bmatrix}\mathbf{A}\begin{bmatrix}1&-1&0\\0&1&-1\\0&0&1\end{bmatrix}.$$

$$\therefore\quad \mathbf{PAQ}=\begin{bmatrix}\mathbf{I}_2&\mathbf{O}\\\mathbf{O}&\mathbf{O}\end{bmatrix},$$

where $\mathbf{P}=\begin{bmatrix}1&0&0\\\tfrac{1}{2}&-\tfrac{1}{2}&0\\-2&-1&1\end{bmatrix}$, $\mathbf{Q}=\begin{bmatrix}1&-1&0\\0&1&-1\\0&0&1\end{bmatrix}$.

Since $\mathbf{A}\sim\begin{bmatrix}\mathbf{I}_2&\mathbf{O}\\\mathbf{O}&\mathbf{O}\end{bmatrix}$, therefore rank $\mathbf{A}=2$.

Ex. 23. *Determine non-singular matrices* **P** *and* **Q** *such that* **PAQ** *is in the normal form* $\begin{bmatrix}\mathbf{I}_r&\mathbf{O}\\\mathbf{O}&\mathbf{O}\end{bmatrix}$,

where $\mathbf{A}=\begin{bmatrix}3&2&-1&5\\5&1&4&-2\\1&-4&11&-19\end{bmatrix}$.

Solution. We write $\mathbf{A}=\mathbf{I}_3\mathbf{A}\mathbf{I}_4$ *i.e.*

$$\begin{bmatrix}3&2&-1&5\\5&1&4&-2\\1&-4&11&-19\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\mathbf{A}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

Now we go on applying suitable E-operations on the matrix **A** (the left hand member of the above equation) unitl it is reduced to the normal form. Every E-row operation will also be applied to the pre-factor $\mathbf{I}_3$ (or its transform) of the product on the right hand member of the above equation and every E-column operation to post-factor $\mathbf{I}_4$ (or its transform).

Performing $R_1\leftrightarrow R_3$, we get

$$\begin{bmatrix}1&-4&11&-19\\5&1&4&-2\\3&2&-1&5\end{bmatrix}=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\mathbf{A}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

Performing $C_2 \to C_2 + 4C_1$, $C_3 \to C_3 - 11C_1$, $C_4 \to C_4 + 19C_1$, $R_2 \to R_2 - 5R_1$, $R_3 \to R_3 - 3R_1$, we g

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 21 & -51 & 93 \\ 0 & 14 & -34 & 62 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -5 \\ 1 & 0 & -3 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 4 & -11 & 19 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing $C_2 \to \frac{1}{7} C_2$, $C_3 \to -\frac{1}{17} C_3$, $C_4 \to \frac{1}{31} C_4$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 3 & 3 \\ 0 & 2 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -5 \\ 1 & 0 & -3 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & \frac{4}{7} & \frac{11}{17} & \frac{19}{31} \\ 0 & \frac{1}{7} & 0 & 0 \\ 0 & 0 & -\frac{1}{17} & 0 \\ 0 & 0 & 0 & \frac{1}{31} \end{bmatrix}.$$

Performing $R_2 \to \frac{1}{3} R_2$, $R_3 \to \frac{1}{2} R_3$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \frac{1}{3} & -\frac{5}{3} \\ \frac{1}{2} & 0 & -\frac{3}{2} \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & \frac{4}{7} & \frac{11}{17} & \frac{19}{31} \\ 0 & \frac{1}{7} & 0 & 0 \\ 0 & 0 & -\frac{1}{17} & 0 \\ 0 & 0 & 0 & \frac{1}{31} \end{bmatrix}.$$

Performing $C_3 \to C_3 - C_2$, $C_4 \to C_4 - C_2$, $R_3 \to R_3 - R_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \frac{1}{3} & -\frac{5}{3} \\ \frac{1}{2} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & \frac{4}{7} & \frac{9}{119} & \frac{9}{217} \\ 0 & \frac{1}{7} & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & -\frac{1}{17} & 0 \\ 0 & 0 & 0 & \frac{1}{31} \end{bmatrix}.$$

Thus $\mathbf{PAQ} = \begin{bmatrix} \mathbf{I}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$, where $\mathbf{P} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \frac{1}{3} & -\frac{5}{3} \\ \frac{1}{2} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix}$ and

$$\mathbf{Q} = \begin{bmatrix} 1 & \frac{4}{7} & \frac{9}{119} & \frac{9}{217} \\ 0 & \frac{1}{7} & -\frac{1}{7} & -\frac{1}{7} \\ 0 & 0 & -\frac{1}{17} & 0 \\ 0 & 0 & 0 & \frac{1}{31} \end{bmatrix}.$$

Since the matrix $\mathbf{A} \sim \begin{bmatrix} \mathbf{I}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}$, therefore the rank of A is

Exercises

Find the rank of each of the following matrices :

1. $\begin{bmatrix} 1 & 2 & -4 & 5 \\ 2 & -1 & 3 & 6 \\ 8 & 1 & 9 & 7 \end{bmatrix}$. **2.** $\begin{bmatrix} 1 & -1 & 3 & 6 \\ 1 & 3 & -3 & -4 \\ 5 & 3 & 3 & 11 \end{bmatrix}$.

(Meerut 197

3. $\begin{bmatrix} 8 & 0 & 0 & 1 \\ 1 & 0 & 8 & 1 \\ 0 & 0 & 1 & 8 \\ 0 & 8 & 1 & 8 \end{bmatrix}$. **4.** $\begin{bmatrix} 0 & 1 & -3 & -1 \\ 1 & 0 & 1 & 1 \\ 3 & 1 & 0 & 2 \\ 1 & 1 & -2 & 0 \end{bmatrix}$.

(Meerut 1990; Rohilkhand 81)

5. $\begin{bmatrix} 2 & 3 & 7 \\ 3 & -2 & 4 \\ 1 & -3 & -1 \end{bmatrix}$. **6.** $\begin{bmatrix} 1 & 0 & 2 & 3 \\ 2 & 1 & 0 & 1 \\ 4 & 1 & 4 & 7 \end{bmatrix}$.

(Marathwada 1971) **(Kerala 1970)**

7. $\begin{bmatrix} 2 & 1 & 3 \\ 4 & 7 & 13 \\ 4 & -3 & -1 \end{bmatrix}$. **8.** $\begin{bmatrix} 3 & -1 & 2 \\ -6 & 2 & -4 \\ -3 & 1 & -2 \end{bmatrix}$.

(Jiwaji 1970) **(Jiwaji 1969)**

9. $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 6 & 2 \\ 1 & 2 & 3 & 2 \end{bmatrix}$. **10.** $\begin{bmatrix} 4 & 5 & 6 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{bmatrix}$.

(Kanpur 1981; Agra 80) **(Meerut 1974; Gorakhpur 80)**

11. $\begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 2 & 3 & 4 \\ 3 & 4 & 5 & 2 \end{bmatrix}$. **12.** $\begin{bmatrix} 1 & 3 & 4 & 5 \\ 1 & 2 & 6 & 7 \\ 1 & 5 & 0 & 1 \end{bmatrix}$.

(Kanpur 1982) **(Agra 1974)**

13. $\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & -2 & 1 \\ 1 & -1 & 4 & 0 \\ -2 & 2 & 8 & 0 \end{bmatrix}$. **(Meerut 1975)**

14. Reduce the matrix

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

to normal form and find its rank. **(Kanpur 1983; Agra 88)**

15. Reduce the matrix

$\mathbf{A} = \begin{bmatrix} 9 & 7 & 3 & 6 \\ 5 & -1 & 4 & 1 \\ 6 & 8 & 2 & 4 \end{bmatrix}$ to normal form and find its rank.

(Gujrat 1970)

16. Use elementary row or column operations to find the rank of the matrix

$$\begin{bmatrix} 1 & 2 & -1 & 3 \\ 4 & 1 & 2 & 1 \\ 3 & -1 & 1 & 2 \\ 1 & 2 & 0 & 1 \end{bmatrix}.$$

(Poona 1970)

17. Find two non-singular matrices **P** and **Q** such that **PAQ** is in the normal form where

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

(Sagar 1966)

18. Find matrices **R** and **S** such that

$\mathbf{R}\begin{bmatrix} 2 & 2 & -6 \\ -1 & 2 & 2 \end{bmatrix}\mathbf{S}$ is in the normal form. **(Poona 1970)**

19. Find matrices **P** and **Q** so that **PAQ** is of the normal form where

$$\mathbf{A}=\begin{bmatrix} 1 & 0 & -2 \\ 2 & 3 & -4 \\ 3 & 3 & -6 \end{bmatrix}$$

(Agra 1974)

20. Show that the interchange of a pair of columns does not change the rank of a matrix.

21. Show that the rank of a matrix is not altered if a column of it is multiplied by a non-zero scalar. **(Meerut 1976)**

Answers

1.	3.	**2.**	3.	**3.**	4.	**4.**	2.	**5.**	2.	**6.**	2.
7.	2.	**8.**	1.	**9.**	2.	**10.**	2.	**11.**	2.	**12.**	2.
13.	3.	**14.**	2.	**15.**	3.	**16.**	3.				

§ 11. Row and column equivalence of matrices.

Definition *A matrix* **A** *is said to be row equivalent to* **B** *if* **B** *is obtainable from* **A** *by a finite number of E-row transformations of* **A**. *Symbolically, we then write* $\mathbf{A} \underset{R}{\sim} \mathbf{B}$. *Similarly a matrix* **A** *is said to be column equivalent to* **B** *if* **B** *is obtainable from* **A** *by a finite number of E-column transformations of* **A**. *Symbolically, we then write* $\mathbf{A} \overset{C}{\sim} \mathbf{B}$. **(Nagarjuna 1978)**

§ 12. Employment of only row transformations.

Theorem. *If* **A** *be an* $m \times n$ *matrix of rank* r, *then there exists a non-singular matrix* **P** *such that*

$$\mathbf{PA}=\begin{bmatrix} \mathbf{G} \\ \mathbf{O} \end{bmatrix},$$

where **G** *is an* $r \times n$ *matrix of rank* r *and* **O** *is* $(m-r)\times n$.

Proof. Since **A** is an $m \times n$ matrix of rank r, therefore there exist non-singular matrices **P** and **Q** such that

$$\mathbf{PAQ}=\begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \qquad \ldots\text{(i)}$$

Now every non-singular matrix can be expressed as the product of elementary matrices. So let

$\mathbf{Q}=\mathbf{Q}_1\mathbf{Q}_2\ldots\mathbf{Q}_t$ where $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_t$ are all elementary matrices. Thus the relation (i) can be written as

$$\mathbf{PAQ}_1\mathbf{Q}_2\ldots\mathbf{Q}_t=\begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} \qquad \ldots\text{(ii)}$$

Now every *E*-column transformation of a matrix, is equivalent to post-multiplication with the corresponding elementary matrix. Since no column transformation can affect the last $(m-r)$ rows of the right hand side of (ii), therefore post-multiplying the L.H.S. of (ii) by the elementary matrices $\mathbf{Q}_t^{-1}, \mathbf{Q}_{t-1}^{-1}, \ldots, \mathbf{Q}_2^{-1}, \mathbf{Q}_1^{-1}$ successively and effecting the corresponding column transformations in the right hand side of (ii), we get a relation of the form

$$\mathbf{PA}=\begin{bmatrix} \mathbf{G} \\ \mathbf{O} \end{bmatrix}.$$

Since elementary transformations do not alter the rank, therefore the rank of the matrix **PA** is the same as that of the matrix **A** which is r. Thus the rank of the matrix $\begin{bmatrix} \mathbf{G} \\ \mathbf{O} \end{bmatrix}$ is r and therefore the rank of the matrix **G** is also r as the matrix **G** has r rows and last $m-r$ rows of the matrix $\begin{bmatrix} \mathbf{G} \\ \mathbf{O} \end{bmatrix}$ consist of zeros only.

§ 13. Employment of only column transformations.

Theorem. *If* **A** *be an* $m \times n$ *matrix of rank* r, *then there exists a non-singular matrix* **Q** *such that*

$$\mathbf{AQ}=[\mathbf{H} \quad \mathbf{O}],$$

where **H** *is an* $m \times r$ *matrix of rank* r *and* **O** *is* $m \times (n-r)$.

Proof. Since **A** is an $m \times n$ matrix of rank r, therefore there exist non-singular matrices **P** and **Q** such that

$$\mathbf{PAQ}=\begin{bmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}. \qquad \ldots\text{(i)}$$

Now every non-singular matrix can be expressed as the product of elementary matrices. So let

$P=P_1 P_2 .. P_s$, where $P_1, P_2, \ldots, P_s$ are elementary matrices. Thus the relation (i) can be written as

$$P_1 P_2 . P_s AQ=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}. \qquad \ldots(ii)$$

Now every *E*-row transformation of a matrix is equivalent to pre-multiplication with the corresponding elementary matrix. Again no row transformation can affect the last $n-r$ columns of

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$$

Therefore pre-multiplying the L.H.S. of (ii) by the elementary matrices $P_1^{-1}, P_2^{-1}, \ldots, P_s^{-1}$ successively and effecting the corresponding transformations in the R.H.S. of (ii), we get a relation of the form $AQ=[H \quad O]$.

Now elementary transformations do not alter the rank. Therefore the rank of the matrix AQ is the same as that of A which is r. Thus the rank of the matrix $[H \quad O]$ is r and therefore the rank of the matrix H is also r as the matrix H has r columns and the last $n-r$ columns of the matrix $[H \quad O]$ consist of zeros only.

§ 14. The rank of a product.

Theorem. *The rank of a product of two matrices cannot exceed the rank of either matrix.*

(Nagarjuna 1980; I.C.S. 87; Gujrat 71; Poona 72; Allahabad 78; Andhra 74; Punjab 71)

Let A and B be two $m \times n$ and $n \times p$ matrices respectively. Let r_1, r_2 be the ranks of A and B respectively and let r be the rank of the product AB.

To prove $r \leqslant r_1$ and $r \leqslant r_2$.

Since A is an $m \times n$ matrix of rank r_1, therefore there exists a non-singular matrix P such that

$$PA=\begin{bmatrix} G \\ O \end{bmatrix}$$

where G is an $r_1 \times n$ matrix of rank r_1 and O is $(m-r_1) \times n$.

$$\therefore \quad PAB=\begin{bmatrix} G \\ O \end{bmatrix} B.$$

Since the rank of a matrix does not alter by multiplying it with a non-singular matrix, therefore

Rank (PAB)=Rank (AB)=r.

$\therefore$ Rank of the matrix $\begin{bmatrix} G \\ O \end{bmatrix} B=r$.

Since the matrix G has only r_1 non-zero rows, therefore the matrix $\begin{bmatrix} G \\ O \end{bmatrix}$ B cannot have more than r_1 non-zero rows which arise by multiplying the r_1 non-zero rows of G with the columns of B.

$\therefore$ Rank of the matrix $\begin{bmatrix} G \\ O \end{bmatrix}$ B is $\leqslant r_1$

i.e. $r \leqslant r_1$

i.e. Rank (AB) $\leqslant$ Rank of the prefactor A.

Again r=Rank (AB)=Rank (AB)$'$

=Rank (B$'$A$'$)

$\leqslant$ Rank of the prefactor B$'$ [$\because$ Rank B=Rank B$'$]

=Rank B

$=r_2$.

$\therefore \quad r \leqslant r_2$

i.e. Rank (AB) $\leqslant$ Rank of the post-factor B.

§ 15. Theorem. *Every non-singular matrix is row equivalent to a unit matrix.*

Proof. We shall prove the theorem by induction on n, the order of the matrix.

If the matrix be of order 1 *i.e.* if $A=[a_{11}]$, the theorem obviously holds.

Let us assume that the theorem holds for all non-singular matrices of order $n-1$.

Let $A=[a_{ij}]$ be an $n \times n$ non-singular matrix. The first column of the matrix A has at least one element different from zero, for otherwise we shall have $|A|=0$ and the matrix A will not be non-singular.

Let $a_{p1}=k \neq 0$.

By interchanging the p^{th} row with the first row (if necessary), we obtain a matrix B whose leading element is equal to k which is not equal to zero.

Multiplying the elements of the first row of the matrix B by $1/k$, we obtain a matrix C whose leading element is equal to unity.

$$\text{Let } C=\begin{bmatrix} 1 & c_{12} & c_{13} & \ldots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \ldots & c_{2n} \\ c_{31} & c_{32} & c_{33} & \ldots & c_{3n} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & c_{n3} & \ldots & c_{nn} \end{bmatrix}.$$

Subtracting suitable multiples of the first row of **C** from the remaining rows, we obtain a matrix **D** in which all elements of the first column except the leading element are equal to zero.

Let $$\mathbf{D}=\begin{bmatrix}1 & d_{12} & d_{13} \dots d_{1n}\\ 0 & & \\ 0 & \mathbf{A}_1 & \\ \vdots & & \\ 0 & & \end{bmatrix}$$

where $\mathbf{A}_1$ is an $(n-1)\times(n-1)$ matrix. The matrix $\mathbf{A}_1$ is non-singular, for otherwise $|\mathbf{A}_1|=0$ and so $|\mathbf{D}|$ is also equal to zero. Thus the matrix **D** will not be non-singular, and therefore **A**, which is row equivalent to **D**, will also not be non-singular.

By the inductive hypothesis, $\mathbf{A}_1$ can be transformed to $\mathbf{I}_{n-1}$ by *E*-row operations. If these elementary row operations be applied to **D**, they will not effect the first row and the first column of **D** and we shall obtain a matrix **M** such that

$$\mathbf{M}=\begin{bmatrix}1 & d_{12} & d_{13} \dots d_{1n}\\ 0 & 1 & 0 \dots 0\\ 0 & 0 & 1 \dots 0\\ \dots & \dots & \dots \dots \dots\\ 0 & 0 & 0 \;\; 0 \;\; 1\end{bmatrix}$$

By adding suitable multiples of the second, third,..., n^{th} rows to the first row of **M**, we obtain the matrix $\mathbf{I}_n$.

Thus the matrix **A** has been reduced to $\mathbf{I}_n$ by *E*-row operations only.

The proof is now complete by induction.

Corollary 1. *If* **A** *be an n-rowed non-singular matrix there exist E-matrices* $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_t$ *such that*

$$\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\mathbf{E}_1\,\mathbf{A}=\mathbf{I}_n.$$

If **A** be an *n-rowed* non-singular matrix, it can be reduced to $\mathbf{I}_n$ by *E*-row operations only. Since every *E*-row operation is equivalent to pre-multiplication by the corresponding *E*-matrix, therefore we can say that if **A** be an *n*-rowed non-singular matrix, there exist *E*-matrices $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_t$ such that

$$\mathbf{E}_t\mathbf{E}_{t-1}\dots\mathbf{E}_2\mathbf{E}_1\mathbf{A}=\mathbf{I}_n.$$

Corollary 2. *Every non-singular matrix* **A** *is expressible as the product of elementary matrices.*

(Nagarjuna 1978; Patna 87; I.A.S. 84)

If **A** be an *n*-rowed non-singular matrix, there exist *E*-matrices $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_t$ such that

$$\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\,\mathbf{E}_1\,\mathbf{A}=\mathbf{I}_n. \qquad \dots(i)$$

Premultiplying both sides of the relation (i) by $(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\mathbf{E}_1)^{-1}$, we get

$$(\mathbf{E}_t\mathbf{E}_{t-1}\dots\mathbf{E}_2\,\mathbf{E}_1)^{-1}(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\mathbf{E}_1)\,\mathbf{A}=(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\mathbf{E}_1)^{-1}\,\mathbf{I}_n$$

or $$\mathbf{I}_n\,\mathbf{A}=\mathbf{E}_1^{-1}\,\mathbf{E}_2^{-1}\,\mathbf{E}_3^{-1}\dots\mathbf{E}_{t-1}^{-1}\,\mathbf{E}_t^{-1}\,\mathbf{I}_n$$

or $$\mathbf{A}=\mathbf{E}_1^{-1}\,\mathbf{E}_2^{-1}\,\mathbf{E}_3^{-1}\dots\mathbf{E}_{t-1}^{-1}\,\mathbf{E}_t^{-1}.$$

Since the inverse of an elementary matrix is also an elementary matrix of the same type, hence we get the result.

Corollary 3. ***The rank of a matrix does not alter by premultiplication or post-multiplication with a non-singular matrix.***

(Nagarjuna 1978, Banaras 1968)

Every non-singular matrix can be expressed as the product of elementary matrices. Also *E*-row (column) transformations are equivalent to pre-(post)-multiplication with the corresponding elementary matrices and elementary transformations do not alter the rank of a matrix. Hence we get the result.

§ 16. Use of Elementary transformations to find the inverse of a non-singular matrix.

Let **A** be a non-singular matrix of order *n*. It can be easily shown that **A** can be reduced to the unit matrix $\mathbf{I}_n$ by a finite number of *E*-row transformations only. Now each *E*-row transformation of a matrix is equivalent to pre-multiplication by the corresponding *E*-matrix. Therefore there exist elementary matrices, say, $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \dots, \mathbf{E}_t$ such that

$$(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\,\mathbf{E}_1)\,\mathbf{A}=\mathbf{I}_n.$$

Post-multiplying both sides by $\mathbf{A}^{-1}$, we get

$$(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\,\mathbf{E}_1)\,\mathbf{A}\mathbf{A}^{-1}=\mathbf{I}_n\,\mathbf{A}^{-1}$$

or $$(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\,\mathbf{E}_1)\,\mathbf{I}_n=\mathbf{A}^{-1} \qquad [\because\ \mathbf{A}\mathbf{A}^{-1}=\mathbf{I}_n,\ \mathbf{I}_n\mathbf{A}^{-1}=\mathbf{A}^{-1}]$$

or $$\mathbf{A}^{-1}=(\mathbf{E}_t\,\mathbf{E}_{t-1}\dots\mathbf{E}_2\,\mathbf{E}_1)\,\mathbf{I}_n.$$

Hence we get the following result.

If a non-singular matrix **A** *of order n is reduced to the unit matrix* $\mathbf{I}_n$ *by a sequence of E-row transformations only, then the same sequence of E-row transformations applied to the unit matrix* $\mathbf{I}_n$ *gives the inverse of* **A** *(i.e*, $\mathbf{A}^{-1}$*).* **(Nagarjuna 1978; Allahabad 78)**

§ 17. Working Rule for finding the inverse of a non-singular matrix by *E*-row transformations. Suppose **A** is a non-singular matrix of order *n*. Then we write

$$\mathbf{A}=\mathbf{I}_n\,\mathbf{A}.$$

Now we go on applying *E*-row transformations only to the matrix **A** and the prefactor $\mathbf{I}_n$ of the product $\mathbf{I}_n\,\mathbf{A}$ till we reach the result $\mathbf{I}_n=\mathbf{B}\mathbf{A}$.

Then obviously **B** is the inverse of **A**.

Solved Examples

Ex. 1. *Find the inverse of the matrix*

$$\mathbf{A}=\begin{bmatrix}1 & 2 & 1\\ 3 & 2 & 3\\ 1 & 1 & 2\end{bmatrix}$$

by using E-transformations. **(Gorakhpur 1965)**

Solution. We write $\mathbf{A}=\mathbf{I}_3\mathbf{A}$,

i.e., $$\begin{bmatrix}1 & 2 & 1\\ 3 & 2 & 3\\ 1 & 1 & 2\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}.$$

Now we go on applying E-row transformations to the matrix A (the left hand member of the above equation) until it is reduced to the form $\mathbf{I}_3$. Every E-row transformation will also be applied to the prefactor $\mathbf{I}_3$ (or its transform) of the product on the right hand side of the above equation.

Applying $R_2 \to R_2-3R_1$, $R_3 \to R_3-R_1$, we get

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

Now we should try to make 1 in the place of the second element of the second row of the matrix on the left hand side. So applying $R_2 \to -\frac{1}{4}R_2$, we get

$$\begin{bmatrix}1 & 2 & 1\\ 0 & 1 & 0\\ 0 & -1 & 1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ \frac{3}{4} & -\frac{1}{4} & 0\\ -1 & 0 & 1\end{bmatrix}\mathbf{A}.$$

Now we shall make zeros in the place of the second elements of the first and the third rows with the help of the second row. So applying $R_1 \to R_1-2R_2$, $R_3 \to R_3+R_2$, we get

$$\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}-\frac{1}{2} & \frac{1}{2} & 0\\ \frac{3}{4} & -\frac{1}{4} & 0\\ -\frac{1}{4} & -\frac{1}{4} & 1\end{bmatrix}\mathbf{A}.$$

Now the third element of the third row is already 1. So to make the third element of the first row zero we apply $R_1 \to R_1-R_3$, and we get

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}-\frac{1}{4} & \frac{3}{4} & -1\\ \frac{3}{4} & -\frac{1}{4} & 0\\ -\frac{1}{4} & -\frac{1}{4} & 1\end{bmatrix}\mathbf{A}.$$

Thus $\mathbf{I}_3=\mathbf{BA}$, where $\mathbf{B}=\begin{bmatrix}-\frac{1}{4} & \frac{3}{4} & -1\\ \frac{3}{4} & -\frac{1}{4} & 0\\ -\frac{1}{4} & -\frac{1}{4} & 1\end{bmatrix}$.

$$\therefore\quad \mathbf{A}^{-1}=\mathbf{B}=\begin{bmatrix}-\frac{1}{4} & \frac{3}{4} & -1\\ \frac{3}{4} & -\frac{1}{4} & 0\\ -\frac{1}{4} & -\frac{1}{4} & 1\end{bmatrix}.$$

Ex. 3. *Find the inverse of the matrix*

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

Solution. We write

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

Applying $R_1 \to -1\,R_1$, we get

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

Applying $R_2 \to R_2-R_1$, $R_3 \to R_3-2R_1$, $R_4 \to R_4+R_1$, we get

$$\begin{bmatrix}1 & 3 & -3 & 1\\ 0 & -2 & 2 & -1\\ 0 & -11 & 8 & -5\\ 0 & 4 & -3 & 2\end{bmatrix}=\begin{bmatrix}-1 & 0 & 0 & 0\\ 1 & 1 & 0 & 0\\ 2 & 0 & 1 & 0\\ -1 & 0 & 0 & 1\end{bmatrix}\mathbf{A}.$$

Applying $R_2 \to -\frac{1}{2}R_2$, we get

$$\begin{bmatrix}1 & 3 & -3 & 1\\ 0 & 1 & -1 & \frac{1}{2}\\ 0 & -11 & 8 & -5\\ 0 & 4 & -3 & 2\end{bmatrix}=\begin{bmatrix}-1 & 0 & 0 & 0\\ -\frac{1}{2} & -\frac{1}{2} & 0 & 0\\ 2 & 1 & 1 & 0\\ -1 & 0 & 0 & 1\end{bmatrix}\mathbf{A}.$$

Applying $R_1 \to R_1-3R_2$, $R_3 \to R_3+11R_2$, $R_4 \to R_4-4R_2$, we get

$$\begin{bmatrix}1 & 0 & 0 & -\frac{1}{2}\\ 0 & 1 & -1 & \frac{1}{2}\\ 0 & 0 & -3 & \frac{1}{2}\\ 0 & 0 & 1 & 0\end{bmatrix}=\begin{bmatrix}\frac{1}{2} & \frac{3}{2} & 0 & 0\\ -\frac{1}{2} & -\frac{1}{2} & 0 & 0\\ -\frac{7}{2} & -\frac{11}{2} & 1 & 0\\ 1 & 2 & 0 & 1\end{bmatrix}\mathbf{A}.$$

Applying $R_3 \leftrightarrow R_4$, we get

$$\begin{bmatrix}1 & 0 & 0 & -\frac{1}{2}\\ 0 & 1 & -1 & \frac{1}{2}\\ 0 & 0 & 1 & 0\\ 0 & 0 & -3 & \frac{1}{2}\end{bmatrix}=\begin{bmatrix}\frac{1}{2} & \frac{3}{2} & 0 & 0\\ -\frac{1}{2} & -\frac{1}{2} & 0 & 0\\ 1 & 2 & 0 & 1\\ -\frac{7}{2} & -\frac{11}{2} & 1 & 0\end{bmatrix}\mathbf{A}.$$

Applying $R_2 \to R_2+R_3$, $R_4 \to R_4+3R_3$, we get

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

Applying $R_2 \to R_2 - R_4$, $R_1 \to R_1 + R_4$, we get

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

Applying $R_4 \to 2R_4$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 1 & 1 & -1 & -2 \\ 1 & 2 & 0 & 1 \\ -1 & 1 & 2 & 6 \end{bmatrix} \mathbf{A}.$$

$$\therefore \quad \mathbf{A}^{-1} = \begin{bmatrix} 0 & 2 & 1 & 3 \\ 1 & 1 & -1 & -2 \\ 1 & 2 & 0 & 1 \\ -1 & 1 & 2 & 6 \end{bmatrix}.$$

Exercises

1. Reduce the matrix $\mathbf{A} = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 4 & 0 \\ 3 & 0 & 1 \end{bmatrix}$ to $\mathbf{I}_3$ by E-row transformations only.

2. Compute the inverse of the following matrices by using elementary operations :

(i) $\begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 1 & 2 & 2 \\ 1 & 1 & 2 & 3 \\ 2 & 2 & 2 & 3 \\ 2 & 3 & 3 & 3 \end{bmatrix}$.

Answers

2. (i) $\begin{bmatrix} -1 & 1 & 2 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{bmatrix}$; (ii) $\begin{bmatrix} -3 & 3 & -3 & 2 \\ 3 & -4 & 4 & -2 \\ -3 & 4 & -5 & 3 \\ 2 & -2 & 3 & -2 \end{bmatrix}$.

5

Vector Space of n-Tuples

§ 1. Vectors. Definition. *Any ordered n-tuple of numbers is called an n-vector.* By an ordered n-tuple we mean a set consisting of n numbers in which the place of each number is fixed. If $x_1, x_2, \ldots, x_n$ be any n numbers, then the ordered n-tuple $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ is called an n-vector. The ordered triad (x_1, x_2, x_3) is called a 3-vector. Similarly $(1, 0, 1, -1)$ and $(1, 8, -5, 7)$ are 4-vectors. The n numbers $x_1, x_2, \ldots, x_n$ are called components of the n-vector $\mathbf{X} = (x_1, x_2, \ldots, x_n)$. A vector may be written either as a *row vector* or as a *column vector*. If $\mathbf{A}$ be a matrix of the type $m \times n$, then each row of $\mathbf{A}$ will be an n-vector and each column of $\mathbf{A}$ will be an m-vector. A vector whose components are all zero is called a *zero vector* and will be denoted by $\mathbf{O}$.

If k be any number and $\mathbf{X}$ be any vector, then relative to the vector $\mathbf{X}$, k is called a scalar.

Algebra of vectors. Since an n-vector is nothing but a row matrix or a column matrix, therefore we can develop an algebra of vectors in the same manner as the algebra of matrices.

Equality of two vectors. Two n-vectors $\mathbf{X}$ and $\mathbf{Y}$ where $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{Y} = (y_1, y_2, \ldots, y_n)$ are said to be equal if and only if their corresponding components are equal *i.e.*, if $x_i = y_i$, for all $i = 1, 2, \ldots, n$. For example if $\mathbf{X} = (1, 4, 7)$ and $\mathbf{Y} = (1, 4, 7)$, then $\mathbf{X} = \mathbf{Y}$. But if $\mathbf{X} = (1, 4, 7)$ and $\mathbf{Y} = (4, 1, 7)$, then $\mathbf{X} \neq \mathbf{Y}$.

Addition of two vectors. If $\mathbf{X} = (x_1, x_2, \ldots, x_n)$ and $\mathbf{Y} = (y_1, y_2, \ldots, y_n)$ then by definition $\mathbf{X} + \mathbf{Y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$.

Thus $\mathbf{X} + \mathbf{Y}$ is an n-vector whose components are the sums of corresponding components of $\mathbf{X}$ and $\mathbf{Y}$.

If $\mathbf{X} = (2, 4, -7)$ and $\mathbf{Y} = (1, -3, 5)$ then

$\mathbf{X} + \mathbf{Y} = (2+1, 4-3, -7+5) = (3, 1, -2)$.

Multiplication of a vector by a scalar (number).

If k be any number and $\mathbf{X} = (x_1, x_2, \ldots, x_n)$, then by definition,

$k\mathbf{X} = (kx_1, kx_2, \ldots, kx_n)$.

The vector $k\mathbf{X}$ is called the scalar multiple of the vector $\mathbf{X}$ by the scalar k.

If $\mathbf{X}=(1, 3, 8)$, then $4\mathbf{X}=(4, 12, 32)$ and $0\mathbf{X}=(0, 0, 0)$.

Properties of addition and scalar multiplication of vectors. If $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ be any three n vectors and p, q be any two numbers, then obviously

(i) $\mathbf{X}+\mathbf{Y}=\mathbf{Y}+\mathbf{X}$.

(ii) $\mathbf{X}+(\mathbf{Y}+\mathbf{Z})=(\mathbf{X}+\mathbf{Y})+\mathbf{Z}$.

(iii) $p(\mathbf{X}+\mathbf{Y})=p\mathbf{X}+p\mathbf{Y}$.

(iv) $(p+q)\mathbf{X}=p\mathbf{X}+q\mathbf{X}$.

(v) $p(q\mathbf{X})=(pq)\mathbf{X}$.

§ 2. Linear dependence and linear independence of vectors. Linearly dependent set of vectors. Definition.

A set of r n-vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_r$ *is said to be linearly dependent if there exist r scalars (numbers)* $k_1, k_2, \ldots, k_r$, *not all zero, such that*

$$k_1\mathbf{X}_1+k_2\mathbf{X}_2+\ldots+k_r\mathbf{X}_r=\mathbf{O},$$

where, $\mathbf{O}$, *denotes the n-vector whose components are all zero.*

Linearly independent set of vectors. Definition.

A set of r n-vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_r$ *is said to be linearly independent if every relation of the type*

$$k_1\mathbf{X}_1+k_2\mathbf{X}_2+\ldots+k_r\mathbf{X}_r=\mathbf{O}$$

implies $k_1=k_2=k_3=\ldots=k_r=0$.

Ex. 1. *Show that the vectors* $\mathbf{X}_1=(1, 2, 4)$, $\mathbf{X}_2=(3, 6, 12)$ *are linearly dependent.*

Solution. By a little inspection, we see that

$$3\mathbf{X}_1+(-1)\mathbf{X}_2=(3, 6, 12)+(-3, -6, -12)$$
$$=(0, 0, 0)=\mathbf{O}.$$

Thus there exist numbers $k_1=3$, $k_2=-1$ which are not all zero such that

$$k_1\mathbf{X}_1+k_2\mathbf{X}_2=\mathbf{O}.$$

Hence the vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ are linearly dependent.

Ex. 2. *Show that the set consisting only of the zero vector,* $\mathbf{O}$, *is linearly dependent.*

Solution. Let $\mathbf{X}=(0, 0, 0, \ldots, 0)$ be an n-vector whose components are all zero. Then the relation $k\mathbf{X}=\mathbf{O}$ is true for some non-zero value of the number k. For example,

$$1\mathbf{X}=\mathbf{O} \text{ and } 1\neq 0.$$

Hence the vector $\mathbf{O}$ is linearly dependent.

Ex. 3. *Show that the set of three 3-vectors is linearly independent.*

Solution. Let k_1, k_2, k_3 be three numbers such that

$$k_1\mathbf{X}_1+k_2\mathbf{X}_2+k_3\mathbf{X}_3=\mathbf{O},$$

i.e., $k_1(1, 0, 0)+k_2(0, 1, 0)+k_3(0, 0, 1)=(0, 0, 0)$,

i.e., $(k_1, 0, 0)+(0, k_2, 0)+(0, 0, k_3)=(0, 0, 0)$,

i.e., $(k_1, k_2, k_3)=(0, 0, 0)$.

Obviously this relation is true if and only if

$$k_1=0, k_2=0, k_3=0.$$

Hence the vectors $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ are linearly independent.

A vector as a linear combination of vectors.

Definition. *A vector* $\mathbf{X}$ *which can be expressed in the form*

$$\mathbf{X}=k_1\mathbf{X}_1+k_2\mathbf{X}_2+\ldots+k_r\mathbf{X}_r,$$

is said to be a linear combination of the set of vectors

$$\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_r.$$

Here $k_1, k_2, \ldots, k_r$ are any numbers.

The following two results are quite obvious :

(*i*) *If a set of vectors is linearly dependent, then at least one member of the set can be expressed as a linear combination of the remaining members.*

(*ii*) *If a set of vectors is linearly independent then no member of the set can be expressed as a linear combination of the remaining members.*

§ 3. The n-vector space. The set of all n-vectors of a field F is called the n-vector space over F. It is usually denoted by $V_n(F)$ or simply by V_n if the field is understood. Similarly the set of all 3-vectors is a vector space which is usually denoted by V_3. The elements of the field F are known as *scalars* relatively to the vectors.

§ 4. Sub-space of an n-vector space V_n. Definition.

A non-empty set, S, of vectors of V_n *is called a vector subspace of* V_n, *if* $\mathbf{a}+\mathbf{b}$ *belongs to S whenever* $\mathbf{a}, \mathbf{b}$ *belong to S and* $k\mathbf{a}$ *belongs to S whenever* $\mathbf{a}$ *belongs to S, where k is any scalar.*

It is important to note that every sub-space of V_n contains the zero vector, being the scalar product of any vector with the scalar zero.

Example. If $\mathbf{a}=(a_1, a_2, a_3)$ is any non-zero vector of V_3, then the set S of vectors $k\mathbf{a}$ is a subspace of V_3, where k is a variable scalar which can take any value. The sum of any two members

$k_1\mathbf{a}$, $k_2\mathbf{a}$ of S, is the vector $k_1\mathbf{a}+k_2\mathbf{a}$

i.e., $(k_1a_1, k_1a_2, k_1a_3)+(k_2a_1, k_2a_2, k_2a_3)$

i.e., $(k_1+k_2)\ \mathbf{a}$

which is also a member of S. Also the scalar multiple by any scalar x of any vector $k_1\mathbf{a}$ of S is the vector $(xk_1)\ \mathbf{a}$ which is again a member of S.

Hence the set S of vectors $k\mathbf{a}$ is a subspace of V_3.

§ 5. Vector subspace spanned by a given system of vectors.

Let $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ be any three vectors of V_3. The set S of all vectors of the form $x\mathbf{a}+y\mathbf{b}+z\mathbf{c}$, where x, y, z are any scalars, is a subspace of V_3. For, if $x_1\mathbf{a}+y_1\mathbf{b}+z_1\mathbf{c}$, $x_2\mathbf{a}+y_2\mathbf{b}+z_2\mathbf{c}$ be any two members of S, then

$$(x_1\mathbf{a}+y_1\mathbf{b}+z_1\mathbf{c})+(x_2\mathbf{a}+y_2\mathbf{b}+z_2\mathbf{c})=(x_1+x_2)\ \mathbf{a}$$
$$+(y_1+y_2)\ \mathbf{b}+(z_1+z_2)\ \mathbf{c},$$

which is also a member of S. Also if k be any scalar, then

$$k\ (x_1\mathbf{a}+y_1\mathbf{b}+z_1\mathbf{c})=(kx_1)\ \mathbf{a}+(ky_1)\ \mathbf{b}+(kz_1)\ \mathbf{c},$$

which is again a member of S. Thus S is a vector subspace and we say that S is spanned by the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. More generally;

if $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ be a set of r fixed vectors of V_n, then the set S of all n-vectors of the form $p_1\mathbf{a}_1+p_2\mathbf{a}_2+\ldots+p_r\mathbf{a}_r$ where $p_1, p_2, \ldots, p_r$ are any scalars is a vector subspace of V_n.

This vector space is said to be spanned by the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$. **Thus a vector space which arises as a set of all linear combinations of any given set of vectors, is said to be spanned by the given set of vectors.**

§ 6. Basis and dimension of a subspace.

A set of vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \ldots, \mathbf{a}_k$ belonging to the subspace S is said to be a basis of S, if

(i) the subspace S is spanned by the set $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ and

(ii) the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ are linearly independent.

It can be easily shown that every subspace, S, of V_n possesses a basis.

It can be easily shown that the vectors

$\mathbf{e}_1=(1, 0, 0, \ldots, 0)$, $\mathbf{e}_2=(0, 1, 0, \ldots, 0)$, $\mathbf{e}_3=(0, 0, 1, 0, \ldots, 0), \ldots$
$\mathbf{e}_n=(0, 0, 0, \ldots, 1)$ constitute a basis of V_n.

We have already shown that these vectors are linearly independent. Moreover any vector $\mathbf{a}=(a_1, a_2, \ldots, a_n)$ of V_n is expressible as $\mathbf{a}=a_1\mathbf{e}_1+a_2\mathbf{e}_2+a_3\mathbf{e}_3+\ldots+a_n\mathbf{e}_n$. Hence the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n$ constitute a basis of V_n.

Theorem. *A basis of a subspace, S, can always be selected out of a set of vectors which span S.*

Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ be a set of vectors which spans a subspace S. If these vectors are linearly independent, they already constitute a basis of S as they span S. In case they are linearly dependent, some member of the set is a linear combination of the other members. Deleting this member we obtain another set which also spans S.

Continuing in this manner we shall ultimately, in a finite number of steps, arrive at a basis of S.

A vector subspace may (and in fact does) possess several bases.

For example, if $\mathbf{a}_1=(1, 0, 0)$, $\mathbf{a}_2=(0, 1, 0)$, $\mathbf{a}_3=(0, 0, 1)$, $\mathbf{b}_1=(1, 1, 1)$, $\mathbf{b}_2=(1, 1, 0)$, $\mathbf{b}_3=(1, 0, 0)$, then $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ constitute a basis of V_3. Also $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ constitute a basis of V_3.

But it is important to note that the *number of members in any one basis of a subspace is the same as in any other basis. This number is called the dimension of the subspace.*

We have already shown that one basis of V_n possesses n members. Therefore every basis of V_n must possess n members. Thus V_n is of dimension n. In particular the dimension of V_3 is 3.

Also it can be easily shown that if r be the dimension of a subspace S and if $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k$ be a linearly independent set of vectors belonging to S, then we can always find vectors $\mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \ldots, \mathbf{a}_r$ such that the vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k, \mathbf{a}_{k+1}, \ldots, \mathbf{a}_r$ constitute a basis of S. In other words we can say that every linearly independent set of vectors belonging to a subspace S can always be extended so as to constitute a basis of S.

Moreover if r be the dimension of a subspace S, then every set of more than r members of S will be linearly dependent.

Intersection of subspaces. If S and T be two subspaces of V_n, then the vectors common to both S and T also constitute a subspace. This subspace is called the intersection of the subspaces S and T.

Row rank and column rank of a matrix. (Nagarjuna 1978)

§ 7. Row rank of a matrix. Let $\mathbf{A}=[a_{ij}]$ be any $m\times n$ matrix. Each of the m rows of $\mathbf{A}$ consists of n elements. Therefore the row vectors of $\mathbf{A}$ are n-vectors. These row vectors of $\mathbf{A}$ will span a subspace R of V_n. This subspace R is called the *row space* of the matrix $\mathbf{A}$. The dimension r of R is called the *row rank* of $\mathbf{A}$. In other words the row rank of a matrix A is equal to the maximum number of linearly independent rows of A.

Left nullity of a matrix. Suppose X is an m-vector written in the form of a row vector. Then the matrix product XA is defined. The subspace S of V_m generated by the row vectors X belonging to V_m such that $XA=O$ is called the *row null* space of the matrix A. The dimension s of S is called the left nullity or *row nullity* of the matrix A.

We shall now prove that the *sum of the row rank and the row nullity of a matrix is equal to the number of rows i.e.,*

$$r+s=m.$$

Proof. Since the row space R of A is spanned by the row vectors of A, therefore it will be a set of all vectors of the form $x_1\,(a_{11}, a_{12}, a_{13}, \ldots, a_{1n})+x_2\,(a_{21}, a_{22}, a_{23}, \ldots, a_{2n})+\ldots$

$$+x_m\,(a_{m1}, a_{m2}, a_{m3}, \ldots\ldots, a_{mn})$$

i.e., of the form $(x_1a_{11}+x_2a_{21}+\ldots+x_ma_{m1},\ x_1a_{12}+x_2a_{22}+\ldots+x_ma_{m2},\ x_1a_{13}+x_2a_{23}+\ldots+x_ma_{m3}, \ldots, x_1a_{1n}+x_2a_{2n}+\ldots+x_ma_{mn})$ *i.e.*, of the form XA, where $X=(x_1, x_2, \ldots, x_m)$ is an m-vector. Let $u_1, u_2, \ldots, u_s$ be a basis of the subspace S of V_m generated by all vectors X such that $XA=O$. Then, we have

$$u_1A=u_2A=\ldots=u_sA=O.$$

Since the vectors $u_1, u_2, \ldots, u_s$ belong to V_m and form a linearly independent set, therefore we can find vectors $u_{s+1}, u_{s+2}, \ldots, u_m$ in V_m such that the vectors $u_1, u_2, \ldots, u_s, u_{s+1}, \ldots, u_m$ constitute a basis of V_m. Then every vector X belonging to V_m can be expressed in the form

$$=h_1u_1+h_2u_2+\ldots+h_mu_m.$$

Now every member of the subspace R is expressible as XA *i.e.*, as $(h_1u_1+h_2u_2+\ldots+h_mu_m)\,A$

i.e., as $h_1u_1A+h_2u_2A+\ldots+h_su_sA+h_{s+1}u_{s+1}A+h_{s+2}u_{s+2}A+\ldots$

$$+h_mu_mA$$

i.e., as $h_{s+1}u_{s+1}A+h_{s+2}u_{s+2}A+\ldots+h_mu_mA.$

Therefore the $m-s$ n-vectors $u_{s+1}A, u_{s+2}A, \ldots, u_mA$ span R. In fact, these vectors form a basis of R. For any relation of the form

$$k_{s+1}u_{s+1}A+k_{s+2}u_{s+2}A+\ldots+k_mu_mA=O$$

implies $(k_{s+1}u_{s+1}+k_{s+2}u_{s+2}+\ldots+k_mu_m)A=O$, which shows that $k_{s+1}u_{s+1}+k_{s+2}u_{s+2}+\ldots+k_mu_m$ is a member of the subspace S and as such it can be linearly expressed in terms of the basis $u_1, u_2, \ldots, u_s$ of S.

But the vectors $u_1, u_2, \ldots, u_m$ are linearly independent. Therefore a relation of the form $k_{s+1}\,u_{s+1}+k_{s+2}\,u_{s+2}+\ldots+k_mu_m=p_1u_1+p_2u_2+\ldots+p_su_s$ will exist if and only if $k_{s+1}=k_{s+2}=\ldots=k_m=0$. Hence the vectors $u_{s+1}A, u_{s+2}A, \ldots, u_mA$ are linearly independent and form a basis of R. Thus the dimension of R is $m-s$.

Hence $r=m-s$ or $r+s=m$.

§ 8. Column rank of a matrix. Let $A=[a_{ij}]$ be any $m\times n$ matrix. Each of the n columns of A consists of m elements. Therefore the column vectors of A are m-vectors. These column vectors of A will span a subspace C of V_m. This subspace C is called the *column space* of the matrix A. The dimension c of C is called the *column rank* of A. In other words the column rank of a matrix A is equal to the maximum number of linearly independent columns of A.

Right nullity of a matrix. Suppose Y is an n-vector written in the form of a column vector. Then the matrix product AY is defined. The subspace T of V_n generated by the column vectors Y belonging to V_n such that $AY=O$ is called the *column null space* of the matrix A. The dimension t of T is called the right nullity or column nullity of the matrix A.

As in § 7, we can show that $c+t=n$.

§ 9. Invariance of row rank under *E*-row operations.

Theorem. *Row equivalent matrices have the same row rank.*

Proof. Let A be any given $m\times n$ matrix. Let B be a matrix row equivalent to A. Since B is obtainable from A by a finite chain of E-row operations and every E-row operation is equivalent to pre-multiplication by the corresponding E-matrix, there exist E-matrices $E_1, E_2, \ldots, E_k$ each of the type $m\times m$ such that

$$B=E_k\,E_{k-1}\ldots E_2E_1A,$$

i.e. $\quad B=PA,$

where $P=E_k\,E_{k-1}\;\;E_2E_1$ is a non-singular matrix of the type $m\times m$.

Let us write

$$B=PA=\begin{bmatrix} p_{11} & p_{12} & \ldots & p_{1m} \\ p_{21} & p_{22} & \ldots & p_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ p_{m1} & p_{m2} & \ldots & p_{mm} \end{bmatrix}\begin{bmatrix} R_1 \\ R_2 \\ \ldots \\ \ldots \\ R_m \end{bmatrix} \quad \ldots(i)$$

where the matrix **A** has been expressed as a matrix of its row sub-matrices $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_m$.

From the product of the matrices on R.H.S. of (i), we observe that the rows of the matrix **B** are

$$p_{11}\mathbf{R}_1+p_{12}\mathbf{R}_2+\ldots+p_{1m}\mathbf{R}_m,$$
$$p_{21}\mathbf{R}_1+p_{22}\mathbf{R}_2+\ldots+p_{2m}\mathbf{R}_m,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$p_{m1}\mathbf{R}_1+p_{m2}\mathbf{R}_2+\ldots+p_{mm}\mathbf{R}_m.$$

Thus we see that the rows of **B** are all linear combinations of the rows $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_m$ of **A**. Therefore every member of the row space of **B** is also a member of the row space of **A**.

Similarly by writing $\mathbf{A}=\mathbf{P}^{-1}\mathbf{B}$ and giving the same reasoning we can prove that every member of the row space of **A** is also a member of the row space of **B**. Therefore the row spaces of **A** and **B** are identical.

Thus we see that elementary row operations do not alter the row space of a matrix. Hence the row rank of a matrix remains invariant under *E*-row transformations.

Note. From the above theorem we also conclude that *pre-multiplication by a non-singular matrix does not alter the row rank of a matrix.*

§ 10. Invariance of column rank under *E*-column operations.

Theorem. *Column equivalent matrices have the same column rank.*

Or

Post-multiplication by a non-singular matrix do s not alter the column rank of a matrix.

Proof. Proceeding in the same way as in § 9, we can show that post multiplication with a non-singular matrix does not alter the column space and therefore the column rank of a matrix.

Note. Since every n-rowed *E*-matrix is obtainable from $\mathbf{I}_n$ by a single *E*-operation (row or column operation as may be desired), therefore the row rank and column rank of an *E*-matrix are each equal to n.

§ 11. Invariance of a column rank under *E*-row operations.

Theorem. *Row equivalent matrices have the same column rank.* (I.A.S. 1984)

Let **A** be any given $m\times n$ matrix and let **B** be a matrix row equivalent to **A**. Then there exists a non-singular matrix **P** such that $\mathbf{B}=\mathbf{PA}$.

For every column vector **X** such that $\mathbf{AX}=\mathbf{O}$, we have

$$\mathbf{BX}=(\mathbf{PA})\mathbf{X}=\mathbf{P}\,(\mathbf{AX})=\mathbf{PO}=\mathbf{O}.$$

Since $\mathbf{B}=\mathbf{PA}$, therefore $\mathbf{A}=\mathbf{P}^{-1}\mathbf{B}$.

Therefore for every vector **X** such that $\mathbf{BX}=\mathbf{O}$, we have

$$\mathbf{AX}=(\mathbf{P}^{-1}\mathbf{B})\,\mathbf{X}=\mathbf{P}^{-1}\,(\mathbf{BX})=\mathbf{P}^{-1}\mathbf{O}=\mathbf{O}.$$

Thus we see that the matrices **A** and **B** have the same right nullities and consequently their column ranks are equal.

Similarly we can prove that *column equivalent matrices have the same row rank.*

§ 12. Theorem. *If r, be the row rank of an $m\times n$ matrix* **A** *then there exists a non-singular matrix,* **P** *such that*

$$\mathbf{PA}=\begin{bmatrix}\mathbf{K}\\ \mathbf{O}\end{bmatrix},$$

where **K** *is an $r\times n$ matrix consisting of a set of r linearly independent rows of* **A**.

Proof. If the row rank r of **A** is zero, we have nothing to prove. Therefore let us assume that $r>0$. The matrix **A** has then r linearly independent rows. By elementary row operations on **A** we can bring these linearly independent rows in the first r places. Since the last $m-r$ rows are now linear combinations of the first r rows, they can be made zero by *E*-row operations without altering the first r rows.

Thus we see that the matrix **A** is row equivalent to a matrix **B** such that $\mathbf{B}=\begin{bmatrix}\mathbf{K}\\ \mathbf{O}\end{bmatrix}$,

where **K** is an $r\times n$ matrix consisting of a set of r linearly independent rows of **A**.

Since every elementary row operation is equivalent to pre-multiplication by the corresponding *E*-matrix and the product of *E*-matrices is a non-singular matrix, therefore there exists a non-singular matrix **P** such that

$$\mathbf{PA}=\begin{bmatrix}\mathbf{K}\\ \mathbf{O}\end{bmatrix}.$$

Similarly considering column transformations instead of row transformations, we can show *that if c be the column rank of a matrix* **A**, *then there exists a non-singular matrix* **R** *such that*

$$\mathbf{AR}=[\mathbf{L}\quad \mathbf{O}],$$

where **L** *is an* $m\times c$ *matrix consisting of, c, linearly independent columns of* **A**.

§ 13. Equality of row rank, column rank, and rank.

Theorem 1. *The row rank of a matrix is the same as its rank.* **(Allahabad 1977, Andhra 81)**

Let, s, be the row rank and r, the rank of an $m\times n$ matrix **A**. Since the matrix **A** is of row rank s, therefore by § 12, page 203 there exists a non-singular matrix **P** such that

$$\mathbf{PA}=\begin{bmatrix}\mathbf{K}\\ \mathbf{O}\end{bmatrix},$$

where **K** is an $s\times n$ matrix.

Now we know that pre-multiplication by a non-singular matrix does not alter the rank of a matrix.

$\therefore$ Rank (**PA**)=Rank **A**$=r$.

But each minor of order $(s+1)$ of the matrix **PA** involves at-least one row of zeros.

$\therefore$ Rank (**PA**) $\leqslant s$.

$\therefore$ $r\leqslant s$.

Again, since the rank of the matrix **A** is r, therefore by § 12, page 186 there exists a non-singular matrix **R** such that

$$\mathbf{RA}=\begin{bmatrix}\mathbf{G}\\ \mathbf{O}\end{bmatrix},$$

where **G** is an $r\times n$ matrix.

Now we know that pre-multiplication by a non-singular matrix does not alter the row rank of a matrix.

$\therefore$ Row rank (**RA**)=Row rank **A**$=s$.

But the matrix **RA** has only r non-zero rows. Therefore the row rank of **RA** can, at the most, be equal to r.

$\therefore$ $s\leqslant r$.

Hence $r=s$.

Theorem 2. *The column rank of a matrix is the same as its rank.* **(Allahabad 1977; Andhra 81)**

Let the matrix **A**$'$ be the transpose of the matrix **A**. Then the columns of **A** are the rows of **A**$'$.

$\therefore$ the column rank of **A**=row rank of **A**$'$
=rank of **A**$'$=rank of **A**.

Thus from theorems 1 and 2, we conclude that the rank, row

rank and column rank of a matrix are all equal. In other words we can say that *the maximum number of linearly independent rows of a matrix is equal to the maximum number of its linearly independent columns and is equal to the rank of the matrix.*

§ 14. Rank of a sum.

Theorem. *Rank of the sum of two matrices cannot exceed the sum of their ranks.* **(Gujrat 1971)**

Proof. Let **A**, **B** be two matrices of the same type. Let $S_{\mathbf{A}}$, $S_{\mathbf{B}}$, $S_{\mathbf{A+B}}$ denote the row-spaces of the matrices **A**, **B**, **A**+**B** respectively. Let S, denote the subspace generated jointly by the rows of **A** as well as the rows of **B**.

Now the number of members in a basis of S must be less than or equal to the sum of the numbers of members in the bases of **A** and **B**.

$\therefore$ Dimension $S\leqslant$ Dimension $S_{\mathbf{A}}+$Dimension $S_{\mathbf{B}}$.

Again the row space $S_{\mathbf{A+B}}$ is a subspace of S.

$\therefore$ Dimension $S_{\mathbf{A+B}}\leqslant$ Dimension S.

$\therefore$ Dimension $S_{\mathbf{A+B}}\leqslant$ Dimension $S_{\mathbf{A}}+$Dimension $S_{\mathbf{B}}$.

$\therefore$ Row rank (**A**+**B**) $\leqslant$ Row rank (**A**)+Row rank (**B**).

$\therefore$ Rank (**A**+**B**) $\leqslant$ Rank (**A**)+Rank (**B**).

[Since the rank and row rank of a matrix are equal]

§ 15. Theorem. *If* **A**, **B** *are two n-rowed square matrices then Rank* (**AB**) $\geqslant$ (*Rank* **A**)+(*Rank* **B**)$-n$. **(Nagarjuna 1980; I.C.S. 87)**

Proof. Let r be the rank of the matrix **A**. Then by § 9, page 164, there exist two non-singular matrices **P** and **Q** such that

$$\mathbf{PAQ}=\begin{bmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix} \quad \ldots(i)$$

Pre-multiplying both sides of (i) by $\mathbf{P}^{-1}$, we get

$$\mathbf{P}^{-1}\,\mathbf{PAQ}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}$$

or $$\mathbf{IAQ}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}$$

or $$\mathbf{AQ}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}. \quad \ldots(ii)$$

Similarly post-multiplying both sides of (ii) by $\mathbf{Q}^{-1}$, we get

$$\mathbf{A}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{O}\end{bmatrix}\mathbf{Q}^{-1}.$$

Let us now consider another matrix **C** denoted as

$$\mathbf{C}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{O}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{I}_{n-r}\end{bmatrix}\mathbf{Q}^{-1}.$$

Then $\mathbf{A}+\mathbf{C}=\mathbf{P}^{-1}\begin{bmatrix}\mathbf{I}_r & \mathbf{O}\\ \mathbf{O} & \mathbf{I}_{n-r}\end{bmatrix}\mathbf{Q}^{-1}=\mathbf{P}^{-1}\,\mathbf{I}_n\,\mathbf{Q}^{-1}=\mathbf{P}^{-1}\,\mathbf{Q}^{-1}.$

Thus $\mathbf{A}+\mathbf{C}$ is a non-singular matrix of order n, it being the product of two non-singular matrices.

Therefore Rank $(\mathbf{A}+\mathbf{C})=n$.

Also Rank $\mathbf{C}=n-r=n-$Rank $(\mathbf{A})$.

Now Rank $\{(\mathbf{A}+\mathbf{C})\,\mathbf{B}\}$=Rank **B**, as the matrix $\mathbf{A}+\mathbf{C}$ is non-singular and the rank of a matrix is not altered by pre-multiplication with a non-singular matrix.

Thus Rank **B**=Rank $\{(\mathbf{A}+\mathbf{C})\,\mathbf{B}\}$=Rank $(\mathbf{AB}+\mathbf{CB})$.

$\therefore$ Rank $\mathbf{B}\leqslant$ Rank $(\mathbf{AB})$+Rank $(\mathbf{CB})$ ...(iii)

[§ 14 page 205]

Again Rank $(\mathbf{CB})\leqslant$ Rank **C**. [Since the rank of the product of two matrices is less than or equal to the rank of either matrix]

$\therefore$ Rank $(\mathbf{CB})\leqslant n-$Rank $(\mathbf{A})$ ...(iv)

[$\because$ Rank $\mathbf{C}=n-$Rank **A**]

Now from (iii) and (iv), we get

Rank $(\mathbf{B})\leqslant$ Rank $(\mathbf{AB})+n-$Rank $(\mathbf{A})$

or Rank $(\mathbf{AB})\geqslant$ Rank **A**+Rank $\mathbf{B}-n$.

Solved Examples

Ex. 1. *If* **A** *be any non-singular matrix and* **B** *a matrix such that* **AB** *exists, then show that* **AB** *and* **B** *have the same rank.* **(Poona 1970; Nagarjuna 78)**

Solution. Let $\mathbf{C}=\mathbf{AB}$.

Since **A** is non-singular, therefore $\mathbf{B}=\mathbf{A}^{-1}\,\mathbf{C}$.

Now we know that the rank of the product of two matrices cannot exceed the rank of either matrix.

$\therefore$ Rank **C**=Rank $(\mathbf{AB})\leqslant$ Rank **B**,

and Rank **B**=Rank $(\mathbf{A}^{-1}\,\mathbf{C})\leqslant$ Rank **C**.

$\therefore$ Rank **B**=Rank **C**=Rank **AB**.

Ex. 2 *Show that, if the* $n\times n$ *matrix* **A** *satisfies the equation* $\mathbf{A}^2=\mathbf{A}$, *then rank* **A**+ *rank* $(\mathbf{I}-\mathbf{A})=n$. **(Gujrat 1971)**

Solution. **A** is an $n\times n$ matrix such that

$\mathbf{A}-\mathbf{A}^2=\mathbf{O}$ *i.e.*, $\mathbf{A}\,(\mathbf{I}_n-\mathbf{A})=\mathbf{O}$.

Now the sum of the matrices **A** and $\mathbf{I}_n-\mathbf{A}$ is the matrix $\mathbf{I}_n$ and we know that the rank of the sum of two matrices cannot exceed the sum of their ranks.

$\therefore$ Rank $(\mathbf{A}+\mathbf{I}_n-\mathbf{A})\leqslant$ Rank **A**+Rank $(\mathbf{I}_n-\mathbf{A})$

or Rank $\mathbf{I}_n\leqslant$ Rank **A**+Rank $(\mathbf{I}_n-\mathbf{A})$

or $n\leqslant$ Rank **A**+Rank $(\mathbf{I}_n-\mathbf{A})$. ...(i)

Again since the product of two n-rowed matrices **A** and $\mathbf{I}_n-\mathbf{A}$ is a zero matrix *i.e.* is a matrix of rank zero, therefore by § 15, page 205, we have

$0\geqslant$ Rank **A**+Rank $(\mathbf{I}_n-\mathbf{A})-n$

i.e. $n\geqslant$ Rank **A**+Rank $(\mathbf{I}_n-\mathbf{A})$. ...(ii)

Hence from (i) and (ii), we get

Rank **A**+Rank $(\mathbf{I}_n-\mathbf{A})=n$.

Ex. 3. *If* **A** *be an* $n\times n$ *matrix, show that the rank of Adj* **A** *is* n, 1 *or* 0 *according as the rank of* **A** *is* n, $n-1$ *or less than* $n-1$.

Solution. (i) Let **A** be an $n\times n$ matrix. Then

$\mathbf{A}\,(\text{Adj. }\mathbf{A})=|\,\mathbf{A}\,|\,\mathbf{I}_n$.

$\therefore$ $|\,\mathbf{A}\,|\,|\,\text{Adj. }\mathbf{A}\,|=|\,\mathbf{A}\,|\,|\,\mathbf{I}_n\,|=|\,\mathbf{A}\,|$.

Since the matrix **A** is of rank n, therefore $|\,\mathbf{A}\,|\neq 0$.

$\therefore$ $|\,\mathbf{A}\,|\,|\,\text{Adj }\mathbf{A}\,|=|\,\mathbf{A}\,|$ gives $|\,\text{Adj. }\mathbf{A}\,|=1$.

Thus the matrix Adj. **A** is also non-singular. Hence it is of rank n.

(ii) If the rank of **A** is $n-1$, then at least one minor of order $n-1$ of the matrix **A** is not equal to zero. Therefore the matrix Adj. **A** will be a non-zero matrix and thus the rank of the matrix Adj. **A** will be greater than zero.

Again the rank of the matrix **A** is $n-1$. Therefore $|\,\mathbf{A}\,|=0$. Therefore **A** (Adj. **A**) is a zero matrix and therefore is of rank zero.

Hence by theorem on page 205, we have

$0\geqslant$ Rank **A**+Rank Adj. $\mathbf{A}-n$

or Rank **A**+Rank Adj. $\mathbf{A}\leqslant n$

or $n-1+$Rank Adj. $\mathbf{A}\leqslant n$

or Rank Adj. $\mathbf{A}\leqslant 1$.

But we have shown that Rank Adj. $\mathbf{A}>0$.

Hence Rank Adj. $\mathbf{A}=1$.

(iii) If the rank of **A** is less than $n-1$, then all minors of order $n-1$ of the matrix **A** will be zero. Therefore the matrix Adj **A** will be a zero matrix and hence Rank Adj. **A** will be zero.

Ex. 4. *Show that the vectors* $\mathbf{X}_1=(1,2,3)$, $\mathbf{X}_2=(2,-2,0)$ *form a linearly independent set.* **(Nagarjuna 1980; Agra 79)**

Solution. Consider the matrix

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

The minor $\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix}$ of A is not equal to zero. Therefore rank $A=2$.

$\therefore$ row rank of $A=2=$the maximum number of linearly independent rows of A. Hence the vectors $(1, 2, 3)$ and $(2, -2, 0)$ are linearly independent.

Ex. 5. *Show that the vectors* $X_1=(3, 1, -4)$, $X_2=(2, 2, -3)$ *form a linearly independent set.* (Agra 1970)

Solution. Consider the matrix

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

The minor $\begin{vmatrix} 3 & 1 \\ 2 & 2 \end{vmatrix}$ of A is not equal to zero because its value is $6-2$ *i.e.* 4. Therefore rank $A=2$.

$\therefore$ row rank of $A=2=$the maximum number of linearly independent rows of A.

Hence the rows of A form a linearly independent set of vectors. Thus the vectors X_1 and X_2 are linearly independent.

Ex. 6. *Show that the vectors* $X_1=(1, 2, 3)$, $X_2=(3, -2, -1)$, $X_3=(1, -6, -5)$ *form a linearly independent system.* (Agra 1988)

6

Linear Equations

We shall devote this chapter to the study of the nature of solutions of a system of linear equations with the help of the theory developed in the preceding chapters. We shall first consider systems of homogeneous linear equations and proceed to discuss systems of non-homogeneous linear equations.

§ 1. Homogeneous Linear Equations. Suppose

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n=0, \\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n=0 \end{array}\right\} \quad \ldots(1)$$

is a system of m homogeneous equations in n unknowns $x_1, x_2, \ldots, x_n$. Let

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ a_{31} & a_{32} & \ldots & a_{3n} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}_{m\times n}, \quad X=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \ldots \\ \ldots \\ x_n \end{bmatrix}_{n\times 1}, \quad O=\begin{bmatrix} 0 \\ 0 \\ 0 \\ \ldots \\ \ldots \\ 0 \end{bmatrix}_{m\times 1},$$

where A, X, O are $m\times n$, $n\times 1$, $m\times 1$ matrices respectively. Then obviously we can write the system of equations (1) in the form of a single matrix equation

$$AX=O. \quad \ldots(2)$$

The matrix A is called the **coefficient matrix** of the system of equations (1).

Obviously $x_1=0, x_2=0, \ldots, x_n=0$ *i.e.*, $X=O$ is a solution of (1). It is a trivial (self-obvious) solution of (1).

Again suppose X_1 *and* X_2 *are two solutions of* (2). *Then their linear combination,* $k_1X_1+k_2X_2$, *where* k_1 *and* k_2 *are any arbitrary numbers, is also a solution of* (2).

We have $\mathbf{A}(k_1\mathbf{X}_1+k_2\mathbf{X}_2)=k_1(\mathbf{AX}_1)+k_2(\mathbf{AX}_2)$

$=k_1\mathbf{O}+k_2\mathbf{O}$ $\quad [\because \mathbf{AX}_1=\mathbf{O}$ and $\mathbf{AX}_2=\mathbf{O}]$

$=\mathbf{O}$.

Hence $k_1\mathbf{X}_1+k_2\mathbf{X}_2$ is also a solution of (2).

Therefore the collection of all the solutions of the system of equations $\mathbf{AX}=\mathbf{O}$ *forms a sub-space of the n-vector space* V_n.

(I. A. S. 1970)

Theorem. *The number of linearly independent solutions of m homogeneous linear equations in n variables,* $\mathbf{AX}=\mathbf{O}$, *is* $(n-r)$, *where r is the rank of the matrix* **A**.

(Nagarjuna 1978; Kerala 70, I.A.S. 70; Poona 70)

Let $$\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}_{m\times n}, \text{ and } \mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ \dots \\ x_n \end{bmatrix}_{n\times 1}.$$

Since the rank of the coefficient matrix **A** is r, therefore it has r linearly independent columns. Without loss of generality we can suppose that the first r columns from the left of the matrix **A** are linearly independent, because it amounts only to renaming the components of **X**.

The matrix **A** can be written as

$$\mathbf{A}=[\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_r, \dots, \mathbf{C}_n]_{1\times n},$$

where $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_n$ are the column vectors of the matrix **A** each of them being an m-vector.

The equation $\mathbf{AX}=\mathbf{O}$ can now be written as the vector equation

$$x_1\mathbf{C}_1+x_2\mathbf{C}_2+\dots+x_r\mathbf{C}_r+x_{r+1}\mathbf{C}_{r+1}+\dots+x_n\mathbf{C}_n=\mathbf{O}. \quad \dots(1)$$

Since each of the vectors $\mathbf{C}_{r+1}, \mathbf{C}_{r+2}, \dots, \mathbf{C}_n$ is a linear combination of the vectors $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_r$, therefore we have relations of the type

$$\left.\begin{array}{l} \mathbf{C}_{r+1}=p_{11}\mathbf{C}_1+p_{12}\mathbf{C}_2+\dots+p_{1r}\mathbf{C}_r, \\ \mathbf{C}_{r+2}=p_{21}\mathbf{C}_1+p_{22}\mathbf{C}_2+\dots+p_{2r}\mathbf{C}_r, \\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots \\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots \\ \mathbf{C}_n=p_{k1}\mathbf{C}_1+p_{k2}\mathbf{C}_2+\dots+p_{kr}\mathbf{C}_r, \text{ where } k=n-r. \end{array}\right\} \quad \dots(2)$$

The relations (2) can be written in the form

$$\left.\begin{array}{l} p_{11}\mathbf{C}_1+p_{12}\mathbf{C}_2+\dots+p_{1r}\mathbf{C}_r-1.\mathbf{C}_{r+1}+0.\mathbf{C}_{r+2}+\dots+0.\mathbf{C}_n=\mathbf{O}, \\ p_{21}\mathbf{C}_1+p_{22}\mathbf{C}_2+\dots+p_{2r}\mathbf{C}_r+0.\mathbf{C}_{r+1}-1.\mathbf{C}_{r+2}+\dots+0.\mathbf{C}_n=\mathbf{O}, \\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots \\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots \\ p_{k1}\mathbf{C}_1+p_{k2}\mathbf{C}_2+\dots+p_{kr}\mathbf{C}_r+0.\mathbf{C}_{r+1}+0.\mathbf{C}_{r+2}+\dots-1.\mathbf{C}_n=\mathbf{O}. \end{array}\right\} \quad \dots(3)$$

Comparing (1) and (3), we find that the vectors

$$\mathbf{X}_1=\begin{bmatrix} p_{11} \\ p_{12} \\ \dots \\ p_{1r} \\ -1 \\ 0 \\ 0 \\ \dots \\ \dots \\ 0 \end{bmatrix}, \ \mathbf{X}_2=\begin{bmatrix} p_{21} \\ p_{22} \\ \dots \\ p_{2r} \\ 0 \\ -1 \\ 0 \\ \dots \\ \dots \\ 0 \end{bmatrix}, \dots, \ \mathbf{X}_{n-r}=\begin{bmatrix} p_{k1} \\ p_{k2} \\ \dots \\ p_{kr} \\ 0 \\ 0 \\ 0 \\ \dots \\ \dots \\ -1 \end{bmatrix}$$

are $(n-r)$ solutions of the equation $\mathbf{AX}=\mathbf{O}$.

The vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-r}$ form a linearly independent set. For, if we have a relation of the type

$$l_1\mathbf{X}_1+l_2\mathbf{X}_2+\dots+l_{n-r}\mathbf{X}_{n-r}=\mathbf{O}, \quad \dots(4)$$

then comparing the $(r+1)^{th}, (r+2)^{th}, \dots, n^{th}$ components on both sides of (4), we get

$$-l_1=0, -l_2=0, \dots, -l_{n-r}=0,$$

i.e. the vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-r}$ are linearly independent.

It can now be easily seen that every solution of the equation $\mathbf{AX}=\mathbf{O}$ is some suitable linear combination of these $n-r$ solutions $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{n-r}$.

Suppose the vector **X**, with components $x_1, x_2, \dots, x_n$ is any solution of the equation $\mathbf{AX}=\mathbf{O}$. Then the vector

$$\mathbf{X}+x_{r+1}\mathbf{X}_1+x_{r+2}\mathbf{X}_2+\dots+x_n\mathbf{X}_{n-r}, \quad \dots(5)$$

which, being a linear combination of solutions, is also a solution. It is quite obvious that the last $n-r$ components of the vector (5) are all equal to zero. Let $z_1, z_2, \dots, z_r$ be the first r components of the vector (5). Then the vector whose components are $(z_1, z_2, \dots, z_r, 0, 0, \dots, 0)$ is a solution of the equation $\mathbf{AX}=\mathbf{O}$.

Therefore from (1), we have

$$z_1C_1+z_2C_2+\ldots+z_rC_r=O.$$

But the vectors $C_1, C_2, \ldots, C_r$ are linearly independent. Therefore, we have $z_1=0, z_2=0, \ldots, z_r=0$. Hence (5) is a zero vector. Therefore

$$X=-x_{r+1}X_1-x_{r+2}X_2-\ldots-x_nX_{n-r}.$$

Thus every solution X is a linear combination of the $n-r$ linearly independent solutions $X_1, X_2, \ldots, X_{n-r}$.

Therefore the set of solutions $\{X_1, X_2, \ldots, X_{n-r}\}$ *forms a basis of the vector space of all the solutions of the system of equations* $AX=O$. (I.A.S. 1970)

§ 2. Some Important conclusions about the nature of solutions of the equation AX=O.

Suppose we have m equations in n unknowns. Then the coefficient matrix A will be of the type $m \times n$. Let r be the rank of the matrix A. Obviously r cannot be greater than n (the number of columns of the matrix A). Therefore we have either $r=n$ or $r<n$.

Case I. If $r=n$, the equation $AX=O$ will have $n-n$ *i.e.*, no linearly independent solutions. In this case the zero solution will be the only solution. We know that zero vector forms a linearly dependent set.

Case II. If $r<n$, we shall have $n-r$ linearly independent solutions. Any linear combination of these $n-r$ solutions will also be a solution of $AX=O$. Thus in this case the equation $AX=O$ will have an infinite number of solutions.

Case III. Suppose $m<n$ *i.e.*, the number of solutions is less than the number of unknowns. Since $r \leqslant m$, therefore r is definitely less than n. Hence in this case the given system of equations must possess a non-zero solution. The number of solutions of the equation $AX=O$ will be infinite.

§ 3. Fundamental set of solutions of the equation AX=O. Suppose the rank r of the coefficient matrix A is less than the number of unknowns n. In this case the given equations have a set of $n-r$ linearly independent solutions and every possible solution is a linear combination of these $n-r$ solutions. This set of $n-r$ solutions is called a fundamental set of solutions of the equation $AX=O$.

Definition. *A set of linearly independent solutions* $X_1, X_2, \ldots, X_k$ *of the system of homogeneous equations* $AX=O$ *is called the fundamental system of solutions of* $AX=O$, *if every solution* X *of* $AX=O$ *can be written as a linear combination of these vectors i.e., in the form*

$$X=c_1X_1+c_2X_2+\ldots+c_kX_k,$$

where $c_1, c_2, \ldots, c_k$ *are suitable numbers.*

§ 4. Working rule for finding the solutions of the equation AX=O. Reduce the coefficient matrix A to **Echelon form** by applying **elementary row transformations only.** This Echelon form will help us to know the rank of the matrix A. Suppose the matrix A is of the type $m \times n$ and its rank comes out to be r. If $r<m$, then in the process of reducing the matrix A to Echelon form, $(m-r)$ equations will be eliminated. The given system of m equations will thus be replaced by an equivalent system of r equations. Solving these r equations (by Cramer's rule or otherwise), we can express the values of some r unknowns in terms of the remaining $n-r$ unknowns. These $n-r$ unknowns can be given any arbitrarily chosen values.

If $r=n$, the zero solution (trivial solution) will be the only solution. If $r<n$, there will be an infinity of solutions.

Solved Examples

Ex. 1. *Does the following system of equations possess a common non-zero solution?*

$$x+2y+3z=0,$$
$$3x+4y+4z=0,$$
$$7x+10y+12z=0.$$

Solution. The given system of equations can be written in the form of the single matrix equation

$$AX=\begin{bmatrix}1 & 2 & 3\\3 & 4 & 4\\7 & 10 & 12\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}=O.$$

We shall start reducing the coefficient matrix A to triangular form by applying only E-row transformations on it. Applying $R_2 \to R_2-3R_1$, $R_3 \to R_3-7R_1$, the given system of equations is equivalent to

$$\begin{bmatrix}1 & 2 & 3\\0 & -2 & -5\\0 & -4 & -9\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=O.$$

Here we find that the determinant of the matrix on the left hand side of this equation is not equal to zero. Therefore the rank of this matrix is 3. So there is no need of further applying E-row transformations on the coefficient matrix. The rank of the coefficient matrix A is 3, *i.e.*, equal to the number of unknowns. Therefore the given system of equations does not possess any linearly independent solution. The zero solution, *i.e.* $x=y=z=0$ is the only solution of the given system of equations.

Ex. 2. *Solve completely the system of equations*

$$x+3y-2z=0,\ 2x-y+4z=0,\ x-11y+14z=0.$$

Solution. The given system of equations is equivalent to the single matrix equation

$$\mathbf{AX}=\begin{bmatrix}1 & 3 & -2\\ 2 & -1 & 4\\ 1 & -11 & 14\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\mathbf{O}.$$

We shall reduce the coefficient matrix **A** to Echelon form by applying only E-row operations on it. Performing $R_2\to R_2-2R_1$, $R_3\to R_3-R_1$, we have

$$\begin{bmatrix}1 & 3 & -2\\ 0 & -7 & 8\\ 0 & -14 & 16\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\mathbf{O}.$$

Performing $R_3\to R_3-2R_2$, we have

$$\begin{bmatrix}1 & 3 & -2\\ 0 & -7 & 8\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\mathbf{O}.$$

The coefficient matrix is now triangular. The coefficient matrix being of rank 2, the given system of equations possesses $3-2=1$ linearly independent solution. We shall assign arbitrary values to $n-r=3-2=1$ variable and the remaining $r=2$ variables shall be found in terms of these. The given system of equations is equivalent to

$$x+3y-2z=0,$$
$$-7y+8z=0.$$

Thus $\quad y=\frac{8}{7}z,\ x=-\frac{10}{7}z.$

Choose $\quad z=c.$

Then $\quad y=\frac{8}{7}c,\ x=-\frac{10}{7}c.$

Hence $x=-\frac{10}{7}c, y=\frac{8}{7}c, z=c$ constitute the general solution of the given system, where c is an arbitrary parameter.

Ex. 3. *Solve completely the system of equations*

$$x+y+z=0,$$
$$2x-y-3z=0,$$

$$3x-5y+4z=0,$$
$$x+17y+4z=0.$$ **(Delhi 1970)**

Solution. The given system of equations is equivalent to the single matrix equation

$$\mathbf{AX}=\begin{bmatrix}1 & 1 & 1\\ 2 & -1 & -3\\ 3 & -5 & 4\\ 1 & 17 & 4\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\mathbf{O}.$$

We shall first find the rank of the coefficient matrix **A** by reducing it to Echelon form by applying elementary row transformations only. Applying $R_2\to R_2-2R_1$, $R_3\to R_3-3R_1$, $R_4\to R_4-R_1$, we get

$$\mathbf{A}\sim\begin{bmatrix}1 & 1 & 1\\ 0 & -3 & -5\\ 0 & -8 & 1\\ 0 & 16 & 3\end{bmatrix}$$

$$\sim\begin{bmatrix}1 & 1 & 1\\ 0 & -3 & -5\\ 0 & -24 & 3\\ 0 & 48 & 9\end{bmatrix}\text{ by } R_3\to 3R_3,\ R_4\to 3R_4$$

$$\sim\begin{bmatrix}1 & 1 & 1\\ 0 & -3 & -5\\ 0 & 0 & 43\\ 0 & 0 & -71\end{bmatrix}\text{ by } R_3\to R_3-8R_2,\ R_4\to R_4+16R_2.$$

$$\sim\begin{bmatrix}1 & 1 & 1\\ 0 & -3 & -5\\ 0 & 0 & 43\\ 0 & 0 & 0\end{bmatrix}\text{ by } R_4\to R_4+\tfrac{71}{43}R_3.$$

Above is the Echelon form of the coefficient matrix **A**. We have rank **A** = the number of non-zero rows in this Echelon form = 3. The number of unknowns is also 3. Since rank **A** is equal to the number of unknowns, therefore the given system of equations possesses no non-zero solution. Hence the zero solution *i.e.* $x=y=z=0$ is the only solution of the given system of equations.

Ex. 4. *Solve completely the system of equations*

$$2x-2y+5z+3w=0,$$
$$4x-y+z+w=0,$$
$$3x-2y+3z+4w=0,$$
$$x-3y+7z+6w=0.$$ **(Kanpur 1970)**

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix}2 & -2 & 5 & 3\\ 4 & -1 & 1 & 1\\ 3 & -2 & 3 & 4\\ 1 & -3 & 7 & 6\end{bmatrix}\begin{bmatrix}x\\ y\\ z\\ w\end{bmatrix}=O.$$

We shall first find the rank of the coefficient matrix A by reducing it to Echelon form by applying elementary row transformations only. Applying $R_1\leftrightarrow R_4$, we get

$$A\sim\begin{bmatrix}1 & -3 & 7 & 6\\ 4 & -1 & 1 & 1\\ 3 & -2 & 3 & 4\\ 2 & -2 & 5 & 3\end{bmatrix}$$

$$\sim\begin{bmatrix}1 & -3 & 7 & 6\\ 0 & 11 & -27 & -23\\ 0 & 7 & -18 & -14\\ 0 & 4 & -9 & -9\end{bmatrix} \text{ by } R_2\to R_2-4R_1,\ R_3\to R_3-3R_1,\ R_4\to R_4-2R_1$$

$$\sim\begin{bmatrix}1 & -3 & 7 & 6\\ 0 & 4 & -9 & -9\\ 0 & 7 & -18 & -14\\ 0 & 4 & -9 & -9\end{bmatrix} \text{ by } R_2\to R_2-R_3$$

$$\sim\begin{bmatrix}1 & -3 & 7 & 6\\ 0 & 4 & -9 & -9\\ 0 & 28 & -72 & -56\\ 0 & 4 & -9 & -9\end{bmatrix} \text{ by } R_3\to 4R_3$$

$$\sim\begin{bmatrix}1 & -3 & 7 & 6\\ 0 & 4 & -9 & -9\\ 0 & 0 & -9 & 7\\ 0 & 0 & 0 & 0\end{bmatrix} \text{ by } R_3\to R_3-7R_2,\ R_4\to R_4-R_2.$$

Above is the Echelon form of the coefficient matrix A. We have rank A = the number of non-zero rows in this Echelon form = 3. The number of unknowns is 4. Since rank A is less than the number of unknowns, therefore the given system of equations possesses non-zero solutions. The given system of

shall assign arbitrary values to $n-r=4-3=1$ variable and the remaining $r=3$ variables shall be found in terms of these. The given system of equations is equivalent to

$$\begin{bmatrix}1 & -3 & 7 & 6\\ 0 & 4 & -9 & -9\\ 0 & 0 & -9 & 7\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x\\ y\\ z\\ w\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}.$$

This gives the equations

$$x-3y+7z+6w=0,$$
$$4y-9z-9w=0,$$
$$-9z+7w=0.$$

From these, we get

$$z=\tfrac{7}{9}w,\ y=\tfrac{9}{4}z+\tfrac{9}{4}w=\tfrac{9}{4}(\tfrac{7}{9}w)+\tfrac{9}{4}w=\tfrac{7}{4}w+\tfrac{9}{4}w=4w,$$
$$x=3y-7z-6w=3\,(4w)-7\,(\tfrac{7}{9}w)-6w$$
$$=12w-\tfrac{49}{9}w-6w=6w-\tfrac{49}{9}w=\tfrac{5}{9}w.$$

Taking $w=c$, we see that $x=\frac{5}{9}c$, $y=4c$, $z=\frac{7}{9}c$, $w=c$ constitute the general solution of the given system.

Ex. 5. *Find all the solutions of the following system of equations :*

$$3x+4y-z-6w=0,$$
$$2x+3y+2z-3w=0,$$
$$2x+y-14z-9w=0,$$
$$x+3y+13z+3w=0.$$

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix}3 & 4 & -1 & -6\\ 2 & 3 & 2 & -3\\ 2 & 1 & -14 & -9\\ 1 & 3 & 13 & 3\end{bmatrix}\begin{bmatrix}x\\ y\\ z\\ w\end{bmatrix}=O.$$

We shall first find the rank of the coefficient matrix A by reducing it to Echelon form by applying *E*-row transformations only.

Applying $R_4\leftrightarrow R_1$, we get

$$A\sim\begin{bmatrix}1 & 3 & 13 & 3\\ 2 & 3 & 2 & -3\\ 2 & 1 & -14 & -9\\ 3 & 4 & -1 & -6\end{bmatrix}$$

$$\sim\begin{bmatrix}1 & 3 & 13 & 3\\0 & -3 & -24 & -9\\0 & -5 & -40 & -15\\0 & -5 & -40 & -15\end{bmatrix} \text{ by } R_2\to R_2-2R_1,\ R_3\to R_3-2R_1,\ R_4\to R_4-3R_1$$

$$\sim\begin{bmatrix}1 & 3 & 13 & 3\\0 & 1 & 8 & 3\\0 & 1 & 8 & 3\\0 & 1 & 8 & 3\end{bmatrix} \text{ by } R_2\to -\tfrac{1}{3}R_2,\ R_3\to -\tfrac{1}{5}R_3,\ R_4\to -\tfrac{1}{5}R_4.$$

$$\sim\begin{bmatrix}1 & 3 & 13 & 3\\0 & 1 & 8 & 3\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix} \text{ by } R_3\to R_3-R_2,\ R_4\to R_4-R_2.$$

The rank of A is obviously 2 which is less than the number of unknowns 4. Therefore the given system of equations possesses $4-2$ *i.e.*, 2 linearly independent solutions. The given system of equations is equivalent to the equation

$$\begin{bmatrix}1 & 3 & 13 & 3\\0 & 1 & 8 & 3\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\\w\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

Thus the given system of four equations is equivalent to the system of two equations, *i.e.*

$$\left.\begin{aligned}x+3y+13z+3w=0,\\ y+8z+3w=0\end{aligned}\right\}.$$

From these equations, we get

$$y=-8z-3w,\ x=-3(-8z-3w)-13z-3w$$

i.e. $\quad y=-8z-3w,\ x=11z+6w.$

Hence $x=11c_1+6c_2,\ y=-8c_1-3c_2,\ z=c_1,\ w=c_2$ constitute the general solution of the given system of equations, where c_1 and c_2 are arbitrary numbers. Since we can give any arbitrary values to c_1 and c_2 therefore the given system of equations has an infinite number of solutions.

Ex. 6. *Solve completely the system of equations*

$$4x+2y+z+3u=0,$$
$$6x+3y+4z+7u=0,$$
$$2x+y+u=0. \qquad \textbf{(Meerut 1976)}$$

Solution. The matrix form of the given system is

$$\begin{bmatrix}4 & 2 & 1 & 3\\6 & 3 & 4 & 7\\2 & 1 & 0 & 1\end{bmatrix}\begin{bmatrix}x\\y\\z\\u\end{bmatrix}=\mathbf{O}$$

or $$\begin{bmatrix}1 & 2 & 4 & 3\\4 & 3 & 6 & 7\\0 & 1 & 2 & 1\end{bmatrix}\begin{bmatrix}z\\y\\x\\u\end{bmatrix}=\mathbf{O},$$ **interchanging the variables x and z.**

Performing $R_2\to R_2-4R_1$, we get

$$\begin{bmatrix}1 & 2 & 4 & 3\\0 & -5 & -10 & -5\\0 & 1 & 2 & 1\end{bmatrix}\begin{bmatrix}z\\y\\x\\u\end{bmatrix}=\mathbf{O}.$$

Performing $R_2\to -\frac{1}{5}R_2$, we get

$$\begin{bmatrix}1 & 2 & 4 & 3\\0 & 1 & 2 & 1\\0 & 1 & 2 & 1\end{bmatrix}\begin{bmatrix}z\\y\\x\\u\end{bmatrix}=\mathbf{O}.$$

Performing $R_3\to R_3-R_2$, we get

$$\begin{bmatrix}1 & 2 & 4 & 3\\0 & 1 & 2 & 1\\0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}z\\y\\x\\u\end{bmatrix}=\mathbf{O}.$$

The coefficient matrix is of rank 2 and therefore the given system will have $4-2$ *i.e.* 2 linearly independent solutions. The given system of equations is equivalent to

$$z+2y+4x+3u=0,$$
$$y+2x+u=0.$$

$$\therefore\quad y=-2x-u,\ z=-4x-3u+4x+2u=-u.$$

$$\therefore\quad x=c_1,\ u=c_2,\ y=-2c_1-c_2,\ z=-c_2$$

constitute the general solution where c_1 and c_2 are arbitrary constants.

Ex. 7. *Prove that a necessary and sufficient condition that values, not all zero, may be assigned to the n variables $x_1, x_2, \ldots, x_n$ so that the n homogeneous equations*

$$a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n=0 \qquad (i=1, 2, \ldots, n)$$

hold simultaneously, is that the determinant

$$|a_{ij}|_{n\times n}=0. \qquad \text{(I. A. S. 1969, Allahabad 82)}$$

Solution. Let $A=[a_{ij}]_{n\times n}=$ the coefficient matrix of the equations.

The condition is necessary. Suppose the given system of equations possesses a non-zero solution. Then we are to prove that the determinant $|a_{ij}|=0$. We shall prove it by contradiction. Let $|a_{ij}|\neq 0$. Then rank $A=n$. Consequently the number of linearly independent solutions of the given equations $=n-n=0$. Thus the given equations possess no linearly independent solution *i.e.*, the zero solution is the only solution. But this contradicts our hypothesis that the given equations possess a non-zero solution. Hence we must have $|a_{ij}|=0$.

The condition is sufficient. Let the determinant $|a_{ij}|=0$. Then we are to prove that the equations must possess a non-zero solution. Since $|a_{ij}|=0$, therefore rank $A<n$. Let rank $A=m$. Then the given equations have $n-m$ linearly independent solutions. But a linearly independent solution can never be the zero solution. Therefore the given equations must have a non-zero solution.

Ex. 8. *Discuss for all values of k the system of equations*

$$\begin{aligned} 2x+3ky+(3k+4)\,z&=0,\\ x+(k+4)\,y+(4k+2)\,z&=0,\\ x+2\,(k+1)\,y+(3k+4)\,z&=0. \end{aligned}$$

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix} 2 & 3k & 3k+4\\ 1 & k+4 & 4k+2\\ 1 & 2k+2 & 3k+4 \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix}=O.$$

Performing $R_1\leftrightarrow R_2$, we have

$$\begin{bmatrix} 1 & k+4 & 4k+2\\ 2 & 3k & 3k+4\\ 1 & 2k+2 & 3k+4 \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix}=O.$$

Performing $R_2\to R_2-2R_1$, $R_3\to R_3-R_1$, we have

$$\begin{bmatrix} 1 & k+4 & 4k+2\\ 0 & k-8 & -5k\\ 0 & k-2 & -k+2 \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix}=O. \qquad \ldots\text{(i)}$$

If the given system of equations is to possess any linearly independent solution the coefficient matrix A must be of rank less than 3. For the matrix A to be of rank less than 3, we must have

$$(k-8)\,(-k+2)+5k\,(k-2)=0$$

i.e., $\quad -k^2+2k+8k-16+5k^2-10k=0$

i.e., $\quad 4k^2-16=0$

i.e., $\quad k=\pm 2.$

Now three cases arise.

Case I. When $k\neq\pm 2$, the given system of equations possesses no linearly independent solution and $x=y=z=0$ is the only solution.

Case II. If $k=2$, the equation (i) reduces to

$$\begin{bmatrix} 1 & 6 & 10\\ 0 & -6 & -10\\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix}=O.$$

The coefficient matrix being of rank 2, the given system of equations now possesses $3-2=1$ linearly independent solution. The given system of equations is now equivalent to

$$-6y-10z=0,\ x+6y+10z=0.$$

Thus $\quad y=-\frac{5}{3}z,\ x=0.$

Hence $\quad x=0,\ y=-\frac{5}{3}c,\ z=c$

or $x=0$, $y=-5k$, $z=3k$ constitute the general solution of the given system where k is an arbitrary parameter.

Case III. If $k=-2$, the equation (i) reduces to

$$\begin{bmatrix} 1 & 2 & -6\\ 0 & -10 & 10\\ 0 & -4 & 4 \end{bmatrix}\begin{bmatrix} x\\ y\\ z \end{bmatrix}=O.$$

The coefficient matrix being of rank 2, the given system of equations now possesses $3-2=1$ linearly independent solution. The given system of equations is now equivalent to

$$-4y+4z=0,\ -10y+10z=0,\ x+2y-6z=0.$$

Thus $\quad y=z,\ x=4z.$

Hence $x=4c$, $y=c$, $z=c$ constitute the general solution of the given system where c is an arbitrary parameter.

Ex. 9. *Show that the only real value of λ for which the following equations have non-zero solution is 6 :*

$$x+2y+3z=\lambda x,\ 3x+y+2z=\lambda y,\ 2x+3y+z=\lambda z.$$

Solution. The given system of equations is equivalent to the single matrix equation

$$\mathbf{AX}=\begin{bmatrix}1-\lambda & 2 & 3\\ 3 & 1-\lambda & 2\\ 2 & 3 & 1-\lambda\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\mathbf{O}.$$

If the given system of equations is to possess a non-zero solution, the coefficient matrix **A** must be of rank less than 3. If the matrix **A** is to be of rank less than 3 its determinant must be equal to zero. Thus we must have

$$\begin{vmatrix}1-\lambda & 2 & 3\\ 3 & 1-\lambda & 2\\ 2 & 3 & 1-\lambda\end{vmatrix}=0$$

or $\begin{vmatrix}6-\lambda & 6-\lambda & 6-\lambda\\ 3 & 1-\lambda & 2\\ 2 & 3 & 1-\lambda\end{vmatrix}=0$, adding the second and the third rows to the first

or $(6-\lambda)\begin{vmatrix}1 & 1 & 1\\ 3 & 1-\lambda & 2\\ 2 & 3 & 1-\lambda\end{vmatrix}=0$

or $(6-\lambda)\begin{vmatrix}1 & 0 & 0\\ 3 & -\lambda-2 & -1\\ 1 & 1 & -\lambda-1\end{vmatrix}=0 \quad C_2-C_1,\ C_3-C_1$

or $(6-\lambda)\,[(\lambda+2)\,(\lambda+1)+1]=0,$

or $(6-\lambda)\,[\lambda^2+3\lambda+3]=0.$

The roots of the equation $\lambda^2+3\lambda+3=0$ are

$$\lambda=\frac{-3\pm\sqrt{(9-12)}}{2}\ \text{ i.e., are imaginary.}$$

Hence the only real value of λ for which the system of equations is to have a non-zero solution is 6.

Exercises

Find all the solutions of the following systems of linear homogeneous equations.

1. $2x-3y+z=0,\ x+2y-3z=0,\ 4x-y-2z=0.$
2. $x+y-3z+2w=0,$
$2x-y+2z-3w=0,$
$3x-2y+z-4w=0,$
$-4x+y-3z+w=0.$
3. $x+y+z=0,\ 2x+5y+7z=0,\ 2x-5y+3z=0.$ **(Poona 1970)**
4. $x+2y+3z=0,\ 2x+3y+4z=0,\ 7x+13y+19z=0.$
5. $x-2y+z-w=0,$
$x+y-2z+3w=0,$
$4x+y-5z+8w=0,$
$5x-7y+2z-w=0.$ **(Meerut 1984)**

Answers

1. $x=0,\ y=0,\ z=0.$ 2. $x=0,\ y=0,\ z=0,\ w=0.$
3. $x=0,\ y=0,\ z=0.$ 4. $x=c,\ y=-2c,\ z=c.$
5. $x=c_1-\frac{5}{3}c_2,\ y=c_1-\frac{4}{3}c_2,\ z=c_1,\ w=c_2.$

§ 5. Systems of linear Non-homogeneous equations. Sometimes we think that we can solve every two simultaneous equations of the type

$$\left.\begin{aligned}a_1x+b_1y=c_1\\ a_2x+b_2y=c_2\end{aligned}\right\}.$$

But it is not so. For example, consider the simultaneous equations

$$\left.\begin{aligned}3x+4y=5\\ 6x+8y=13\end{aligned}\right\}.$$

There is no set of values of x and y which satisfies both these equations. Such equations are said to be **inconsistent.**

Let us take another example. Consider the simultaneous equations

$$\left.\begin{aligned}3x+4y=5\\ 6x+8y=10\end{aligned}\right\}.$$

These equations are consistent since there exist values of x and y which satisfy both of these equations. We see that $x=-\frac{4}{3}c+\frac{5}{3}$, $y=c$ constitute a solution of these equations, where c is arbitrary. Thus these equations possess an infinite number of solutions.

Now we shall discuss the nature of solutions of a system of non-homogeneous linear equations.

Let

$$\left.\begin{aligned}a_{11}x_1+a_{12}x_2+\ldots+a_{1n}x_n&=b_1\\ a_{21}x_1+a_{22}x_2+\ldots+a_{2n}x_n&=b_2\\ \ldots\quad\ldots\quad\ldots\quad&\ldots\\ a_{m1}x_1+a_{m2}x_2+\ldots+a_{mn}x_n&=b_m\end{aligned}\right\}\quad\ldots(1)$$

be a system of m non-homogeneous equations in n unknowns $x_1, x_2, \ldots, x_n$.

If we write

$$A=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}_{m\times n}, \quad X=\begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix}_{n\times 1}, \quad B=\begin{bmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{bmatrix}_{m\times 1},$$

where **A**, **X**, **B** are $m\times n$, $n\times 1$ and $m\times 1$ matrices respectively, the above equations can be written in the form of a single matrix equation **AX=B**.

Any set of values of $x_1, x_2, \ldots, x_n$ which simultaneously satisfy all these equations is called a solution of the system (1). When the system of equations has one or more solutions, the equations are said to be **consistent**, *otherwise they are said to be* **inconsistent.**

The matrix

$$[A \quad B]=\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{bmatrix}$$

is called the **augmented matrix** of the given system of equations.

§ 6. Condition for Consistency. Theorem. *The system of equations* **AX=B** *is consistent i.e., possesses a solution, if and only if the coefficient matrix* **A** *and the augmented matrix* **[A B]** *are of the same rank.* **(Nagarjuna 1990; Andhra 90; Kanpur 86; Meerut 90; Gujrat 70; Allahabad 76; Madras 81, I.A.S. 73)**

Proof. Let $C_1, C_2, \ldots, C_n$ denote the column vectors of the matrix **A**. The equation **AX=B** is then equivalent to

$$[C_1 \; C_2 \ldots C_n]\begin{bmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{bmatrix}=B$$

i.e., $x_1C_1+x_2C_2+\ldots+x_nC_n=B.$ ...(1)

Let now r be the rank of the matrix **A**. The matrix **A** has then r linearly independent columns and, without loss of generality, we can suppose that the first r columns $C_1, C_2, \ldots, C_r$ form a linearly independent set so that each of the remaining $n-r$ columns is a linear combination of these r columns.

The condition is necessary. If the given system of equations is consistent, there must exist n scalars (numbers) $k_1, k_2, \ldots, k_n$ such that

$$k_1C_1+k_2C_2+\ldots+k_nC_n=B. \quad \ldots(2)$$

Since each of the $n-r$ columns $C_{r+1}, C_{r+2}, \ldots, C_n$ is a linear combination of first r columns $C_1, C_2, \ldots, C_r$ it is obvious from (2) that **B** is also a linear combination of $C_1, C_2, \ldots, C_r$. Thus the maximum number of linearly independent columns of the matrix [A B] is also r. Therefore the matrix [A B] is also of rank r. Hence the matrices **A** and [A B] are of the same rank.

The condition is sufficient. Now suppose that the matrices **A** and [A B] are of the same rank r. The maximum number of linearly independent columns of the matrix [A B] is then r. But the first r columns $C_1, C_2, \ldots, C_r$ of the matrix [A B] already form a linearly independent set. Therefore the column **B** should be expressed as a linear combination of the columns $C_1, C_2, \ldots, C_r$.

Thus there exist r scalars $k_1, k_2, \ldots, k_r$ such that

$$k_1C_1+k_2C_2+\ldots+k_rC_r=B. \quad \ldots(3)$$

Now (3) may be written as

$$k_1C_1+k_2C_2+\ldots+k_rC_r+0\,C_{r+1}+0.C_{r+2}+\ldots+0.C_n=B. \quad \ldots(4)$$

Comparing (1) and (4), we see that

$x_1=k_1, x_2=k_2, \ldots x_r=k_r, x_{r+1}=0, x_{r+2}=0, \ldots, x_n=0$

constitute a solution of the equation **AX=B**.

Therefore the given system of equations is consistent.

§ 7. Condition for a system of n-equations in n-unknowns to have a unique solution.

Theorem. *If* **A** *be an n-rowed non-singular matrix,* **X** *be an $n\times 1$ matrix,* **B** *be an $n\times 1$ matrix, the system of equations* **AX=B** *has a unique solution.* **(Andhra 1990)**

Proof. If **A** be an n-rowed non-singular matrix, the ranks of the matrices **A** and [A B] are both n. Therefore the system of equations **AX=B** is consistent *i.e.*, possesses a solution.

Pre-multiplying both sides of **AX=B** by A^{-1}, we have

$$A^{-1}AX=A^{-1}B$$

i.e., $IX=A^{-1}B$

i.e., $X=A^{-1}B$

is a solution of the equation **AX=B**.

To show that the solution is unique, let us suppose that X_1 and X_2 be two solutions of **AX=B**.

Then $AX_1=B,\ AX_2=B \Rightarrow AX_1=AX_2 \Rightarrow A^{-1}AX_1=A^{-1}AX_2$
$\Rightarrow IX_1=IX_2 \Rightarrow X_1=X_2.$

Hence the solution is unique.

§ 8. Working Rule for finding the solution of the equation AX=B Suppose the coefficient matrix A is of the type $m \times n$, *i.e.*, we have m equations in n unknowns. Write the augmented matrix [A B] and reduce it to a Echelon form by applying only *E*-row transformations on it. This Echelon form will enable us to know the ranks of the augmented matrix [A B] and the coefficient matrix A. Then the following different cases arise:

Case I. Rank A < Rank [A B].

In this case the equations AX=B are inconsistent *i.e.*, they have no solution.

Case II. Rank A=Rank [A B]=r (say).

In this case the equations AX=B are consistent *i.e.*, they possess a solution. If $r < m$, then in the process of reducing the matrix [A B] to Echelon form, $(m-r)$ equations will be eliminated. The given system of m equations will then be replaced by an equivalent system of r equations. From these r equations we shall be able to express the values of some r unknowns in terms of the remaining $n-r$ unknowns which can be given any arbitrary chosen values.

If $r=n$, then $n-r=0$, so that no variable is to be assigned arbitrary values and therefore in this case there will be a unique solution.

If $r < n$, then $n-r$ variables can be assigned arbitrary values. So in this case there will be an infinite number of solutions. Only $n-r+1$ solutions will be linearly independent and the rest of the solutions will be linear combinations of them.

If $m < n$, then $r \leqslant m < n$. Thus in this case $n-r > 0$. Therefore when the number of equations is less than the number of unknowns, the equations will always have an infinite number of solutions, provided they are consistent.

Solved Examples

Ex. 1. *Show that the equations*

$$2x+6y+11=0,$$
$$6x+20y-6z+3=0,$$
$$6y-18z+1=0$$

are not consistent.

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix}2 & 6 & 0\\ 6 & 20 & -6\\ 0 & 6 & -18\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-11\\ -3\\ -1\end{bmatrix}=B.$$

We shall reduce the coefficient matrix A to triangular form by *E*-row operations on it and apply the same operations on the right hand side *i.e.* on the matrix B.

Performing $R_2 \to R_2-3R_1$, we have

$$\begin{bmatrix}2 & 6 & 3\\ 0 & 2 & -6\\ 0 & 6 & -18\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-11\\ 30\\ -1\end{bmatrix}.$$

Performing $R_3 \to R_3-3R_2$, we have

$$\begin{bmatrix}2 & 6 & 0\\ 0 & 2 & -6\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-11\\ 30\\ -91\end{bmatrix}.$$

The last equation of this system is $0x+0y+0z=-91$. This shows that the given system is not consistent.

Ex. 2. *Show that the equations*

$$x+y+z=-3,\ 3x+y-2z=-2,\ 2x+4y+7z=7$$

are not consistent. **(Meerut 1983)**

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix}1 & 1 & 1\\ 3 & 1 & -2\\ 2 & 4 & 7\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}-3\\ -2\\ 7\end{bmatrix}=B.$$

The augmented matrix

$$[A\ \ B]=\left[\begin{array}{ccc:c}1 & 1 & 1 & -3\\ 3 & 1 & -2 & -2\\ 2 & 4 & 7 & 7\end{array}\right].$$

We shall reduce the augmented matrix [A B] to Echelon form by applying *E*-row transformations only. Applying $R_2 \to R_2-3R_1,\ R_3 \to R_3-2R_1$, we get

$$[A\ \ B] \sim \left[\begin{array}{ccc:c}1 & 1 & 1 & -3\\ 0 & -2 & -5 & 7\\ 0 & 2 & 5 & 13\end{array}\right]$$

$$\sim \left[\begin{array}{ccc:c}1 & 1 & 1 & -3\\ 0 & -2 & -5 & 7\\ 0 & 0 & 0 & 20\end{array}\right], \text{ applying } R_3 \to R_3+R_2.$$

Above is the Echelon form of the matrix [A B]. We have rank [A B] = the number of non-zero rows in this Echelon form $=3$.

Also by the same E-row transformations, we get

$$A \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -5 \\ 0 & 0 & 0 \end{bmatrix}.$$

Obviously rank $A=2$.

Since Rank $A \neq$ Rank [A B], therefore the given equations are inconsistent *i.e.*, they have no solution.

Ex. 3. *Show that the equations*

$x+y+z=6,\ x+2y+3z=14,\ x+4y+7z=30$

are consistent and solve them. **(Meerut 1983; Agra 73)**

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 6 \\ 14 \\ 30 \end{bmatrix}=B.$$

The augmented matrix

$$[A\ \ B]=\left[\begin{array}{ccc:c} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 14 \\ 1 & 4 & 7 & 30 \end{array}\right]$$

We shall reduce the augmented matrix [A B] to Echelon form by applying elementary row transformations only. Applying $R_2 \to R_2 - R_1$, $R_3 \to R_3 - R_1$, we get

$$[A\ \ B] \sim \left[\begin{array}{ccc:c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 3 & 6 & 24 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ by } R_3 \to R_3 - 3R_2.$$

Above is the Echelon form of the matrix [A B]. We have rank [A B] = the number of non-zero rows in this Echelon form $=2$.

By the same elementary transformations, we get

$$A \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Obviously rank $A=2$. Since rank A = rank [A B], therefore the given equations are consistent. Here the number of unknowns

is 3. Since rank A is less than the number of unknowns, therefore the given system will have an infinite number of solutions. We see that the given system of equations is equivalent to the matrix equation

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 6 \\ 8 \\ 0 \end{bmatrix}.$$

This matrix equation is equivalent to the system of equations

$$\left.\begin{array}{r} x+y+z=6, \\ y+2z=8. \end{array}\right\}$$

$\therefore\ \ y=8-2z,\ x=6-y-z=6-(8-2z)-z=z-2.$

Taking $z=c$, we see that $x=c-2,\ y=8-2c,\ z=c$ constitute the general solution of the given system, where c is an arbitrary constant.

Ex. 4. *Apply the test of rank to examine if the following equations are consistent* :

$$2x-y+3z=8$$
$$-x+2y+z=4$$
$$3x+y-4z=0$$

and if consistent, find the complete solution.

(Rohilkhand 1991; Meerut 88)

Solution. The given system of equations is equivalent to the single matrix equation

$$AX=\begin{bmatrix} 2 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & -4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 8 \\ 4 \\ 0 \end{bmatrix}=B.$$

The augmented matrix

$$[A\ \ B]=\left[\begin{array}{ccc:c} 2 & -1 & 3 & 8 \\ -1 & 2 & 1 & 4 \\ 3 & 1 & -4 & 0 \end{array}\right].$$

We shall reduce the augmented matrix to Echelon form by applying elementary row transformations only. Applying $R_1 \leftrightarrow R_2$, we get

$$A \sim \left[\begin{array}{ccc:c} -1 & 2 & 1 & 4 \\ 2 & -1 & 3 & 8 \\ 3 & 1 & -4 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:c} -1 & 2 & 1 & 4 \\ 0 & 3 & 5 & 16 \\ 0 & 7 & -1 & 12 \end{array}\right] \text{ by } R_2 \to R_2 + 2R_1,\ R_3 \to R_3 + 3R_1$$

$$\sim\left[\begin{array}{rrr:r} -1 & 2 & 1 & 4 \\ 0 & 3 & 5 & 16 \\ 0 & 21 & -3 & 36 \end{array}\right], \text{ by } R_3 \to 3R_3$$

$$\sim\left[\begin{array}{rrr:r} -1 & 2 & 1 & 4 \\ 0 & 3 & 5 & 16 \\ 0 & 0 & -38 & -76 \end{array}\right], \text{ by } R_3 \to R_3 - 7R_2$$

$$\sim\left[\begin{array}{rrr:r} -1 & 2 & 1 & 4 \\ 0 & 3 & 5 & 16 \\ 0 & 0 & 1 & 2 \end{array}\right], \text{ by } R_3 \to -\tfrac{1}{38} R_3.$$

Above is the Echelon form of the matrix [A B]. We have rank [A B] = the number of non-zero rows in this Echelon form = 3.

By the same transformations, we get

$$\mathbf{A} \sim \begin{bmatrix} -1 & 2 & 1 \\ 0 & 3 & 5 \\ 0 & 0 & 1 \end{bmatrix}.$$

Obviously rank A = 2. Since rank A = rank [A B], therefore the given equations are consistent. Here the number of unknowns is 3. Since rank A is equal to the number of unknowns, therefore the given equations have a unique solution. We see that the given equations are equivalent to the matrix equation

$$\begin{bmatrix} -1 & 2 & 1 \\ 0 & 3 & 5 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 16 \\ 2 \end{bmatrix}.$$

The matrix equation is equivalent to the equations

$$-x+2y+z=4,$$
$$3y+5z=16,$$
$$z=2.$$

These give $z=2$, $y=2$, $x=2$.

Ex. 5. *Show that the equations*

$$x+2y-z=3,$$
$$3x-y+2z=1,$$
$$2x-2y+3z=2,$$
$$x-y+z=-1$$

are consistent and solve them. **(Nagarjuna 1980; Meerut 85)**

Solution. The given system of equations is equivalent to the single matrix equation

$$\mathbf{AX} = \begin{bmatrix} 1 & 2 & -1 \\ 3 & -1 & 2 \\ 2 & -2 & 3 \\ 1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 2 \\ -1 \end{bmatrix} = \mathbf{B}.$$

The augmented matrix

$$[\mathbf{A}\ \ \mathbf{B}] = \left[\begin{array}{rrr:r} 1 & 2 & -1 & 3 \\ 3 & -1 & 2 & 1 \\ 2 & -2 & 3 & 2 \\ 1 & -1 & 1 & -1 \end{array}\right].$$

Performing $R_2 \to R_2 - 3R_1$, $R_3 \to R_3 - 2R_1$, $R_4 \to R_4 - R_1$, we get

$$[\mathbf{A}\ \ \mathbf{B}] \sim \left[\begin{array}{rrr:r} 1 & 2 & -1 & 3 \\ 0 & -7 & 5 & -8 \\ 0 & -6 & 5 & -4 \\ 0 & -3 & 2 & -4 \end{array}\right]$$

$$\sim \left[\begin{array}{rrr:r} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & -6 & 5 & -4 \\ 0 & -3 & 2 & -4 \end{array}\right] \text{ by } R_2 \to R_2 - R_3$$

$$\sim \left[\begin{array}{rrr:r} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 5 & 20 \\ 0 & 0 & 2 & 8 \end{array}\right] \text{ by } R_3 \to R_3 - 6R_2,\ R_4 \to R_4 - 3R_2,$$

$$\sim \left[\begin{array}{rrr:r} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 4 \end{array}\right], \text{ by } R_3 \to \tfrac{1}{5}R_3,\ R_4 \to \tfrac{1}{2}R_4$$

$$\sim \left[\begin{array}{rrr:r} 1 & 2 & -1 & 3 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ by } R_4 \to R_4 - R_3.$$

Thus the matrix [A B] has been reduced to Echelon form. We have rank [A B] = the number of non-zero rows in this Echelon form = 3. Also

$$\mathbf{A} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

We have rank A = 3. Since rank [A B] = rank A, therefore the given equations are consistent. Since rank A = 3 = the number of unknowns, therefore the given equations have unique solution. The given equations are equivalent to the equations

$$x+2y-z=3,\quad -y=-4,\quad z=4.$$

These give $z=4$, $y=4$, $x=-1$.

Ex. 6. *State the conditions under which a system of non-homogeneous equations will have (i) no solution (ii) a unique solution (iii) infinity of solutions.* **(Meerut M.Sc. 1967)**

Solution. Let $\mathbf{AX}=\mathbf{B}$ be a system of linear non-homogeneous equations, where $\mathbf{A}, \mathbf{X}, \mathbf{B}$ are $m\times n$, $n\times 1$, $m\times 1$ matrices respectively.

(i) These equations will have no solution if the coefficient matrix $\mathbf{A}$ and the augmented matrix $[\mathbf{A}\ \ \mathbf{B}]$ are not of the same rank.

(ii) These equations will possess a unique solution if the matrices $\mathbf{A}$ and $[\mathbf{A}\ \ \mathbf{B}]$ are of the same rank and the rank is equal to the number of variables. In particular if $\mathbf{A}$ is a square matrix, these equations will possess a unique solution if and only if the matrix $\mathbf{A}$ is non-singular.

(iii) These equations will have infinity of solutions if the matrices $\mathbf{A}$ and $[\mathbf{A}\ \ \mathbf{B}]$ are of the same rank and the rank is less than the number of variables.

Ex. 7. *By the use of matrices, solve the equations :*

$$x+y+z=9,\ 2x+5y+7z=52,\ 2x+y-z=0.$$

(Meerut 1970; Agra 72)

Solution. The given system of equations is equivalent to the single matrix equation

$$\mathbf{AX}=\begin{bmatrix}1 & 1 & 1\\ 2 & 5 & 7\\ 2 & 1 & -1\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}9\\ 52\\ 0\end{bmatrix}=\mathbf{B}.$$

The augmented matrix

$$[\mathbf{A}\ \ \mathbf{B}]=\left[\begin{array}{ccc:c}1 & 1 & 1 & 9\\ 2 & 5 & 7 & 52\\ 2 & 1 & -1 & 0\end{array}\right]$$

$$\sim\left[\begin{array}{ccc:c}1 & 1 & 1 & 9\\ 0 & 3 & 5 & 34\\ 0 & -1 & -3 & -18\end{array}\right],\text{ by } R_2\to R_2-2R_1,\ R_3\to R_3-2R_1$$

$$\sim\left[\begin{array}{ccc:c}1 & 1 & 1 & 9\\ 0 & -1 & -3 & -18\\ 0 & 3 & 5 & 34\end{array}\right],\text{ by } R_2\leftrightarrow R_3$$

$$\sim\left[\begin{array}{ccc:c}1 & 1 & 1 & 9\\ 0 & -1 & -3 & -18\\ 0 & 0 & -4 & -20\end{array}\right],\text{ by } R_3\to R_3+3R_2.$$

Above is the Echelon form of the matrix $[\mathbf{A}\ \ \mathbf{B}]$. We have rank $[\mathbf{A}\ \ \mathbf{B}]$ = the number of non-zero rows in this Echelon form $=3$.

Also by the same E-row transformations, we get

$$\mathbf{A}\sim\begin{bmatrix}1 & 1 & 1\\ 0 & -1 & -3\\ 0 & 0 & -4\end{bmatrix}.$$

$\therefore$ rank $\mathbf{A}=3$.

Since rank $\mathbf{A}$ = rank $[\mathbf{A}\ \ \mathbf{B}]$, therefore the given equations are consistent. Also rank $\mathbf{A}=3$ and the number of unknowns is also 3. Hence the given equations will have a unique solution. To find the solution, we see that the given system of equations is equivalent to the matrix equation

$$\begin{bmatrix}1 & 1 & 1\\ 0 & -1 & -3\\ 0 & 0 & -4\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}9\\ -18\\ -20\end{bmatrix}.$$

The matrix equation is equivalent to the system of equations

$$x+y+z=9,\ -y-3z=-18,\ -4z=-20.$$

Solving these, we get $z=5$, $y=3$, $x=1$.

Ex. 8. *Investigate for what values of λ, μ the simultaneous equations*

$$x+y+z=6,\ x+2y+3z=10,\ x+2y+\lambda z=\mu$$

have (i) no solution, (ii) a unique solution, (iii) an infinite number of solutions. **(Meerut 1990; I.A.S. 81; Kerala 70; Rohilkhand 90; Agra 79; Kanpur 85)**

Solution. The matrix form of the given system of equations is

$$\mathbf{AX}=\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 3\\ 1 & 2 & \lambda\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix}=\begin{bmatrix}6\\ 10\\ \mu\end{bmatrix}=\mathbf{B}.$$

The augmented matrix

$$[\mathbf{A}\ \ \mathbf{B}]=\left[\begin{array}{ccc:c}1 & 1 & 1 & 6\\ 1 & 2 & 3 & 10\\ 1 & 2 & \lambda & \mu\end{array}\right]$$

$$\sim\left[\begin{array}{ccc:c}1 & 1 & 1 & 6\\ 0 & 1 & 2 & 4\\ 0 & 1 & \lambda-1 & \mu-6\end{array}\right]\text{ by } R_2\to R_2-R_1,\ R_3\to R_3-R_1$$

$$\sim\left[\begin{array}{ccc:c}1 & 1 & 1 & 6\\ 0 & 1 & 2 & 4\\ 0 & 0 & \lambda-3 & \mu-10\end{array}\right]\text{ by } R_3\to R_3-R_2.$$

If $\lambda \neq 3$, we have rank [A B]=3=rank A. So in this case the given system of equations is consistent. Since rank A=the number of unknowns, therefore the given system of equations possesses a unique solution. Thus if $\lambda \neq 3$, the given system of equations possesses a unique solution for any value of μ.

If $\lambda=3$ and $\mu \neq 10$, we have rank [A B]=3 and rank A=2. Thus in this case rank [A B]$\neq$rank A and so the given system of equations is inconsistent *i.e.*, possesses no solution.

If $\lambda=3$ and $\mu=10$, we have rank [A B]=rank A. So in this case the given system of equations is again consistent. Since rank A < the number of unknowns, therefore in this case the given system of equations possesses an infinite number of solutions.

Ex. 9. *For what values of the parameter λ will the following equations fail to have unique solution*

$$3x-y+\lambda z=1,$$
$$2x+y+z=2,$$
$$x+2y-\lambda z=-1\ ?$$

Will the equations have any solutions for these values of λ?

(Meerut 1983)

Solution. The matrix form of the given system of equations is

$$\begin{bmatrix} 3 & -1 & \lambda \\ 2 & 1 & 1 \\ 1 & 2 & -\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

The given system of equations will have a unique solution if and only if the coefficient matrix is non-singular.

Performing $R_1 \leftrightarrow R_3$, we get

$$\begin{bmatrix} 1 & 2 & -\lambda \\ 2 & 1 & 1 \\ 3 & -1 & \lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Performing $R_2 \to R_2-2R_1$, $R_3 \to R_3-3R_1$, we get

$$\begin{bmatrix} 1 & 2 & -\lambda \\ 0 & -3 & 1+2\lambda \\ 0 & -7 & 4\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 4 \end{bmatrix}. \quad \ldots(i)$$

Therefore the coefficient matrix will be non-singular if and only if

$$-12\lambda+7+14\lambda \neq 0$$

i.e. if and only if $\lambda \neq -\frac{7}{2}$.

Thus the given system will have a unique solution if $\lambda \neq -\frac{7}{2}$. In case $\lambda=-\frac{7}{2}$, the equation (i) becomes

$$\begin{bmatrix} 1 & 2 & \frac{7}{2} \\ 0 & -3 & -6 \\ 0 & -7 & -14 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ 4 \end{bmatrix}.$$

Performing $R_3 \to R_3 - \frac{7}{3}R_2$, we get

$$\begin{bmatrix} 1 & 2 & \frac{7}{2} \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ -\frac{16}{3} \end{bmatrix},$$

showing that given equations are inconsistent in this case.

Thus if $\lambda=-\frac{7}{2}$, no solution exists.

Ex. 10. *For what values of η the equations*

$$x+y+z=1,$$
$$x+2y+4z=\eta,$$
$$x+4y+10z=\eta^2,$$

have a solution and solve them completely in each case.

(Kanpur 1981; Meerut 89)

Solution. The matrix form of the given system is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \eta \\ \eta^2 \end{bmatrix}.$$

Performing $R_2 \to R_2-R_1$, $R_3 \to R_3-R_1$, we get

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 3 & 9 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \eta-1 \\ \eta^2-1 \end{bmatrix}$$

Performing $R_3 \to R_3-3R_2$, we get

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \eta-1 \\ \eta^2-3\eta+2 \end{bmatrix}. \quad \ldots(i)$$

Now the given equations will be consistent if and only if

$$\eta^2-3\eta+2=0,$$
$$\text{iff} \quad (\eta-2)(\eta-1)=0,$$
$$\text{iff} \quad \eta=2 \text{ or } \eta=1.$$

Case I. If $\eta=2$, the equation (i) becomes

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

The above system of equations is equivalent to

$$y+3z=1,\ x+y+z=1.$$
$$\therefore \quad y=1-3z,\ x=2z.$$

Thus $x=2k$, $y=1-3k$, $z=k$ constitute the general solution where k is an arbitrary constant.

Case II. If $\eta=1$, the equation (i) becomes

$$\begin{bmatrix}1 & 1 & 1\\0 & 1 & 3\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}.$$

The above system of equations is equivalent to

$$y+3z=0,\ x+y+z=1.$$

$$\therefore\quad y=-3z,\ x=1+2z.$$

Thus $x=1+2c$, $y=-3c$, $z=c$ constitute the general solution, where c is an arbitrary constant.

Ex. 11. *Discuss for all values of* λ, *the system of equations*

$$x+y+4z=6,$$
$$x+2y-2z=6,$$
$$\lambda x+y+z=6,$$

as regards existence and nature of solutions.

Solution. The matrix form of the given system is

$$\begin{bmatrix}1 & 1 & 4\\1 & 2 & -2\\\lambda & 1 & 1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}6\\6\\6\end{bmatrix}.$$

The given system of equations will have a unique solution if and only if the coefficient matrix is non-singular.

Performing $R_2\to R_2-R_1$, $R_3\to R_3-\lambda R_1$, we get

$$\begin{bmatrix}1 & 1 & 4\\0 & 1 & -6\\0 & 1-\lambda & 1-4\lambda\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}6\\0\\6-6\lambda\end{bmatrix}. \quad \text{...(i)}$$

Therefore the coefficient matrix will be non-singular if and only if

$$1-4\lambda+6-6\lambda\neq 0$$

i.e., $\quad \lambda\neq 7/10.$

Thus the given system will have a unique solution if

$$\lambda\neq 7/10.$$

In case $\lambda=7/10$, the equation (i) becomes

$$\begin{bmatrix}1 & 1 & 4\\0 & 1 & -6\\0 & \frac{3}{10} & -\frac{18}{10}\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}6\\0\\\frac{18}{10}\end{bmatrix}.$$

Performing $R_3\to R_3-\frac{3}{10}R_2$, we get

$$\begin{bmatrix}1 & 1 & 4\\0 & 1 & -6\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}6\\0\\\frac{18}{10}\end{bmatrix},$$

showing that the equations are not consistent in this case.

Ex. 12. *Solve the equations*

$$\lambda x+2y-2z-1=0,$$
$$4x+2\lambda y-z-2=0,$$
$$6x+6y+\lambda z-3=0,$$

considering specially the case when $\lambda=2$. **(Kanpur 1971)**

Solution. The matrix form of the given system is

$$\begin{bmatrix}\lambda & 2 & -2\\4 & 2\lambda & -1\\6 & 6 & \lambda\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}. \quad \text{...(i)}$$

The given system of equations will have a unique solution if and only if the coefficient matrix is non-singular, *i.e.*, iff

$$\begin{vmatrix}\lambda & 2 & -2\\4 & 2\lambda & -1\\6 & 6 & \lambda\end{vmatrix}\neq 0$$

i.e., iff $\quad \lambda^3+11\lambda-30\neq 0$

i.e., iff $\quad (\lambda-2)(\lambda^2+2\lambda+15)\neq 0.$

Now the only real root of the equation

$$(\lambda-2)(\lambda^2+2\lambda+15)=0 \text{ is } \lambda=2.$$

Therefore if $\lambda\neq 2$, the given system of equations will have a unique solution given by

$$\frac{x}{\begin{vmatrix}1 & 2 & -2\\2 & 2\lambda & -1\\3 & 6 & \lambda\end{vmatrix}}=\frac{y}{\begin{vmatrix}\lambda & 1 & -2\\4 & 2 & -1\\6 & 3 & \lambda\end{vmatrix}}=\frac{z}{\begin{vmatrix}\lambda & 2 & 1\\4 & 2\lambda & 2\\6 & 6 & 3\end{vmatrix}}=\frac{1}{\begin{vmatrix}\lambda & 2 & -2\\4 & 2\lambda & -1\\6 & 6 & \lambda\end{vmatrix}}$$

In case $\lambda=2$, the equation (i) becomes

$$\begin{bmatrix}2 & 2 & -2\\4 & 4 & -1\\6 & 6 & 2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}.$$

Performing $R_2\to R_2-2R_1$, $R_3\to R_3-3R_1$, we get

$$\begin{bmatrix}2 & 2 & -2\\0 & 0 & 3\\0 & 0 & 8\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\0\\0\end{bmatrix}.$$

The above system of equations is equivalent to

$$8z=0,\ 3z=0,\ 2x+2y-2z=1.$$

$\therefore\ x=\frac{1}{2}-c,\ y=c,\ z=0$ constitute the general solution of the given system of equations in case $\lambda=2$.

Exercises

1. Show that the equations
$$x-4y+7z=14, 3x+8y-2z=13, 7x-8y+26z=5$$
are not consistent.

2. Use the test of rank to show that the following equations are not consistent :
$$2x-y+z=4,$$
$$3x-y+z=6,$$
$$4x-y+2z=7,$$
$$-x+y-z=9.$$

3. Apply the test of rank to examine if the following system of equations is consistent, and if consistent, find the complete solution :
$$x+y+z=6, x+2y+3z=10, x+2y+4z=1.$$
(Meerut 1979)

4. Solve completely the equations
$$3x-2y-w=2$$
$$2y+2z+w=1$$
$$x-2y-3z+2w=3$$
$$y+2z+w=1.$$
(Gujrat 1971)

5. Show that the equations
$$3x+y+z=8,$$
$$-x+y-2z=-5,$$
$$2x+2y+2z=12,$$
$$-2x+2y-3z=-7$$
are consistent and solve the same.

6. Show that the equations
$$x-3y-8z+10=0, 3x+y-4z=0, 2x+5y+6z-13=0$$
are consistent and solve the same.

7. Solve completely the equations
$$2x+3y+z=9, x+2y+3z=6, 3x+y+2z=8.$$
(Ravi Shanker 1971; I.A.S. 85)

8. Solve completely the equations
$$x+y+z=1$$
$$x+2y+3z=4$$
$$x+3y+5z=7$$
$$x+4y+7z=10.$$

9. Show that the equations
$$x+2y-5z=-9$$
$$3x-y+2z=5$$

$$2x+3y-z=3$$
$$4x-5y+z=-3$$
are consistent and solve the same.

10. Express the following system of equations into the matrix equation form **AX=B** :
$$4x-y+6z=16$$
$$x-4y-3z=-16$$
$$2x+7y+12z=48$$
$$5x-5y+3z=0,$$
Determine if this system of equations is consistent and if so find its solution. **(Meerut 1975)**

11. Solve the following equations using matrix methods :
$$2x-y+3z-9=0, x+y+z-6=0, x-y+z-2=0.$$
(Meerut 1991)

12. Examine if the system of equations :
$$x+y+4z=6, 3x+2y-2z=9, 5x+y+2z=13$$
is consistent ? Find also the solution if it is consistent ? **(Meerut 1983)**

13. Show that the equations :
$$3x+7y+5z=4, 26x+2y+3z=9, 2x+10y+7z=5$$
are consistent and solve them. **(Gujrat 1971)**

14. Examine for consistency and solve (if consistent) the system of equations
$$x-y+2z=4, 3x+y+4z=6, x+y+z=1.$$
(Meerut 1973; Delhi 79)

15. Show that the three equations
$$-2x+y+z=a, x-2y+z=b, x+y-2z=c$$
have no solutions unless $a+b+c=0$, in which case they have infinitely many solutions. Find these solutions when
$a=1, b=1, c=-2.$ **(Poona 1970)**

16. Show that there are two values of k for which the equations
$$kx+3y+2z=1,$$
$$x+(k-1)y=4,$$
$$10y+3z=-2,$$
$$2x-ky-z=5,$$
are consistent. Find their common solution for that value of k which is an integer.

17. Investigate for what values of a, b the equations
$$x+2y+3z=4, x+3y+4z=5, x+3y+az=b,$$
have (i) no solution, (ii) a unique solution and (iii) an infinite number of solutions **(I. A. S. 1971)**

Answers

3. $x=-7, y=22, z=-9$. 4. $x=1, y=0, z=0, w=1$.
5. $x=1, y=2, z=3$.
6. $x=-1+2c, y=3-2c, z=c$, where c is arbitrary.
7. $x=\frac{35}{18}, y=\frac{29}{18}, z=\frac{5}{18}$.
8. $x=c-2, y=3-2c, z=c$.
9. $x=\frac{1}{2}, y=\frac{3}{2}, z=\frac{5}{2}$.
10. Consistent ; $x=-\frac{9}{5}c+\frac{16}{3}, y=-\frac{6}{5}c+\frac{16}{3}, z=c$.
11. $x=1, y=2, z=3$.
12. Consistent ; $x=2, y=2, z=\frac{1}{2}$.
13. $x=\frac{5}{16}-\frac{1}{16}c, y=\frac{7}{16}-\frac{11}{16}c, z=c$.
14. Consistent ; $x=\frac{5}{2}-\frac{3}{2}c, y=-\frac{3}{2}+\frac{1}{2}c, z=c$.
15. $x=c-1, y=c-1, z=c$, (c arbitrary).
16. $k=3, -\frac{1}{25}$; $x=2, y=1, z=-4$.
17. (i) no solution when $a=4, b\neq5$; (ii) a unique solution for all values of b provided $a\neq4$; (iii) an infinite number of solutions when $a=4, b=5$.

7

Eigenvalues and Eigenvectors

§ 1. **Matric Polynomials. Def.** *An expression of the form*

$$F(\lambda)=A_0+A_1\lambda+A_2\lambda^2+\ldots+A_{m-1}\lambda^{m-1}+A_m\lambda^m,$$

where $A_0, A_1, A_2, \ldots, A_m$ *are all square matrices of the same order is called a* **Matric Polynomial** *of degree m provided* A_m *is not a null matrix.* The symbol λ is called *indeterminate.* **If the order of each of the matrix coefficients $A_0, A_1, \ldots, A_m$ is n, then we say that the matric polynomial is n-rowed. According to this definition of a matric polynomial, each square matrix can be expressed as a matrix polynomial with zero degree. For example, if A be any square matrix, we can write $A=\lambda^0 A$.**

Equality of polynomials. Two matric polynomials are equal iff (if and only if), the coefficients of the like powers of λ are the same.

Theorem. *Every square matrix whose elements are ordinary polynomials in λ, can essentially be expressed as a matrix polynomial in λ of degree m, where m is the highest power of λ occurring in any element of the matrix.* **We shall illustrate this theorem by the following example :**

Consider the matrix

$$A=\begin{bmatrix}1+2\lambda+3\lambda^2 & \lambda^2 & 4-6\lambda\\ 1+\lambda^3 & 3+4\lambda^2 & 1-2\lambda+4\lambda^3\\ 2-3\lambda+2\lambda^3 & 5 & 6\end{bmatrix}$$

in which the highest power of λ occurring in any element is 3. Rewriting each element as a cubic in λ, supplying missing coefficients with zero, we get

$$A=\begin{bmatrix}1+2.\lambda+3.\lambda^2+0.\lambda^3 & 0+0.\lambda+1.\lambda^2+0.\lambda^3 & 4-6.\lambda+0.\lambda^2+0.\lambda^3\\ 1+0.\lambda+0.\lambda^2+1.\lambda^3 & 3+0.\lambda+4.\lambda^2+0.\lambda^3 & 1-2.\lambda+0.\lambda^2+4.\lambda^3\\ 2-3.\lambda+0.\lambda^2+2.\lambda^3 & 5+0.\lambda+0.\lambda^2+0.\lambda^3 & 6+0.\lambda+0.\lambda^2+0.\lambda^3\end{bmatrix}.$$

Obviously A can be written as the matric polynomial

$$A=\begin{bmatrix}1 & 0 & 4\\ 1 & 3 & 1\\ 2 & 5 & 6\end{bmatrix}+\lambda\begin{bmatrix}2 & 0 & -6\\ 0 & 0 & -2\\ -3 & 0 & 0\end{bmatrix}$$
$$+\lambda^2\begin{bmatrix}3 & 1 & 0\\ 0 & 4 & 0\\ 0 & 0 & 0\end{bmatrix}+\lambda^3\begin{bmatrix}0 & 0 & 0\\ 1 & 0 & 4\\ 2 & 0 & 0\end{bmatrix}.$$

§ 2. Characteristic values and Characteristic vectors of a matrix.

Let $\mathbf{X}=[a_{ij}]_{n\times n}$ be a given n-rowed square matrix. Let

$$\mathbf{X}=\begin{bmatrix}x_1\\ x_2\\ \dots\\ \dots\\ x_n\end{bmatrix}$$

be a column vector. Consider the vector equation

$$\mathbf{AX}=\lambda\mathbf{X} \qquad \dots(1)$$

where λ is a scalar (*i.e.*, number).

It is obvious that the zero vector $\mathbf{X}=\mathbf{O}$ is a solution of (1) for any value of λ. Now let us see whether there exist scalars λ and non-zero vectors $\mathbf{X}$ which satisfy (1).

If $\mathbf{I}$ denotes the unit matrix of order n, then the equation (1) may be written as

$$\mathbf{AX}=\lambda\mathbf{IX}$$

or $$(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{X}=\mathbf{O}. \qquad \dots(2)$$

The matrix equation (2) represents the following system of n homogeneous equations in n unknowns :

$$\left.\begin{array}{l}(a_{11}-\lambda)\,x_1+a_{12}\,x_2+\dots+a_{1n}x_n=0\\ a_{21}\,x_1+(a_{22}-\lambda)\,x_2+\dots+a_{2n}x_n=0\\ \dots\quad\dots\quad\dots\quad\dots\quad\dots\quad\dots\\ a_{n1}x_1+a_{n2}x_2+\dots+(a_{nn}-\lambda)\,x_n=0\end{array}\right\} \qquad \dots(3)$$

The coefficient matrix of the equations (3) is $\mathbf{A}-\lambda\mathbf{I}$. The necessary and sufficient condition for equations (3) to possess a non-zero solution ($\mathbf{X}\neq\mathbf{O}$) is that the coefficient matrix $\mathbf{A}-\lambda\mathbf{I}$ should be of rank less than the number of unknowns n. But this will be so if and only if the matrix $\mathbf{A}-\lambda\mathbf{I}$ is singular *i.e.*, if and only if $|\mathbf{A}-\lambda\mathbf{I}|=0$. Thus the scalars λ for which

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

are of special importance.

Definitions. **(Kerala 1970; Meerut 86; Delhi 79; Gujrat 70; Kanpur 71)**

Let $\mathbf{A}=[a_{ij}]_{n\times n}$ be any n-rowed square matrix and λ an indeterminate. The matrix $\mathbf{A}-\lambda\mathbf{I}$ is called the characteristic matrix of A where I is the unit matrix of order n.

Also the determinant

$$|\mathbf{A}-\lambda\mathbf{I}|=\begin{vmatrix}a_{11}-\lambda & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22}-\lambda & \dots & a_{2n}\\ \dots & \dots & \dots & \dots\\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda\end{vmatrix}$$

which is an ordinary polynomial in λ of degree n, **is called the characteristic polynomial of A.** The equation

$$|\mathbf{A}-\lambda\mathbf{I}|=0,$$

is called the characteristic equation of A and the roots of this equation are called the characteristic roots or characteristic values or eigen values or latent roots or proper values of the matrix A. The set of the eigen values of A is called the spectrum of A.

If λ is a characteristic root of the matrix $\mathbf{A}$, then

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

and the matrix $\mathbf{A}-\lambda\mathbf{I}$ is singular. Therefore there exists a non-zero vector $\mathbf{X}$ such that

$$(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{X}=\mathbf{O} \text{ or } \mathbf{AX}=\lambda\mathbf{X}.$$

Characteristic vectors. Definition. *If λ is a characteristic root of an $n\times n$ matrix* $\mathbf{A}$, *then a non-zero vector* $\mathbf{X}$ *such that*

$$\mathbf{AX}=\lambda\mathbf{X}$$

is called a characteristic vector or eigenvector of $\mathbf{A}$ *corresponding to the characteristic root* λ. **(Kerala 1970)**

§ 3. Certain relations between characteristic roots and characteristic vectors.

Theorem 1. *λ is a characteristic root of a matrix* $\mathbf{A}$ *if and only if there exists a non-zero vector* $\mathbf{X}$ *such that* $\mathbf{AX}=\lambda\mathbf{X}$. **(Agra 1977)**

Proof. Suppose λ is a characteristic root of the matrix $\mathbf{A}$. Then $|\mathbf{A}-\lambda\mathbf{I}|=0$ and the matrix $\mathbf{A}-\lambda\mathbf{I}$ is singular. Therefore the matrix equation $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{X}=\mathbf{O}$ possesses a non-zero solution *i.e.* there exists a non-zero vector $\mathbf{X}$ such that

$$(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{X}=\mathbf{O} \text{ or } \mathbf{AX}=\lambda\mathbf{X}.$$

Conversely suppose there exists a non-zero vector $\mathbf{X}$ such that $\mathbf{AX}=\lambda\mathbf{X}$ *i.e.*, $(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{X}=\mathbf{O}$. Since the matrix equation

$$(\mathbf{A}-\lambda\mathbf{I})\,\mathbf{X}=\mathbf{O}$$

possesses a non-zero solution, therefore the coefficient matrix $\mathbf{A}-\lambda\mathbf{I}$

must be singular *i.e.*, $|A-\lambda I|=0$. Hence λ is a characteristic root of the matrix **A**.

Theorem 2. *If* **X** *is a characteristic vector of a matrix* **A** *corresponding to the characteristic value* λ, *then* kX *is also a characteristic vector of* **A** *corresponding to the same characteristic value* λ. *Here* k *is any non-zero scalar.*

Proof. Suppose **X** is a characteristic vector of **A** corresponding to the characteristic value λ. Then $X \neq O$ and

$$AX=\lambda X.$$

If k is any non-zero scalar, then $kX \neq O$. Also

$$A(kX)=k(AX)=k(\lambda X)=\lambda(kX).$$

Now kX is a non-zero vector such that $A(kX)=\lambda(kX)$. Hence kX is a characteristic vector of **A** corresponding to the characteristic value λ. Thus corresponding to a characteristic value λ, there corresponds more than one characteristic vectors.

Theorem 3. *If* **X** *is a characteristic vector of a matrix* **A**, *then* **X** *cannot correspond to more than one characteristic values of* **A**.

Proof. Let **X** be a characteristic vector of a matrix **A** corresponding to two characteristic values λ_1 and λ_2. Then $AX=\lambda_1 X$ and $AX=\lambda_2 X$. Therefore

$$\lambda_1 X=\lambda_2 X$$
$$\Rightarrow (\lambda_1-\lambda_2)X=O$$
$$\Rightarrow \lambda_1-\lambda_2=0 \qquad [\because X\neq O]$$
$$\Rightarrow \lambda_1=\lambda_2.$$

Theorem 4. Linear independence of characteristic vectors corresponding to distinct characteristic roots.

The characteristic vectors corresponding to distinct characteristic roots of a matrix are linearly independent.

(Agra 1974; Nagarjuna 78)

Proof. Let $X_1,\ldots,X_m$ be the characteristic vectors of a matrix **A** corresponding to distinct characteristic values $\lambda_1,\ldots,\lambda_m$. Then $AX_i=\lambda_i X_i$, $i=1,\ldots,m$. To prove that the vectors $X_1,\ldots,X_m$ are linearly independent.

If the vectors $X_1,\ldots,X_m$ are linearly dependent we can choose r so that $1\leq r<m$ and $X_1,\ldots,X_r$ are linearly independent but $X_1,\ldots,X_r,X_{r+1}$ are linearly dependent. Hence we can choose scalars $a_1,\ldots,a_{r+1}$, not all zero such that

$$a_1X_1+\ldots+a_{r+1}X_{r+1}=O \qquad \ldots(1)$$

$$\Rightarrow A(a_1X_1+\ldots+a_{r+1}X_{r+1})=AO$$
$$\Rightarrow a_1AX_1+\ldots+a_{r+1}AX_{r+1}=O$$
$$\Rightarrow a_1(\lambda_1X_1)+\ldots+a_{r+1}(\lambda_{r+1}X_{r+1})=O. \qquad \ldots(2)$$

Multiplying (1) by the scalar λ_{r+1} and subtracting from (2), we get

$$a_1(\lambda_1-\lambda_{r+1})X_1+\ldots+a_r(\lambda_r-\lambda_{r+1})X_r=O. \qquad \ldots(3)$$

Since $X_1,\ldots,X_r$ are linearly independent according to our assumption and $\lambda_1,\ldots,\lambda_{r+1}$ are distinct, therefore from (3), we get

$$a_1=0,\ldots,a_r=0.$$

Putting $a_1=0,\ldots,a_r=0$ in (1), we get

$$a_{r+1}X_{r+1}=O$$
$$\Rightarrow a_{r+1}=0 \text{ since } X_{r+1}\neq O.$$

Thus the relation (1) implies that

$$a_1=0,\ldots,a_r=0,\ a_{r+1}=0.$$

But this contradicts our assumption that the scalars

$$a_1,\ldots,a_{r+1}$$

are not all zero.

Hence our initial assumption is wrong and the vectors $X_1,\ldots,X_m$ are linearly independent.

§ 4. Algebraic and geometric multiplicity of a characteristic root. If λ_1 be a characteristic root of order t of the characteristic equation $|A-\lambda I|=0$, then t is called the *algebraic multiplicity of* λ_1. If s be the number of linearly independent eigenvectors corresponding to the eigenvalue λ_1, then s is called the *geometric multiplicity of* λ_1. In this case number of linearly independent solutions of $(A-\lambda_1 I)X=O$ will be s and the matrix $A-\lambda_1 I$ will be of rank $n-s$.

The geometric multiplicity of a characteristic root cannot exceed its algebraic multiplicity i.e. $s\leq t$.

§ 5. Nature of the characteristic roots of special types of matrices.

Theorem 1. *The characteristic roots of a Hermitian matrix are real.* **(I.C.S. 1988; Kanpur 85; Gujrat 70; Allahabad 82)**

Proof. Suppose **A** is a Hermitian matrix, λ a characteristic root of **A** and **X** a corresponding eigenvector. Then

$$AX=\lambda X. \qquad \ldots(i)$$

Premultiplying both sides of (i) by X^θ, we get

$$X^\theta AX=\lambda X^\theta X. \qquad \ldots(ii)$$

Taking conjugate transpose of both sides of (ii), we get

$$(X^{\theta}AX)^{\theta}=(\lambda X^{\theta}X)^{\theta}$$

or $$X^{\theta}A^{\theta}(X^{\theta})^{\theta}=\bar{\lambda}X^{\theta}(X^{\theta})^{\theta}$$

or $$X^{\theta}AX=\bar{\lambda}X^{\theta}X \qquad \ldots\text{(iii)}$$

$$[\because (X^{\theta})^{\theta}=X \text{ and } A^{\theta}=A, A \text{ being Hermitian}]$$

From (ii) and (iii), we have

$$\lambda X^{\theta}X=\bar{\lambda}X^{\theta}X$$

or $$(\lambda-\bar{\lambda})\,X^{\theta}X=O.$$

But X is not a zero vector, therefore $X^{\theta}X\neq O$.

Hence $\lambda-\bar{\lambda}=0$, so that $\lambda=\bar{\lambda}$ and consequently λ is real.

Corollary 1. *The characterstic roots of a real symmetric matrix are all real.* **(Nagarjuna 1978; Allahabad 71; Gujrat 70, Nagpur 70)**

If the elements of a Hermitian matrix A are all real, then A is a real symmetric matrix. Thus a real symmetric matrix is Hermitian and therefore the result follows.

Corollary 2. *The characteristic roots of a skew-Hermitian matrix are either pure imaginary or zero.* **(Kanpur 1989; I.C.S. 88)**

Suppose A is a skew-Hermitian matrix. Then iA is Hermitian. Let λ be a characteristic root of A. Then

$$AX=\lambda X$$

or $$(iA)X=(i\lambda)\,X.$$

From this it follows that $i\lambda$ is a characteristic root of iA which is Hermitian. Hence $i\lambda$ is real. Therefore either λ must be zero or pure imaginary.

Corollary 3. *The characteristic roots of a real symmetric matrix are either pure imaginary or zero, for every such matrix is skew-Hermitian.*

Theorem 2. *The characteristic roots of a unitary matrix are of unit modulus.* **(Madurai 1985)**

Proof. **Suppose A is a unitary matrix. Then**

$$A^{\theta}A=I.$$

Let λ be a characteristic root of A. Then

$$AX=\lambda X. \qquad \ldots\text{(i)}$$

Taking conjugate transpose of both sides of (i), we get

$$(AX)^{\theta}=(\lambda X)^{\theta}$$

or $$X^{\theta}A^{\theta}=\bar{\lambda}X^{\theta} \qquad \ldots\text{(ii)}$$

From (i) and (ii), we have

$$(X^{\theta}A^{\theta})(AX)=\bar{\lambda}\lambda X^{\theta}X$$

or $$X^{\theta}(A^{\theta}A)\,X=\bar{\lambda}\lambda X^{\theta}X$$

or $$X^{\theta}IX=\bar{\lambda}\lambda X^{\theta}X$$

or $$X^{\theta}X=\bar{\lambda}\lambda X^{\theta}X$$

or $$X^{\theta}X\,(\lambda\bar{\lambda}-1)=O. \qquad \ldots\text{(iii)}$$

Since $X^{\theta}X\neq O$, therefore (iii) gives

$$\lambda\bar{\lambda}-1=0 \text{ or } \lambda\bar{\lambda}=1 \text{ or } |\lambda|^2=1.$$

Corollary. *The characteristic roots of an orthogonal matrix are of unit modulus.* **(Madurai 1985)**

We know that if the elements of a unitary matrix A are all real, then A is said to be an orthogonal matrix. Hence the result follows.

§ 6. The process of finding the eigenvalues and eigenvectors of a matrix.

Let $A=[a_{ij}]_{n\times n}$ be a square matrix of order n. First we should write the characteristic equation of the matrix A *i.e.* the equation $|A-\lambda I|=0$. This equation will be of degree n in λ. So it will have n roots. These n roots will give us the eigenvalues of the matrix A. If λ_1 is an eigenvalue of A, then the corresponding eigenvectors of A will be given by the non-zero vectors

$$X=[x_1, x_2, \ldots, x_n]'$$

satisfying the equation

$$AX=\lambda_1 X \text{ or } (A-\lambda_1 I)\,X=O.$$

Solved Examples

Ex. 1. *Determine the characteristic roots of the matrix*

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

Solution. **The characteristic matrix of A**

$$=A-\lambda I=\begin{bmatrix}0 & 1 & 2\\ 1 & 0 & -1\\ 2 & -1 & 0\end{bmatrix}-\lambda\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

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

It should be noted that in order to obtain the characteristic matrix of a matrix A, we should simply subtract λ from each of its principal diagonal elements.

The characteristic polynomial of A

$$=|A-\lambda I|=\begin{vmatrix}-\lambda & 1 & 2\\ 1 & -\lambda & -1\\ 2 & -1 & -\lambda\end{vmatrix}$$

$$=-\lambda(\lambda^2-1)-1(-\lambda+2)+2(-1+2\lambda)$$

$$=-\lambda^3+\lambda+\lambda-2-2+4\lambda=-\lambda^3+6\lambda-4.$$

$\therefore$ the characteristic equation of **A** is $|\mathbf{A}-\lambda\mathbf{I}|=0$

i.e., $\lambda^3-6\lambda+4=0$ *i.e.,* $(\lambda-2)(\lambda^2+2\lambda-2)=0.$

The roots of this equation are $\lambda=2,\ -1\pm\sqrt{3}$.

Hence the characteristic roots of the matrix **A** are $2,\ -1\pm\sqrt{3}$.

Ex. 2. *Determine the eigenvalues of the matrix*

$$\mathbf{A}=\begin{bmatrix} a & h & g \\ 0 & b & 0 \\ 0 & c & c \end{bmatrix}.$$

(Kanpur 1979)

Solution. Here $|\mathbf{A}-\lambda\mathbf{I}|=\begin{vmatrix} a-\lambda & h & g \\ 0 & b-\lambda & 0 \\ 0 & c & c-\lambda \end{vmatrix}$

$$=(a-\lambda)(b-\lambda)(c-\lambda).$$

The characteristic equation of **A** is

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

i.e., $(a-\lambda)(b-\lambda)(c-\lambda)=0.$

The roots of this equation are $\lambda=a, b, c$. Hence the eigenvalues of **A** are a, b, c.

Ex. 3. *Verify that the matrices*

$$\mathbf{A}=\begin{bmatrix} 0 & h & g \\ h & 0 & f \\ g & f & 0 \end{bmatrix},\ \mathbf{B}=\begin{bmatrix} 0 & f & h \\ f & 0 & g \\ h & g & 0 \end{bmatrix},\ \mathbf{C}=\begin{bmatrix} 0 & g & f \\ g & 0 & h \\ f & h & 0 \end{bmatrix}$$

have the same characteristic equation

$$\lambda^3-(f^2+g^2+h^2)\lambda-2fgh=0.$$

Solution. The characteristic equation of the matrix **A** is

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

i.e., $\begin{vmatrix} 0-\lambda & h & g \\ h & 0-\lambda & f \\ g & f & 0-\lambda \end{vmatrix}=0$

i.e. $-\lambda(\lambda^2-f^2)-h(-h\lambda-gf)+g(fh+g\lambda)=0$

i.e. $\lambda^3-\lambda(f^2+g^2+h^2)-2fgh=0.$

Similarly show that the matrices **B** and **C** have the same characteristic equation.

Ex. 4. *Determine the eigenvalues and eigenvectors of the matrix*

$$\mathbf{A}=\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}.$$

Solution. The characteristic equation of **A** is

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

i.e., $\begin{vmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{vmatrix}=0$

i.e., $(5-\lambda)(2-\lambda)-4=0$ *i.e.,* $\lambda^2-7\lambda+6=0.$

The roots of this equation are $\lambda_1=6, \lambda_2=1$. Therefore the eigenvalues of **A** are 6, 1.

The eigenvectors $\mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ of **A** corresponding to the eigenvalue 6 are given by the non-zero solutions of the equation

$$(\mathbf{A}-6\mathbf{I})\,\mathbf{X}=\mathbf{O}$$

or $\begin{bmatrix} 5-6 & 4 \\ 1 & 2-6 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or $\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or $\begin{bmatrix} -1 & 4 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, applying $R_2\to R_2+R_1$.

The coefficient matrix of these equations is of rank 1. Therefore these equations have $2-1$ *i.e.,* 1 linearly independent solution. These equations reduce to the single equation $-x_1+4x_2=0$. Obviously $x_1=4,\ x_2=1$ is a solution of this equation. Therefore $\mathbf{X}_1=\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 6. The set of all eigenvectors of **A** corresponding to the eigenvalue 6 is given by $c_1\mathbf{X}_1$ where c_1 is any non-zero scalar.

The eigenvectors **X** of **A** corresponding to the eigenvalue 1 are given by the non-zero solutions of the equation

$$(\mathbf{A}-1\mathbf{I})\,\mathbf{X}=\mathbf{O}$$

or $\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

or $4x_1+4x_2=0,\ x_1+x_2=0.$

From these $x_1=-x_2$. Let us take $x_1=1,\ x_2=-1$. Then $\mathbf{X}_2=\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 1. Every non-zero multiple of the vector $\mathbf{X}_2$ is an eigenvector of **A** corresponding to the eigenvalue 1.

Ex. 5. *Determine the characteristic roots and the corresponding characteristic vectors of the matrix*

$$\mathbf{A}=\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}.$$

(Agra 1977; Kanpur 85; Delhi 80; Rohilkhand 81; I.A.S. 83)

Solution. The characteristic equation of the matrix A is

$$|A-\lambda I|=0$$

$$\text{i.e.,}\quad \begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix}=0$$

$$\text{or}\quad (8-\lambda)\{(7-\lambda)(3-\lambda)-16\}+6\{-6(3-\lambda)+8\}+2\{24-2(7-\lambda)\}=0$$

$$\text{or}\quad \lambda^3-18\lambda^2+45\lambda=0 \text{ or } \lambda(\lambda-3)(\lambda-15)=0.$$

Hence the characteristic roots of A are 0, 3, 15.

The eigenvectors $X=[x_1, x_2, x_3]'$ of A corresponding to the eigenvalue 0 are given by the non-zero solutions of the equation

$$(A-0I)\,X=O$$

$$\text{or}\quad \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} 2 & -4 & 3 \\ -6 & 7 & -4 \\ 8 & -6 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_1\leftrightarrow R_3$$

$$\text{or}\quad \begin{bmatrix} 2 & -4 & 3 \\ 0 & -5 & 5 \\ 0 & 10 & -10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_2\to R_2+3R_1,\ R_3\to R_3-4R_1$$

$$\text{or}\quad \begin{bmatrix} 2 & -4 & 3 \\ 0 & -5 & 5 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_3\to R_3+2R_2.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations have $3-2=1$ linearly independent solution. Thus there is only one linearly independent eigenvector corresponding to the eigenvalue 0. These equations can be written as

$$2x_1-4x_2+3x_3=0,$$
$$-5x_2+5x_3=0.$$

From the last equation, we get $x_2=x_3$. Let us take $x_2=1$, $x_3=1$. Then the first equation gives $x_1=1/2$. Therefore $X_1=[\frac{1}{2}\ \ 1\ \ 1]'$ is an eigenvector of A corresponding to the eigenvalue 0. If c_1 is any non-zero scalar, then c_1X_1 is also an eigenvector of A corresponding to the eigenvalue 0.

The eigenvectors of A corresponding to the eigenvalue 3 are given by the non-zero solutions of the equation

$$(A-3I)\,X=O$$

$$\text{or}\quad \begin{bmatrix} 5 & -6 & 2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} -1 & -2 & -2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_1\to R_1+R_2$$

$$\text{or}\quad \begin{bmatrix} -1 & -2 & -2 \\ 0 & 16 & 8 \\ 0 & -8 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_2\to R_2-6R_1,\ R_3\to R_3+2R_1$$

$$\text{or}\quad \begin{bmatrix} -1 & -2 & -2 \\ 0 & 16 & 8 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_3\to R_3+\tfrac{1}{2}R_2.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations have $3-2=1$ linearly independent solution. These equations can be written as

$$-x_1-2x_2-2x_3=0,$$
$$16x_2+8x_3=0.$$

From the second equation we get $x_2=-\frac{1}{2}x_3$. Let us take $x_3=4$, $x_2=-2$. Then the first equation gives $x_1=-4$. Therefore

$$X_2=[-4\ \ -2\ \ 4]'$$

is an eigenvector of A corresponding to the eigenvalue 3. Every non-zero multiple of X_2 is an eigenvector of A corresponding to the eigenvalue 3.

The eigenvectors of A corresponding to the eigenvalue 15 are given by the non-zero solutions of the equation

$$(A-15I)\,X=O$$

$$\text{or}\quad \begin{bmatrix} 8-15 & -6 & 2 \\ -6 & 7-15 & -4 \\ 2 & -4 & 3-15 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} -7 & -6 & 2 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} -1 & 2 & 6 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ \text{by } R_1\to R_1-R_2$$

$$\text{or}\quad \begin{bmatrix} -1 & 2 & 6 \\ 0 & -20 & -40 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 - 6R_1,\ R_3 \to R_3 + 2R_1.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations have $3-2=1$ linearly independent solution. These equations can be written as

$$-x_1 + 2x_2 + 6x_3 = 0,$$
$$-20x_2 - 40x_3 = 0.$$

The last equation gives $x_2 = -2x_3$. Let us take $x_3 = 1$, $x_2 = -2$. Then the first equation gives $x_1 = 2$. Therefore

$$X_3 = [2 \ \ -2 \ \ 1]'$$

is an eigenvector of A corresponding to the eigenvalue 15. If k is any non-zero scalar, then kX_3 is also an eigenvector of A corresponding to the eigenvalue 15.

Ex. 6. *Determine the characteristic roots and the corresponding characteristic vectors of the matrix*

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

(Agra 1970)

Solution. The characteristic equation of A is

$$|A - \lambda I| = 0$$

$$\text{or}\quad \begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0$$

$$\text{or}\quad \begin{vmatrix} 6-\lambda & -2 & 0 \\ -2 & 3-\lambda & 2-\lambda \\ 2 & -1 & 2-\lambda \end{vmatrix} = 0, \text{ by } C_3 \to C_3 + C_2$$

$$\text{or}\quad (2-\lambda) \begin{vmatrix} 6-\lambda & -2 & 0 \\ -2 & 3-\lambda & 1 \\ 2 & -1 & 1 \end{vmatrix} = 0$$

$$\text{or}\quad (2-\lambda) \begin{vmatrix} 6-\lambda & -2 & 0 \\ -4 & 4-\lambda & 0 \\ 2 & -1 & 1 \end{vmatrix} = 0, \text{ by } R_2 \to R_2 - R_3$$

$$\text{or}\quad (2-\lambda)\,[(6-\lambda)(4-\lambda) - 8] = 0$$

$$\text{or}\quad (2-\lambda)(\lambda^2 - 10\lambda + 16) = 0$$

$$\text{or}\quad (2-\lambda)(\lambda-2)(\lambda-8) = 0.$$

Therefore the characteristic roots of A are given by

$$\lambda = 2, 2, 8.$$

The characteristic vectors of A corresponding to the characteristic root 8 are given by the non-zero solutions of the equation

$$(A - 8I)\,X = O$$

$$\text{or}\quad \begin{bmatrix} 6-8 & -2 & 2 \\ -2 & -8 & -1 \\ 2 & -1 & 3-8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} -2 & -2 & 2 \\ -2 & -5 & -1 \\ 2 & -1 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & -3 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_2 \to R_2 - R_1,\ R_3 \to R_3 + R_1$$

$$\text{or}\quad \begin{bmatrix} -2 & -2 & 2 \\ 0 & -3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_3 \to R_3 - R_2.$$

The coefficient matrix of these equations is of rank 2. Therefore these equations possess $3-2=1$ linearly independent solution. These equations can be written as

$$-2x_1 - 2x_2 + 2x_3 = 0,\ -3x_2 - 3x_3 = 0.$$

The last equation gives $x_2 = -x_3$. Let us take $x_3 = 1$, $x_2 = -1$. Then the first equation gives $x_1 = 2$. Therefore $X_1 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 8. Every non-zero multiple of X_1 is an eigenvector of A corresponding to the eigenvalue 8.

The eigenvectors of A corresponding to the eigenvalue 2 are given by the non-zero solutions of the equation

$$(A - 2I)\,X = O$$

$$\text{or}\quad \begin{bmatrix} 4 & -2 & 2 \\ -2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix} -2 & 1 & -1 \\ 4 & -2 & 2 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ by } R_1 \leftrightarrow R_2$$

or $\begin{bmatrix} -2 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, by $R_2 \to R_2+2R_1$, $R_3 \to R_3+R_1$.

The coefficient matrix of these equations is of rank 1. Therefore these equations possess $3-1=2$ linearly independent solutions. We see that these equations reduce to the single equation

$$-2x_1+x_2-x_3=0.$$

Obviously

$$\mathbf{X}_2=\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \qquad \mathbf{X}_3=\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

are two linearly independent solutions of this equation. Therefore $\mathbf{X}_2$ and $\mathbf{X}_3$ are two linearly independent eigenvectors of **A** corresponding to the eigenvalue 2. If c_1, c_2 are scalars not both equal to zero, then $c_1\mathbf{X}_2+c_2\mathbf{X}_3$ gives all the eigenvectors of **A** corresponding to the eigenvalue 2.

Ex. 7. *Determine the eigenvectors of the matrix*

$$\mathbf{A}=\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

Solution. The characteristic equation of **A** is

$$\begin{vmatrix} 2-\lambda & 1 & 0 \\ 0 & 2-\lambda & 1 \\ 0 & 0 & 2-\lambda \end{vmatrix}=0 \text{ i.e. } (2-\lambda)^3=0.$$

Thus, 2 is the only distinct eigenvalue of **A**. The corresponding eigenvectors are given by the non-zero solutions of the equation

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The coefficient matrix of these equations is of rank 2. Therefore there is only one linearly independent solution of these equations. These equations may be written as $x_2=0$, $x_3=0$.

Therefore $x_1=1$, $x_2=0$, $x_3=0$ is a non-zero solution of these equations. So $\mathbf{X}=[1 \ \ 0 \ \ 0]'$ is an eigenvector of **A** corresponding to the eigenvalue 2. Any non-zero multiple of this vector is an eigenvector of **A** corresponding to $\lambda=2$.

Ex. 8. *Show that the matrices* **A** *and* **A′** *have the same eigenvalues.*

Solution. We have $(\mathbf{A}-\lambda\mathbf{I})'=\mathbf{A}'-\lambda\mathbf{I}'=\mathbf{A}'-\lambda\mathbf{I}$.

$\therefore \quad |(\mathbf{A}-\lambda\mathbf{I})'|=|\mathbf{A}'-\lambda\mathbf{I}|$

or $\quad |\mathbf{A}-\lambda\mathbf{I}|=|\mathbf{A}'-\lambda\mathbf{I}|$. $\quad [\because \ |\mathbf{B}'|=|\mathbf{B}|]$

$\therefore \quad |\mathbf{A}-\lambda\mathbf{I}|=0$ if and only if $|\mathbf{A}'-\lambda\mathbf{I}|=0$

i.e., λ is an eigenvalue of **A** if and only if λ is an eigenvalue of **A′**.

Ex. 9. *Show that the characteristic roots of* $\mathbf{A}^\theta$ *are the conjugates of the characteristic roots of* **A**.

Solution. We have $|\mathbf{A}^\theta-\bar{\lambda}\mathbf{I}|=|(\mathbf{A}-\lambda\mathbf{I})^\theta|=\overline{|\mathbf{A}-\lambda\mathbf{I}|}$

[Note that $|\mathbf{B}^\theta|=|(\overline{\mathbf{B}'})|=\overline{|\mathbf{B}'|}=\overline{|\mathbf{B}|}$]

$\therefore \quad |\mathbf{A}^\theta-\bar{\lambda}\mathbf{I}|=0$ iff $\overline{|\mathbf{A}-\lambda\mathbf{I}|}=0$

or $\quad |\mathbf{A}^\theta-\bar{\lambda}\mathbf{I}|=0$ iff $|\mathbf{A}-\lambda\mathbf{I}|=0 \quad [\because$ if z is a complex number, then $z=0$ if and only if $\bar{z}=0]$

or $\quad \bar{\lambda}$ is an eigenvalue of $\mathbf{A}^\theta$ if and only if λ is an eigenvalue of **A**.

Ex. 10. *Show that 0 is a characteristic root of a matrix if and only if the matrix is singular.* **(Madurai 1985)**

Solution. We have

0 is an eigenvalue of **A** $\Rightarrow \lambda=0$ satisfies the equation $|\mathbf{A}-\lambda\mathbf{I}|=0$

$\Rightarrow |\mathbf{A}|=0 \Rightarrow$ **A** is singular.

Conversely, **A** is singular $\Rightarrow |\mathbf{A}|=0$

$\Rightarrow \lambda=0$ satisfies the equation $|\mathbf{A}-\lambda\mathbf{I}|=0$

$\Rightarrow$ 0 is an eigenvalue of **A**.

Ex. 11. *Show that the characteristic roots of a triangular matrix are just the diagonal elements of the matrix.* **(I.A.S. 1983)**

Solution. Let $\mathbf{A}=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$ be a triangular matrix of order n.

We have $|\mathbf{A}-\lambda\mathbf{I}|=\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22}-\lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn}-\lambda \end{vmatrix}$

$=(a_{11}-\lambda)(a_{22}-\lambda)\ldots\ldots(a_{nn}-\lambda).$

$\therefore$ the roots of the equation $|A-\lambda I|=0$ are

$\lambda=a_{11}, a_{22}, \ldots, a_{nn}.$

Hence the characteristic roots of A are $a_{11}, a_{22}, \ldots, a_{nn}$ which are just the diagonal elements of A.

Note. Similarly we can show that the characteristic roots of a diagonal matrix are just the diagonal elements of the matrix.

Ex. 12. *If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of* A, *then show that* $k\lambda_1, \ldots, k\lambda_n$ *are the eigenvalues of* kA. **(Kanpur 1987)**

Solution. If $k=0$, then $kA=O$ and each eigenvalue of O is 0. Thus $0\lambda_1, \ldots, 0\lambda_n$ are the eigenvalues of kA if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

So let us suppose that $k\neq 0$.

We have $|kA-\lambda kI|=|k(A-\lambda I)|$

$=k^n|A-\lambda I|$ $\quad [\because \ |kB|=k^n|B|]$

$\therefore$ if $k\neq 0$, then $|kA-\lambda kI|=0$ if and only if $|A-\lambda I|=0$ *i.e.*, $k\lambda$ is an eigenvalue of kA if and only if λ is an eigenvalue of A.

Thus $k\lambda_1, \ldots, k\lambda_n$ are the eigenvalues of kA if $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

Ex. 13. *If* A *is non-singular, prove that the eigenvalues of* A^{-1} *are the reciprocals of the eigenvalues of* A.

(Kerala 1970; Rohilkhand 78; Nagarjuna 78)

Solution. Let λ be an eigenvalue of A and X be a corresponding eigenvector. Then

$AX=\lambda X \Rightarrow X=A^{-1}(\lambda X)=\lambda(A^{-1}X)$

$\Rightarrow \frac{1}{\lambda}X=A^{-1}X \quad [\because$ A is non-singular $\Rightarrow \lambda\neq 0]$

$\Rightarrow A^{-1}X=\frac{1}{\lambda}X$

$\Rightarrow \frac{1}{\lambda}$ is an eigenvalue of A^{-1} and X is a corresponding eigenvector.

Conversely suppose that k is an eigenvalue of A^{-1}. Since A is non-singular $\Rightarrow A^{-1}$ is non-singular and $(A^{-1})^{-1}=A$, therefore it follows from the first part of this question that $\frac{1}{k}$ is an eigenvalue of A. Thus each eigenvalue of A^{-1} is equal to the reciprocal of some eigenvalue of A.

Hence the eigenvalues of A^{-1} are nothing but the reciprocals of the eigenvalues of A.

Ex. 14. *Show that if λ is a characteristic root of the matrix* A *then $k+\lambda$ is a characteristic root of the matrix* $A+kI$.

(Kerala 1970)

Solution. Let λ be a characteristic root of the matrix A and X be a corresponding characteristic vector. Then X is a non-zero vector such that $AX=\lambda X$. ...(1)

Now $\quad (A+kI)X=AX+kIX$

$=\lambda X+kX \quad$ [by (1)]

$=(\lambda+k)X.$ ...(2)

Since $X\neq O$, therefore from the relation (2), we see that the scalar $\lambda+k$ is a characteristic value of the matrix $A+kI$ and X is a corresponding characteristic vector.

Ex. 15. *If $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the characteristic roots of the n-square matrix* A *and k is a scalar, prove that characteristic roots of $A-kI$ are $\alpha_1-k, \alpha_2-k, \ldots, \alpha_n-k$.*

Solution. Since $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the characteristic roots of A, therefore the characteristic polynomial of A is

$|A-\lambda I|=(\alpha_1-\lambda)(\alpha_2-\lambda)\ldots(\alpha_n-\lambda).$ ...(i)

The characteristic polynomial of $A-kI$ is

$|A-kI-\lambda I|=|A-(k+\lambda)I|$

$=\{\alpha_1-(\lambda+k)\}\{\alpha_2-(\lambda+k)\}\ldots\{\alpha_n-(\lambda+k)\}$, from (i)

$=\{(\alpha_1-k)-\lambda\}\{(\alpha_2-k)-\lambda\}\ldots\{(\alpha_n-k)-\lambda\}$

which shows that the characteristic roots of $A-kI$ are

$\alpha_1-k, \alpha_2-k, \ldots, \alpha_n-k.$

Ex. 16. *Show that the two matrices* A, $C^{-1}AC$ *have the same characteristic roots.* **(Nagarjuna 1980; I.A.S. 84; Meerut 81)**

Solution. Let $\quad B=C^{-1}AC.$ Then

$B-\lambda I=C^{-1}AC-\lambda I$

$=C^{-1}AC-C^{-1}\lambda IC \quad [\because \ C^{-1}(\lambda I)C=\lambda C^{-1}C=\lambda I]$

$=C^{-1}(A-\lambda I)C.$

$\therefore \quad |B-\lambda I|=|C^{-1}|\,|A-\lambda I|\,|C|$

$=|A-\lambda I|\,|C^{-1}|\,|C|=|A-\lambda I|\,|C^{-1}C|$

$=|A-\lambda I|\,|I|=|A-\lambda I|.$

Thus the two matrices A and B have the same characteristic determinants and hence the same characteristic equations and the same characteristic roots.

Ex. 17. *Prove that if the characteristic roots of* A *are $\lambda_1, \lambda_2, \ldots, \lambda_n$, then the characteristic roots of A^2 are*

$\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2.$

Solution. Let λ be a characteristic root of the matrix **A**. Then there exists a non-zero vector **X** such that

$$\mathbf{AX}=\lambda\mathbf{X} \Rightarrow \mathbf{A}(\mathbf{AX})=\mathbf{A}(\lambda\mathbf{X})$$
$$\Rightarrow \mathbf{A}^2\mathbf{X}=\lambda(\mathbf{AX}) \Rightarrow \mathbf{A}^2\mathbf{X}=\lambda(\lambda\mathbf{X}) \; [\because \mathbf{AX}=\lambda\mathbf{X}]$$
$$\Rightarrow \mathbf{A}^2\mathbf{X}=\lambda^2\mathbf{X}. \qquad \ldots(1)$$

Since **X** is a non-zero vector, therefore from the relation (1) it is obvious that λ^2 is a characteristic root of the matrix **A**. Therefore if $\lambda_1, \ldots, \lambda_n$ are the characteristic roots of **A**, then $\lambda_1^2, \ldots, \lambda_n^2$ are the characteristic roots of $\mathbf{A}^2$.

Ex. 18. *If α is a characteristic root of a non-singular matrix* **A**, *then prove that $\dfrac{|\mathbf{A}|}{\alpha}$ is a characteristic root of Adj* **A**.

Solution. Since α is a characteristic root of a non-singular matrix, therefore $\alpha \neq 0$. Also α is a characteristic root of **A** implies that there exists a non-zero vector **X** such that

$$\mathbf{AX}=\alpha\mathbf{X}$$
$$\Rightarrow (\text{Adj } \mathbf{A})(\mathbf{AX})=(\text{Adj } \mathbf{A})(\alpha\mathbf{X})$$
$$\Rightarrow [(\text{Adj } \mathbf{A})\mathbf{A}]\mathbf{X}=\alpha(\text{Adj } \mathbf{A})\mathbf{X}$$
$$\Rightarrow |\mathbf{A}|\mathbf{IX}=\alpha(\text{Adj } \mathbf{A})\mathbf{X} \qquad [\because (\text{Adj } \mathbf{A})\mathbf{A}=|\mathbf{A}|\mathbf{I}]$$
$$\Rightarrow |\mathbf{A}|\mathbf{X}=\alpha(\text{Adj } \mathbf{A})\mathbf{X} \qquad [\because \mathbf{IX}=\mathbf{X}]$$
$$\Rightarrow \frac{|\mathbf{A}|}{\alpha}\mathbf{X}=(\text{Adj } \mathbf{A})\mathbf{X} \qquad [\because \alpha\neq 0]$$
$$\Rightarrow (\text{Adj } \mathbf{A})\mathbf{X}=\frac{|\mathbf{A}|}{\alpha}\mathbf{X}.$$

Since **X** is a non-zero vector, therefore from the relation (1) it is obvious that $\dfrac{|\mathbf{A}|}{\alpha}$ is a characteristic root of the matrix Adj **A**.

§ 7. ****The Cayley-Hamilton Theorem.** *Every square matrix satisfies its characteristic equation i.e., if for a square matrix* **A** *of order n,*

$$|\mathbf{A}-\lambda\mathbf{I}|=(-1)^n[\lambda^n+a_1\lambda^{n-1}+a_2\lambda^{n-2}+\ldots+a_n],$$

then the matrix equation

$$\mathbf{X}^n+a_1\mathbf{X}^{n-1}+a_2\mathbf{X}^{n-2}+a_3\mathbf{X}^{n-3}+\ldots+a_n\mathbf{I}=\mathbf{O}$$

is satisfied by $\mathbf{X}=\mathbf{A}$

i.e., $\quad \mathbf{A}^n+a_1\mathbf{A}^{n-1}+a_2\mathbf{A}^{n-2}+\ldots+a_n\mathbf{I}=\mathbf{O}.$

(Nagarjuna 1990; Andhra 81; Meerut 82, 91; I.C.S 86; Agra 88; Madras 80; Kanpur 86; Rohilkhand 90; Patna 86)

Proof. Since the elements of $\mathbf{A}-\lambda\mathbf{I}$ are at most of the first degree in λ, the elements of Adj $(\mathbf{A}-\lambda\mathbf{I})$ are ordinary polynomials in λ of degree $n-1$ or less. Therefore Adj $(\mathbf{A}-\lambda\mathbf{I})$ can be written as a matrix polynomial in λ, given by

$$\text{Adj. }(\mathbf{A}-\lambda\mathbf{I})=\mathbf{B}_0\lambda^{n-1}+\mathbf{B}_1\lambda^{n-2}+\ldots+\mathbf{B}_{n-2}\lambda+\mathbf{B}_{n-1},$$

where $\mathbf{B}_0, \mathbf{B}_1, \ldots, \mathbf{B}_{n-1}$ are matrices of the type $n\times n$ whose elements are functions of a_{ij}'s.

Now $(\mathbf{A}-\lambda\mathbf{I})$ Adj. $(\mathbf{A}-\lambda\mathbf{I})=|\mathbf{A}-\lambda\mathbf{I}|\,\mathbf{I}$ [$\because$ **A** Adj **A** $=|\mathbf{A}|\,\mathbf{I}_n$].

$$\therefore \quad (\mathbf{A}-\lambda\mathbf{I})(\mathbf{B}_0\lambda^{n-1}+\mathbf{B}_1\lambda^{n-2}+\ldots+\mathbf{B}_{n-2}\lambda+\mathbf{B}_{n-1})$$
$$=(-1)^n[\lambda^n+a_1\lambda^{n-1}+\ldots+a_n]\,\mathbf{I}.$$

Comparing coefficients of like powers of λ on both sides, we get

$$-\mathbf{IB}_0=(-1)^n\,\mathbf{I},$$
$$\mathbf{AB}_0-\mathbf{IB}_1=(-1)^n a_1\mathbf{I},$$
$$\mathbf{AB}_1-\mathbf{IB}_2=(-1)^n a_2\mathbf{I},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\mathbf{AB}_{n-1}=(-1)^n a_n\mathbf{I}.$$

Premultiplying these successively by $\mathbf{A}^n, \mathbf{A}^{n-1}, \ldots, \mathbf{I}$ and adding we get

$$\mathbf{O}=(-1)^n[\mathbf{A}^n+a_1\mathbf{A}^{n-1}+a_2\mathbf{A}^{n-2}+\ldots+a_n\mathbf{I}].$$

Thus $\mathbf{A}^n+a_1\mathbf{A}^{n-1}+a_2\mathbf{A}^{n-2}+\ldots+a_{n-1}\mathbf{A}+a_n\mathbf{I}=\mathbf{O}.$ ...(i)

Cor. 1. If **A** be a non-singular matrix, $|\mathbf{A}|\neq 0$. Also $|\mathbf{A}|=(-1)^n a_n$ and therefore $a_n\neq 0$.

Premultiplying (i) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^{n-1}+a_1\mathbf{A}^{n-2}+a_2\mathbf{A}^{n-3}+\ldots+a_{n-1}\mathbf{I}+a_n\mathbf{A}^{-1}=\mathbf{O}$$

or $\quad \mathbf{A}^{-1}=-(1/a_n)[\mathbf{A}^{n-1}+a_1\mathbf{A}^{n-2}+\ldots+a_{n-1}\mathbf{I}].$

Cor. 2. If m be a positive integer such that $m \geqslant n$, then multiplying the result (i) by $\mathbf{A}^{m-n}$, we get

$$\mathbf{A}^m+a_1\mathbf{A}^{m-1}+\ldots+a_n\mathbf{A}^{m-n}=\mathbf{O},$$

showing that any positive integral power $\mathbf{A}^m$ $(m\geqslant n)$ of **A** is linearly expressible in terms of those of lower order.

Solved Examples.

Ex. 1. *Find the characteristic equation of the matrix*

$$\mathbf{A}=\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

and verify that it is satisfied by **A** *and hence obtain* $\mathbf{A}^{-1}$.

(Nagarjuna 1980; Kanpur 84; Agra 83; Meerut 82; Lucknow 85 ;Kerala 69; Rohilkhand 81)

Solution. We have

$$|\mathbf{A}-\lambda \mathbf{I}|=\begin{vmatrix} 2-\lambda & -1 & 1 \\ -1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix}$$

$$=(2-\lambda)\{(2-\lambda)^2-1\}+1\{-1(2-\lambda)+1\}+1\{1-(2-\lambda)\}$$

$$=(2-\lambda)(3-4\lambda+\lambda^2)+(\lambda-1)+(\lambda-1)$$

$$=-\lambda^3+6\lambda^2-9\lambda+4.$$

$\therefore$ the characteristic equation of the matrix **A** is

$$\lambda^3-6\lambda^2+9\lambda-4=0.$$

We are now to verify that

$$\mathbf{A}^3-6\mathbf{A}^2+9\mathbf{A}-4\mathbf{I}=\mathbf{O}. \qquad \text{...(i)}$$

We have

$$\mathbf{I}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \mathbf{A}=\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix},$$

$$\mathbf{A}^2=\mathbf{A}\times\mathbf{A}=\begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix},\ \mathbf{A}^3=\mathbf{A}^2\mathbf{A}=\begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix}.$$

Now we can verify that $\mathbf{A}^3-6\mathbf{A}^2+9\mathbf{A}-4\mathbf{I}$

$$=\begin{bmatrix} 22 & -21 & 21 \\ -21 & 22 & -21 \\ 21 & -21 & 22 \end{bmatrix}-6\begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix}+9\begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}-4\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Multiplying (i) by $\mathbf{A}^{-1}$, we get

$$\mathbf{A}^2-6\mathbf{A}+9\mathbf{I}-4\mathbf{A}^{-1}=\mathbf{O}.$$

$$\therefore \quad \mathbf{A}^{-1}=\tfrac{1}{4}(\mathbf{A}^2-6\mathbf{A}+9\mathbf{I}).$$

Now $\mathbf{A}^2-6\mathbf{A}+9\mathbf{I}$

$$=\begin{bmatrix} 6 & -5 & 5 \\ -5 & 6 & -5 \\ 5 & -5 & 6 \end{bmatrix}+\begin{bmatrix} -12 & 6 & -6 \\ 6 & -12 & 6 \\ -6 & 6 & -12 \end{bmatrix}+\begin{bmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

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

$$\therefore \quad \mathbf{A}^{-1}=\tfrac{1}{4}\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}.$$

Ex. 2. *Obtain the characteristic equation of the matrix*

$$\mathbf{A}=\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$$

and verify that it is satisfied by **A** *and hence find its inverse.*

(Meerut 1988; Kanpur 82; Rohilkhand 91; Delhi 81)

Solution. We have

$$|\mathbf{A}-\lambda \mathbf{I}|=\begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & 2-\lambda & 1 \\ 2 & 0 & 3-\lambda \end{vmatrix}$$

$$=(1-\lambda)(2-\lambda)(3-\lambda)+2[0-2(2-\lambda)]$$

$$=(2-\lambda)[(1-\lambda)(3-\lambda)-4]$$

$$=(2-\lambda)[\lambda^2-4\lambda-1]$$

$$=-(\lambda^3-6\lambda^2+7\lambda+2).$$

$\therefore$ the characteristic equation of **A** is

$$\lambda^3-6\lambda^2+7\lambda+2=0 \qquad \text{...(i)}$$

By the Cayley-Hamilton theorem

$$\mathbf{A}^3-6\mathbf{A}^2+7\mathbf{A}+2\mathbf{I}=\mathbf{O}. \qquad \text{...(ii)}$$

Verification of (ii). We have

$$\mathbf{A}^2=\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}\times\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}=\begin{bmatrix} 5 & 0 & 8 \\ 2 & 4 & 5 \\ 8 & 0 & 13 \end{bmatrix}.$$

Also $\mathbf{A}^3=\mathbf{A}.\mathbf{A}^2=\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}\times\begin{bmatrix} 5 & 0 & 8 \\ 2 & 4 & 5 \\ 8 & 0 & 13 \end{bmatrix}=\begin{bmatrix} 21 & 0 & 34 \\ 12 & 8 & 23 \\ 34 & 0 & 55 \end{bmatrix}.$

Now $\mathbf{A}^3-6\mathbf{A}^2+7\mathbf{A}+2\mathbf{I}$

$$=\begin{bmatrix} 21 & 0 & 34 \\ 12 & 8 & 23 \\ 34 & 0 & 55 \end{bmatrix}-6\begin{bmatrix} 5 & 0 & 8 \\ 2 & 4 & 5 \\ 8 & 0 & 13 \end{bmatrix}$$

$$+7\begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}+2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$=\begin{bmatrix}21 & 0 & 34\\12 & 8 & 23\\34 & 0 & 55\end{bmatrix}-\begin{bmatrix}30 & 0 & 48\\12 & 24 & 30\\48 & 0 & 78\end{bmatrix}$$

$$+\begin{bmatrix}7 & 0 & 14\\0 & 14 & 7\\14 & 0 & 21\end{bmatrix}+\begin{bmatrix}2 & 0 & 0\\0 & 2 & 0\\0 & 0 & 2\end{bmatrix}$$

$$=\begin{bmatrix}30 & 0 & 48\\12 & 24 & 30\\48 & 0 & 78\end{bmatrix}-\begin{bmatrix}30 & 0 & 48\\12 & 24 & 30\\48 & 0 & 78\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{bmatrix}=\mathbf{O}.$$

Hence Cayley-Hamilton theorem is verified. Now we shall compute A^{-1}.

Multiplying (ii) by A^{-1}, we get

$$A^2-6A+7I+2A^{-1}=O.$$

$$\therefore \quad A^{-1}=-\tfrac{1}{2}\,(A^2-6A+7I)$$

$$=-\tfrac{1}{2}\begin{bmatrix}5 & 0 & 8\\2 & 4 & 5\\8 & 3 & 13\end{bmatrix}+3\begin{bmatrix}1 & 0 & 2\\0 & 2 & 1\\2 & 0 & 3\end{bmatrix}-\tfrac{7}{2}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

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

Ex. 3. *Show that the matrix*

$$A=\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}$$

satisfies Cayley-Hamilton theorem. **(Meerut 1987)**

Solution. We have

$$|A-\lambda I|=\begin{vmatrix}0-\lambda & c & -b\\-c & 0-\lambda & a\\b & -a & 0-\lambda\end{vmatrix}$$

$$=-\lambda\,(\lambda^2+a^2)-c\,(c\lambda-ab)-b\,(ac+b\lambda)$$

$$=-\lambda^3-\lambda\,(a^2+b^2+c^2).$$

$\therefore$ the characteristic equation of the matrix A is

$$\lambda^3+\lambda\,(a^2+b^2+c^2)=0.$$

We are to verify that $A^3+(a^2+b^2+c^2)\,A=O$.

We have

$$A^2=\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}\times\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}$$

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

$$\therefore \quad A^3=A^2A=\begin{bmatrix}-c^2-b^2 & ab & ac\\ab & -c^2-a^2 & bc\\ac & bc & -b^2-a^2\end{bmatrix}\times\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}$$

$$=\begin{bmatrix}0 & -c^3-b^2c-a^2c & bc^2+b^3+a^2b\\c^3+a^2c+b^2c & 0 & -ab^2-ac^2-a^3\\-bc^2-b^3-a^2b & ac^2+ab^2+a^3 & 0\end{bmatrix}$$

$$=-(a^2+b^2+c^2)\begin{bmatrix}0 & c & -b\\-c & 0 & a\\b & -a & 0\end{bmatrix}$$

$$=-(a^2+b^2+c^2)\,A.$$

$$\therefore \quad A^3+(a^2+b^2+c^2)\,A$$

$$=-(a^2+b^2+c^2)A+(a^2+b^2+c^2)\,A=0A=\mathbf{O}.$$

Hence A satisfies Cayley-Hamilton theorem.

Ex. 4. *State Cayley-Hamilton theorem. Use it to express* $2A^5-3A^4+A^2-4I$ *as a linear polynomial in A, when*

$$A=\begin{bmatrix}3 & 1\\-1 & 2\end{bmatrix}.$$ **(Meerut 1985)**

Solution. Statement of Cayley-Hamilton theorem. Every square matrix satisfies its characteristic equation.

Now let us find the characteristic equation of the matrix A. We have

$$|A-\lambda I|=\begin{vmatrix}3-\lambda & 1\\-1 & 2-\lambda\end{vmatrix}=(3-\lambda)\,(2-\lambda)+1=\lambda^2-5\lambda+7.$$

The characteristic equation of A is $|A-\lambda I|=0$ *i.e.,* is

$$\lambda^2-5\lambda+7=0 \qquad \ldots(1)$$

By Cayley-Hamilton theorem, the matrix A must satisfy (1). Therefore we have

$$A^2-5A+7I=O \qquad \ldots(2)$$

From (2), we get

$$A^2=5A-7I \qquad \ldots(3)$$

Multiplying both sides of (3) by A, we get

$$A^3=5A^2-7A. \qquad \ldots(4)$$

$$\therefore \quad A^4=5A^3-7A^2 \qquad \ldots(5)$$

and $$A^5=5A^4-7A^3. \qquad \ldots(6)$$

Now $2A^5-3A^4+A^2-4I=2(5A^4-7A^3)-3A^4+A^2-4I$ [Substituting for A^5 from (6)]

$=7A^4-14A^3+A^2-4I$

$=7(5A^3-7A^2)-14A^3+A^2-4I$ [by (5)]

$=21A^3-48A^2-4I$

$=21(5A^2-7A)-48A^2-4I$ [by (4)]

$=57A^2-147A-4I$

$=57(5A-7I)-147A-4I$ [by (3)]

$=138A-403I$, which is a linear polynomial in A.

Ex. 5. *Find the characteristic roots of the matrix* $A=\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}$ *and verify Cayley-Hamilton theorem for this matrix. Find the inverse of the matrix A and also express*

$$A^5-4A^4-7A^3+11A^2-A-10I$$

as a linear polynomial in A.

Solution. The characteristic equation of the matrix A is

$$|A-\lambda I|=0$$

or $$\begin{vmatrix}1-\lambda & 4\\ 2 & 3-\lambda\end{vmatrix}=0$$

or $$(1-\lambda)(3-\lambda)-8=0$$

or $$\lambda^2-4\lambda-5=0 \qquad \ldots(1)$$

or $$(\lambda-5)(\lambda+1)=0.$$

The roots of this equation are $\lambda=5, -1$ and these are the characteristic roots of A.

By Cayley-Hamilton theorem, the matrix A must satisfy its characteristic equation (1). So we must have

$$A^2-4A-5I=O. \qquad \ldots(2)$$

Let us verify it. We have

$$A^2=\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}=\begin{bmatrix}9 & 16\\ 8 & 17\end{bmatrix}$$

Now $$A^2-4A-5I=\begin{bmatrix}9 & 16\\ 8 & 17\end{bmatrix}-\begin{bmatrix}4 & 16\\ 8 & 12\end{bmatrix}-\begin{bmatrix}5 & 0\\ 0 & 5\end{bmatrix}$$

$$=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=O.$$

This verifies the theorem.

Now multiplying (2) by A^{-1}, we get

$$A^2A^{-1}-4AA^{-1}-5IA^{-1}=OA^{-1}$$

or $$A-4I-5A^{-1}=O$$

or $$A^{-1}=\tfrac{1}{5}(A-4I).$$

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

$$\therefore \quad A^{-1}=\tfrac{1}{5}(A-4I)=\tfrac{1}{5}\begin{bmatrix}-3 & 4\\ 2 & -1\end{bmatrix}.$$

The characteristic equation of A is $\lambda^2-4\lambda-5=0$. Dividing the polynomial $\lambda^5-4\lambda^4-7\lambda^3+11\lambda^2-\lambda-10$ by the polynomial $\lambda^2-4\lambda-5$, we get

$$\lambda^5-4\lambda^4-7\lambda^3+11\lambda^2-\lambda-10=(\lambda^2-4\lambda-5)(\lambda^3-2\lambda+3)+\lambda+5.$$

$$\therefore \quad A^5-4A^4-7A^3+11A^2-A-10I$$
$$=(A^2-4A-5I)(A^3-2A+3I)+A+5I.$$

But $A^2-4A-5I=O$.

Therefore we have

$$A^5-4A^4-7A^3+11A^2-A-10I=A+5I,$$

which is a linear polynomial in A.

Exercises

1. Find the characteristic roots of the matrix
$$A=\begin{bmatrix}1 & 2 & 3\\ 0 & -4 & 2\\ 0 & 0 & 7\end{bmatrix}.$$
Verify that the matrix A satisfies its characteristic equation.

2. If $A=\begin{bmatrix}1 & 2\\ -1 & 3\end{bmatrix}$, express $A^6-4A^5+8A^4-12A^3+14A^2$ as a linear polynomial in A.

3. Verify that the matrix $A=\begin{bmatrix}1 & 2 & 1\\ 0 & 1 & -1\\ 3 & -1 & 1\end{bmatrix}$ satisfies its characteristic equation and compute A^{-1}. **(Meerut 1973)**

4. Find the characteristic roots of the matrix
$$A=\begin{bmatrix}1 & 1 & 3\\ 5 & 2 & 6\\ -2 & -1 & -3\end{bmatrix}$$
and verify Cayley-Hamilton theorem.

5. Determine the eigenvalues and eigenvectors of the matrix

$$A=\begin{bmatrix}-2 & 2 & -3\\ 2 & 1 & -6\\ -1 & -2 & 0\end{bmatrix}.$$

6. Show that the matrix $A=\begin{bmatrix}2 & 2 & 1\\ 1 & 3 & 1\\ 1 & 2 & 2\end{bmatrix}$ satisfies Cayley-Hamilton theorem. **(Kanpur 1983; Rohilkhand 80; Agra 79)**
Also determine the characteristic roots and the corresponding characteristic vectors of the matrix A.

7. Verify Cayley-Hamilton theorem for the matrix

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

Hence or otherwise evaluate A^{-1}. **(Meerut 1973, 77, 83)**

8. Verify the Cayley-Hamilton theorem for the matrix

$$A=\begin{bmatrix}1 & \sqrt{2} & 0\\ \sqrt{2} & -1 & 0\\ 0 & 0 & 1\end{bmatrix},$$

and use the result to find A^{-1}. **(I.A.S. 1972)**

9. Define the characteristic equation of a square matrix and show that the matrix $A=\begin{bmatrix}1 & 2 & 3\\ 3 & -2 & 1\\ 4 & 2 & 1\end{bmatrix}$ satisfies its characteristic equation. **(Poona 1970)**

10. Find the characteristic equation of the matrix

$$A=\begin{bmatrix}1 & 3 & 7\\ 4 & 2 & 3\\ 0 & 2 & 1\end{bmatrix}$$

and show that it is satisfied by A. Hence obtain the inverse of the given matrix A. **(Agra 1974)**

11. Verify that the matrix $A=\begin{bmatrix}1 & 2 & 0\\ 2 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}$ satisfies its own characteristic equation. Is it true of every square matrix ? State the theorem that applies here. **(Rohilkhand 1991; Kanpur 82; Agra 80)**

12. Find the characteristic roots and characteristic spaces of the matrix $\begin{bmatrix}1 & 2 & 3\\ 0 & 2 & 3\\ 0 & 0 & 2\end{bmatrix}$. **(Agra 1974)**

13. Show that if the two characteristic roots of a Hermitian matrix of order 2 are equal, the matrix must be a scalar multiple of the unit matrix. **(Agra 1973)**

Answers

1. 1, −4, 7.
2. $-4A+5I$.
3. Characteristic equation is $\lambda^3-3\lambda^2-\lambda+9=0$.

$$A^{-1}=\tfrac{1}{9}\begin{bmatrix}0 & 3 & 3\\ 3 & 2 & -1\\ 3 & -7 & -1\end{bmatrix}.$$

4. All the characteristic roots are zero.
5. Eigenvalues are 5, −3, −3. Corresponding to the eigenvalue 5 an eigenvector is $[1, 2, -1]'$. Two linearly independent eigenvectors corresponding to the eigenvalue −3 are $[-2, 1, 0]'$ and $[3, 0, 1]'$.
6. 5, 1, 1. Corresponding to the characteristic root 5 a characteristic vector is $[1, 1, 1]'$. Two linearly independent characteristic vectors corresponding to the characteristic root 1 are $[1, 0, -1]'$ and $[1, 1, -3]'$.
7. $A^{-1}=\tfrac{1}{5}\begin{bmatrix}4 & 1 & -1\\ -12 & 2 & 3\\ 5 & 0 & 0\end{bmatrix}$.
8. $A^{-1}=-\tfrac{1}{3}\begin{bmatrix}-1 & -\sqrt{2} & 0\\ -\sqrt{2} & 1 & 0\\ 0 & 0 & -3\end{bmatrix}$.
10. $\lambda^3-4\lambda^2-13\lambda-40=0$;

$$A^{-1}=\tfrac{1}{40}\begin{bmatrix}-4 & 11 & -5\\ -4 & 1 & 25\\ 8 & -2 & -10\end{bmatrix}.$$

12. Characteristic roots are 1, 2, 2. Corresponding to the characteristic root 1 a characteristic vector is $[1, 0, 0]'$. Corresponding to the characteristic root 2 a characteristic vector is $[2, 1, 0]'$.

8

Eigenvalues and Eigenvectors

(Continued)

§ 1. Characteristic subspaces of a matrix. Suppose λ is an eigenvalue of a square matrix A of order n. Then every non-zero vector X satisfying the equation

$$(A-\lambda I)\, X=O \qquad \ldots(1)$$

is an eigenvector of A corresponding to the eigenvalue λ. If the matrix $A-\lambda I$ is of rank r, then the equation (1) will possess $n-r$ linearly independent solutions. Each non-zero linear combination of these solutions is also a solution of (1) and therefore it will be an eigenvector of A. The set of all these linear combinations is a subspace of V_n provided we add zero vector also to this set. This subspace of V_n is called **characteristic subspace** of A corresponding to the eigenvalue λ. It is nothing but the column null space of the matrix $A-\lambda I$. Its dimension $n-r$ is the geometric multiplicity of the eigenvalue λ.

§ 2. Relation between algebraic and geometric multiplicities of a characteristic root.

Rank-Multiplicity Theorem. *The geometric multiplicity of a characteristic root cannot exceed its algebraic multiplicity.*

(Punjab 1971)

Proof. Let A be a square matrix of order n. Let λ be an eigenvalue of A with geometric multiplicity m. Then there exist m linearly independent column vectors $X_1, \ldots, X_m$ such that

$$AX_i=\lambda X_i,\ i=1, \ldots, m.$$

The linearly independent set $\{X_1, \ldots, X_m\}$ can be extended to form a basis of V_n. Let $\{X_1, \ldots, X_m, X_{m+1}, \ldots, X_n\}$ be a basis of V_n. Since basis set is linearly independent, therefore the matrix $P=[X_1, \ldots, X_m, X_{m+1}, \ldots, X_n]$ is non-singular.

Now consider the matrix $P^{-1}AP$. Since X_1 is the first column

of P, therefore the first column of $P^{-1}AP$ is

$$\begin{aligned} &=P^{-1}AX_1 \\ &=P^{-1}\lambda X_1 && [\because\ AX_1=\lambda X_1 \text{ from (1)}] \\ &=\lambda P^{-1}X_1. \end{aligned}$$

But $P^{-1}X_1$ is the first column of $P^{-1}P=I$. Therefore the first

column of $P^{-1}AP$ is $=\begin{bmatrix} \lambda \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}$.

Similarly we can show that the 2nd, third,..., mth columns of $P^{-1}AP$ are respectively

$$\begin{bmatrix} 0 \\ \lambda \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}, \begin{bmatrix} 0 \\ 0 \\ \lambda \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}, \ldots, \begin{bmatrix} 0 \\ \vdots \\ \lambda \\ 0 \\ 0 \end{bmatrix}_{n\times 1}$$

Therefore the matrix $P^{-1}AP$ is of the form $\begin{bmatrix} \lambda I_m & A_1 \\ O & B_1 \end{bmatrix}$.

$$\therefore\quad P^{-1}AP-xI=\begin{bmatrix} (\lambda-x)I_m & A_1 \\ O & B_1-xI_{n-m} \end{bmatrix}.$$

$$\therefore\quad |\,P^{-1}AP-xI\,|=(\lambda-x)^m\,|\,B_1-xI_{n-m}\,|. \qquad \ldots(2)$$

From (2), we see that $(\lambda-x)^m$ is a factor of the characteristic polynomial of $P^{-1}AP-xI$. Therefore λ is a characteristic root of $P^{-1}AP$ of the algebraic multiplicity at least m. But A and $P^{-1}AP$ have same characteristic roots. Therefore λ is a characteristic root of A of algebraic multiplicity at least m. Therefore if k is the algebraic multiplicity of λ, then $k \geqslant m$ or $m \leqslant k$.

Note. The above result may also be written as

$$n-\text{rank}\,(A-\lambda I) \leqslant k \qquad [\because\ m=n-\text{rank}\,(A-\lambda I)]$$

or $$\text{rank}\,(A-\lambda I) \geqslant (n-k).$$

Solved Examples

Ex. 1. *If* A *and* B *are two square matrices of the same order, then* AB *and* BA *have the same characteristic roots.*

(Allahabad 1978; I.A.S. 82)

Solution. Suppose r is the rank of the matrix A. Then there exist non-singular matrices P and Q such that $PAQ=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$.

We have $PABP^{-1}=(PAQ)\,(Q^{-1}BP^{-1})$.

Let $Q^{-1}BP^{-1}=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$, where C_{11} is an $r\times r$ matrix.

Then $PABP^{-1}=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}=\begin{bmatrix} C_{11} & C_{12} \\ O & O \end{bmatrix}$ ...(1)

Also $Q^{-1}BAQ=(Q^{-1}BP^{-1})(PAQ)$

$$=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}=\begin{bmatrix} C_{11} & O \\ C_{21} & O \end{bmatrix} \quad ...(2)$$

From (1) and (2) we see that the characteristic roots of

$$P\,(AB)\,P^{-1} \text{ and } Q^{-1}\,(BA)\,Q$$

are the same as those of C_{11} along with $n-r$ roots each equal to 0. But the matrices **AB** and **P (AB) P**$^{-1}$ have the same characteristic roots. Also the matrices **BA** and **Q**$^{-1}$ **(BA) Q** have the same characteristic roots. Hence the matrices **AB** and **BA** have the same characteristic roots

Ex. 2. *Prove that* ± 1 *can be the only real characteristic roots of an orthogonal matrix.*

Solution. We know that the characteristic roots of an orthogonal matrix are of unit modulus. Since ± 1 are the only real numbers of unit modulus, therefore ± 1 are the only real numbers which can be the characteristic roots of an orthogonal matrix.

Ex. 3. *If* **A** *is both real symmetric and orthogonal, prove that all its eigenvalues are* $+1$ *or* -1.

Solution. If **A** is a real symmetric matrix, then all its eigenvalues are real. Further if **A** is orthogonal, then all its eigenvalues must be of unit modulus. Now ± 1 are the only real numbers of unit modulus. Therefore if **A** is both real symmetric and orthogonal, then all its eigenvalues are $+1$ or -1.

Ex. 4. *Show that a characteristic root of every orthogonal matrix of odd order is either* 1 *or* -1.

Solution. Suppose **A** is an orthogonal matrix of odd order n. Since an orthogonal matrix is a real matrix, therefore all the coefficients in the characteristic equation $f(\lambda)\equiv|A-\lambda I|=0$ of **A** are real. Since **A** is of odd order n, therefore the equation $f(\lambda)=0$ is of odd degree n. In a polynomial equation with real coefficients complex roots occur in conjugate pairs. Since the equation $f(\lambda)=0$ is of odd degree, therefore it must have at least one real root. Thus if **A** is an orthogonal matrix of odd order, then **A** must have at least one real characteristic root. But the characteristic roots of an orthogonal matrix are of unit modulus. Hence either 1 or -1 must be a characteristic root of **A**.

Ex. 5. *Find the characteristic roots of the 2-rowed orthogonal matrix* $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ *and verify that they are of unit modulus.*

(Madras 1983)

Solution. Let $A=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

We have $|A-\lambda I|=\begin{vmatrix} \cos\theta-\lambda & -\sin\theta \\ \sin\theta & \cos\theta-\lambda \end{vmatrix}=(\cos\theta-\lambda)^2+\sin^2\theta.$

$\therefore$ The characteristic equation of **A** is

$$(\cos\theta-\lambda)^2+\sin^2\theta=0. \quad ...(1)$$

From (1), we get

$$(\cos\theta-\lambda)^2=-\sin^2\theta$$

or $\cos\theta-\lambda=\pm i\sin\theta$

or $\lambda=\cos\theta\pm i\sin\theta$.

Therefore $\cos\theta\pm i\sin\theta$ are the characteristic roots of **A**. We have $|\cos\theta+i\sin\theta|=\sqrt{(\cos^2\theta+\sin^2\theta)}=1$. Similarly

$$|\cos\theta-i\sin\theta|=1.$$

Ex. 6. *Show that the roots of the equation*

$$\begin{vmatrix} a+x & h & g \\ h & b+x & f \\ g & f & c+x \end{vmatrix}=0$$

are real; *a, b, c, f, g, h being real numbers.*

Solution Let $A=\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$.

Then **A** is a real symmetric matrix. Therefore all the characteristic roots of **A** must be real. Thus all the roots of the equation

$$\begin{vmatrix} a-x & h & g \\ h & b-x & f \\ g & f & c-x \end{vmatrix}=0 \quad ...(1)$$

are real.

If we put $-x$ in place of x in the equation (1), then we get the equation

$$\begin{vmatrix} a+x & h & g \\ h & b+x & f \\ g & f & c+x \end{vmatrix}=0. \quad ...(2)$$

The roots of (2) are then the roots of (1) with their signs changed. Hence all the roots of (2) must also be real.

Ex. 7. *Find the characteristic roots and characteristic vectors of the matrix*

$$\mathbf{A}=\begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}.$$

Also verify the fact that the geometric multiplicity of a characteristic root cannot exceed its algebraic multiplicity.

Solution. The characteristic equation of **A** is

$$\begin{vmatrix} 3-\lambda & 10 & 5 \\ -2 & -3-\lambda & -4 \\ 3 & 5 & 7-\lambda \end{vmatrix}=0. \qquad \ldots(1)$$

On expanding the determinant the equation (1) becomes

$$-\lambda^3+7\lambda^2-16\lambda+12=0. \qquad \ldots(2)$$

The roots of (2) are 2, 2, 3. Therefore the characteristic roots of **A** are 2, 2, 3.

Corresponding to $\lambda=3$, the characteristic vectors are given by

$$\begin{bmatrix} 0 & 10 & 5 \\ -2 & -6 & -4 \\ 3 & 5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$$

or $\begin{bmatrix} 0 & 10 & 5 \\ 1 & 3 & 2 \\ 3 & 5 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$, performing $R_2 \to -\frac{1}{2}R_2$

or $\begin{bmatrix} 0 & 10 & 5 \\ 1 & 3 & 2 \\ 0 & -4 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$, performing $R_3 \to R_3-3R_2$

$\begin{bmatrix} 1 & 3 & 2 \\ 0 & 10 & 5 \\ 0 & -4 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$, performing $R_1 \leftrightarrow R_2$.

The coefficient matrix of these equations is of rank 2. Therefore these equations have one linearly independent solution. These equations are equivalent to the equations

$$x_1+3x_2+2x_3=0,\ 10x_2+5x_3=0,\ -4x_2-2x_3=0.$$

Obviously $x_1=1$, $x_2=1$, $x_3=-2$ is a solution of these equations. The only one linearly independent characteristic vector corresponding to $\lambda=3$ may be taken as

$$\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

We see that the algebraic multiplicity of 3 is 1 and its geometric multiplicity is also 1.

Corresponding to $\lambda=2$, the characteristic vectors are given by

$$\begin{bmatrix} 1 & 10 & 5 \\ -2 & -5 & -4 \\ 3 & 5 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$$

or $\begin{bmatrix} 1 & 10 & 5 \\ 0 & 15 & 6 \\ 0 & -25 & -10 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$, performing $R_2 \to R_2+2R_1,\ R_3 \to R_3-3R_1$

or $\begin{bmatrix} 1 & 10 & 5 \\ 0 & 5 & 2 \\ 0 & -5 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$, by $R_2 \to \frac{1}{3}R_2$ and $R_3 \to \frac{1}{5}R_3$

or $\begin{bmatrix} 1 & 10 & 5 \\ 0 & 5 & 2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\mathbf{O}$, by $R_3 \to R_3+R_2$.

The coefficient matrix of these equations is of rank 2. Therefore these equations have only one linearly independent solution. These equations are equivalent to the equations

$$x_1+10x_2+5x_3=0,\ 5x_2+2x_3=0.$$

Obviously $x_1=5$, $x_2=2$, $x_3=-5$ is a solution of these equations. Therefore the only one linearly independent characteristic vector corresponding to $\lambda=2$ may be taken as

$$\begin{bmatrix} 5 \\ 2 \\ -5 \end{bmatrix}.$$

Since $\lambda=2$ is a multiple root of order 2 of the equation (2), therefore algebraic multiplicity of 2 is 2. The number of linearly independent characteristic vectors corresponding to $\lambda=2$ is one. So the geometric multiplicity of $\lambda=2$ is one. Obviously $1<2$.

Ex. 8. *Let $f(x)$ be any scalar polynomial in x. Show that if* **X** *is a characteristic vector of a matrix* **A** *corresponding to the characteristic root λ of* **A**, *then* **X** *is also a characteristic vector of the matrix f*(**A**) *and $f(\lambda)$ is the corresponding characteristic root of the matrix f*(**A**). **(I. C. S. 1987)**

Solution. Suppose λ is any characteristic root of **A** and **X** is a corresponding characteristic vector of **A**. Then

$$\mathbf{AX}=\lambda\mathbf{X} \qquad \ldots(1)$$

$$\Rightarrow \mathbf{A}(\mathbf{AX})=\mathbf{A}(\lambda\mathbf{X})$$

$$\Rightarrow \mathbf{A}^2\mathbf{X}=\lambda(\mathbf{AX})$$

$$\Rightarrow \quad A^2X=\lambda(\lambda X) \qquad \text{[from (1)]}$$
$$\Rightarrow \quad A^2X=\lambda^2X. \qquad \ldots(2)$$

The equation (2) shows that X is also a characteristic vector of A^2 and λ^2 is the corresponding characteristic root of A^2.

If m is any positive integer, then repeating the above process m times, we obtain
$$A^mX=\lambda^mX. \qquad \ldots(3)$$

The equation (3) shows that X is also a characteristic vector of A^m and λ^m is the corresponding characteristic root of A^m.

Now let $f(x)=a_0+a_1x+a_2x^2+\ldots+a_kx^k$.

Then $f(A)=a_0I+a_1A+a_2A^2+\ldots+a_kA^k$.

$$\begin{aligned}
\therefore \quad f(A)X &= (a_0I+a_1A+a_2A^2+\ldots+a_kA^k)X\\
&= a_0X+a_1AX+a_2A^2X+\ldots+a_kA^kX\\
&= a_0X+a_1\lambda X+a_2\lambda^2X+\ldots+a_k\lambda^kX \qquad \text{[from (3)]}\\
&= (a_0+a_1\lambda+a_2\lambda^2+\ldots+a_k\lambda^k)X\\
&= f(\lambda)X. \qquad \ldots(4)
\end{aligned}$$

The equation (4) shows that X is also a characteristic vector of $f(A)$ and $f(\lambda)$ is the corresponding characteristic value.

Ex. 9. *Let $f(x)$ and $g(x)$ be any two scalar polynomials in x and let X be a characteristic vector of a matrix A corresponding to the characteristic root λ. If $g(A)$ is non-singular, show that X is also a characteristic vector of the matrix*
$$f(A)\,[g(A)]^{-1} \text{ and } f(\lambda)/g(\lambda)$$
is the corresponding characteristic root.

Solution. Let λ be any characteristic root of A and let X be a corresponding characteristic vector. Then as in Ex. 8, $f(\lambda)$ is a characteristic root of $f(A)$ and $g(\lambda)$ is a characteristic root of $g(A)$, X being a characteristic vector in each case.

$$\therefore \quad f(A)X=f(\lambda)X \qquad \ldots(1)$$
and
$$g(A)X=g(\lambda)X. \qquad \ldots(2)$$

Since $g(A)$ is non-singular, therefore $g(\lambda)\neq 0$ and $1/g(\lambda)$ is a characteristic root of the matrix $[g(A)]^{-1}$, X being a corresponding characteristic vector.

$$\therefore \quad [g(A)]^{-1}X=[g(\lambda)]^{-1}X. \qquad \ldots(3)$$

Pre-multiplying (3) throughout by $f(A)$, we get
$$f(A)\,[g(A)]^{-1}X=f(A)\,[g(\lambda)]^{-1}X$$
$$\Rightarrow (f(A)\,[g(A)]^{-1})X=[g(\lambda)]^{-1}f(A)X$$
$$\Rightarrow (f(A)\,[g(A)]^{-1})X=[g(\lambda)]^{-1}f(\lambda)X \qquad \text{[using (1)]}$$
$$\Rightarrow (f(A)\,[g(A)]^{-1})X=[f(\lambda)/g(\lambda)]X. \qquad \ldots(4)$$

The equation (4) shows that X is also a characteristic vector of $f(A)\,[g(A)]^{-1}$ and $f(\lambda)/g(\lambda)$ is the corresponding characteristic root.

The Construction of Orthogonal matrices.

The following example gives a very simple method of constructing orthogonal matrices.

Ex. 10. *Suppose S is an n-rowed real skew-symmetric matrix and I is the unit matrix of order n. Then show that*

(*i*) I−S *is non-singular*; **(I.C.S. 1989)**

(*ii*) $A=(I+S)(I-S)^{-1}$ *is orthogonal*; **(I.C.S. 1989)**

(*iii*) $A=(I-S)^{-1}(I+S)$;

(*iv*) *If X is a charasteristic vector of S corresponding to the characteristic root λ, then X is also a characteristic vector of A and $(1+\lambda)/(1-\lambda)$ is the corresponding characteristic root.*

Solution. (i) Since S is a real skew-symmetric matrix, therefore the characteristic roots of S are either zero or pure imaginaries. Therefore the roots of the equation $|S-\lambda I|=0$ are either pure imaginaries or zero. Therefore 1 is not a root of the equation
$$|S-\lambda I|=0. \text{ So } |S-I|\neq 0$$
$\Rightarrow$ S−I is non-singular $\Rightarrow$ I−S is non-singular.

(ii) Let $A=(I+S)(I-S)^{-1}$. Then $A'=[(I+S)(I-S)^{-1}]'$
$$=[(I-S)^{-1}]'(I+S)'=[(I-S)']^{-1}(I+S)'.$$

But $(I-S)'=I'-S'$
$$=I+S \qquad [\because \text{ S is skew symmetric} \Rightarrow S'=-S]$$

Also $(I+S)'=I'+S'=I-S$.
$$\therefore \quad A'=(I+S)^{-1}(I-S).$$
$$\begin{aligned}
\therefore \quad A'A &= (I+S)^{-1}(I-S)(I+S)(I-S)^{-1}\\
&= (I+S)^{-1}(I+S)(I-S)(I-S)^{-1}
\end{aligned}$$
[$\because$ I−S and I+S commute as is quite obvious]
$$=(I)(I)=I.$$

Thus A is orthogonal.

(iii) Since I+S and I−S commute, therefore
$$(I+S)(I-S)=(I-S)(I+S).$$

Pre-multiplying throughout by $(I-S)^{-1}$ and post-multiplying throughout by $(I-S)^{-1}$, we get
$$(I-S)^{-1}(I+S)(I-S)(I-S)^{-1}=(I-S)^{-1}(I-S)(I+S)(I-S)^{-1}$$
or
$$(I-S)^{-1}(I+S)=(I+S)(I-S)^{-1}$$
or
$$A=(I+S)(I-S)^{-1}=(I-S)^{-1}(I+S).$$

(iv) Suppose λ is a characteristic root of S and X is the corresponding characteristic vector. Then $SX=\lambda X$.

$\therefore \quad X+SX=X+\lambda X$

$\Rightarrow \quad (I+S)X=(1+\lambda)X$ ...(1)

Similarly $\quad (I-S)X=(1-\lambda)X.$ ...(2)

Pre-multiplying (2) throughout by $(I-S)^{-1}$, we get

$$(I-S)^{-1}(I-S)X=(1-\lambda)(I-S)^{-1}X$$

or $\quad X=(1-\lambda)(I-S)^{-1}X$

or $\quad (1-\lambda)^{-1}X=(I-S)^{-1}X \qquad [\because 1-\lambda\neq 0]$

or $\quad (I-S)^{-1}X=(1-\lambda)^{-1}X.$ ...(3)

Now pre-multiplying (1) throughout by $(I-S)^{-1}$, we get

$$(I-S)^{-1}(I+S)X=(1+\lambda)(I-S)^{-1}X$$

or $\quad (I-S)^{-1}(I+S)X=(1+\lambda)(1-\lambda)^{-1}X \qquad$ [from (3)]

$\therefore$ X is a characteristic vector of $A=(I-S)^{-1}(I+S)$ and $(1+\lambda)/(1-\lambda)$ is the corresponding characteristic root.

Ex. 11. *If* A *be an orthogonal matrix with the property that* -1 *is not a characteristic root, then* A *is expressible as*

$$(I+S)(I-S)^{-1}$$

for some suitable real skew-symmetric matrix S.

Solution. The problem will be solved if we show that corresponding to a given orthogonal matrix A, such that, -1, is not a characteristic root of A, the equation

$$A=(I+S)(I-S)^{-1} \qquad ...(1)$$

is solvable for S and the solution is a skew-symmetric matrix.

Post-multiplying both sides of (1) by $I-S$, we get

$A(I-S)=(I+S)$

$\Rightarrow A-AS=I+S$

$\Rightarrow A-I=AS+S$

$\Rightarrow (A-I)=(A+I)S.$ ...(2)

Since -1 is not a characteristic root of A, therefore $|A+I|\neq 0$, *i.e.* $A+I$ is non-singular. Therefore pre-multiplying both sides of (1) by $(A+I)^{-1}$, we get $(A+I)^{-1}(A-I)=S$. Thus (1) is solvable for S. Since A is a real matrix, therefore S is also a real matrix. Now it remains to show that S is a skew-symmetric matrix. We have

$$S'=[(A+I)^{-1}(A-I)]'=(A-I)'[(A+I)^{-1}]'$$
$$=(A-I)'[(A+I)']^{-1}=(A'-I')[(A'+I')]^{-1}$$
$$=(A'-I)(A'+I)^{-1} \qquad ...(3)$$

Now it can be easily seen that $A'-I$ and $A'+I$ commute. Therefore

$$(A'+I)(A'-I)=(A'-I)(A'+I)$$
$$\Rightarrow (A'+I)^{-1}(A'+I)(A'-I)(A'+I)^{-1}=(A'+I)^{-1}(A'-I)(A'+I)(A'+I)^{-1}$$
$$\Rightarrow (A'-I)(A'+I)^{-1}=(A'+I)^{-1}(A'-I).$$

$\therefore$ From (3), we get

$$S'=(A'+I)^{-1}(A'-I)=(A'+A'A)^{-1}(A'-A'A)$$

$[\because$ A is orthogonal $\Rightarrow A'A=I]$

$$=[A'(I+A)]^{-1}[A'(I-A)]=(I+A)^{-1}(A')^{-1}A'(I-A)$$
$$=(I+A)^{-1}(I-A)=(A+I)^{-1}(I-A)=-S.$$

$\therefore$ S is skew-symmetric.

Ex. 12. *If* S *is a skew-Hermitian matrix, show that the matrices* $I-S$ *and* $I+S$ *are both non-singular. Also show that*

$$A=(I+S)(I-S)^{-1}$$

is a unitary matrix.

Solution. Since S is a skew-Hermitian matrix, therefore the eigenvalues of S are either pure imaginaries or zero. So neither 1 nor -1 is a root of the equation $|S-\lambda I|=0$. Therefore neither $|S-I|=0$ nor $|S+I|=0$. So both $I-S$ and $I+S$ are non-singular matrices.

Now let $A=(I+S)(I-S)^{-1}$. Then $A^\theta=[(I-S)^{-1}]^\theta(I+S)^\theta$

$$=[(I-S)^\theta]^{-1}(I+S)^\theta=(I^\theta-S^\theta)^{-1}(I^\theta+S^\theta)$$
$$=(I+S)^{-1}(I-S) \; [\because \text{ S is skew-Hermitian} \Rightarrow S^\theta=-S]$$

$\therefore \quad A^\theta A=(I+S)^{-1}(I-S)(I+S)(I-S)^{-1}$

$$=(I+S)^{-1}(I+S)(I-S)(I-S)^{-1}$$

$[\because I+S$ and $I-S$ commute$]$

$=I.$

$\therefore$ A is unitary.

Exercises

1. Find the eigenvalues and eigenvectors of the matrix

$$A=\begin{bmatrix}3 & 2 & 4\\ 2 & 0 & 2\\ 4 & 2 & 3\end{bmatrix}.$$

2. Verify the fact that the geometric multiplicity of a characteristic value cannot exceed its algebraic multiplicity for the following matrices :

(a) $\begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}$ (b) $\begin{bmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}$.

(c) $\begin{bmatrix} 5 & 4 & -4 \\ 4 & 5 & -4 \\ -1 & -1 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

3. **What are** the eigenvalues and eigenvectors of the identity **matrix ?**

4. **Verify that the matrix** $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 3 & -1 & 1 \end{bmatrix}$ satisfies the charac**teristic equation.** Hence find its inverse. **(Delhi Hons. 1960)**

5. **If** $A=\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$,

show that for every integer $n \geq 3$, $A^n=A^{n-2}+A^2-I$. Hence **determine** A^{50}. **(Rajasthan 1967)**

6. If A is an odd order orthogonal matrix, show that either $A-I$ or $A+I$ is necessarily singular.

7. If $\lambda_1,\dots,\lambda_n$ are the characteristic roots of a matrix A, then, show that $\lambda_1^2,\dots,\lambda_n^2$ are the characteristic roots of A^2.

8. If H is a Hermitian matrix, show that $A=(I+iH)^{-1}(I-iH)$ is a unitary matrix. Also show that $A=(I-iH)(I+iH)^{-1}$. Further show that if λ is an eigenvalue of H, then $(1-i\lambda)/(1+i\lambda)$ is an eigenvalue of A.

9. If S is a real skew-symmetric matrix, then show that $I+S$ is non-singular and $(I-S)(I+S)^{-1}$ is orthogonal. **(Madras 1983)**

10. If A is an orthogonal matrix with the property that -1 is not a characteristic root, then A is expressible as $(I-S)(I+S)^{-1}$ for some suitable skew-symmetric matrix.

11. Show that every unitary matrix A can, by a suitable choice of skew-Hermitian matrix S, be expressed as

$$A=(I+S)(I-S)^{-1},$$

provided that, -1, is not a characteristic root of A.

12. If H is any Hermitian matrix, then

$$A=(H+iI)^{-1}(H-iI)=(H-iI)(H+iI)^{-1}$$

is unitary and every unitary matrix can be thus expressed provided, -1, is not a characteristic root of A.

13. If U is a unitary matrix and $|I-U|\neq 0$, prove that the matrix H defined by setting $iH=(I+U)(I+U)^{-1}$, is Hermitian. If $e^{i\theta_1},\dots,e^{i\theta_n}$ be the eigenvalues of U, show that eigenvalues of H are $\cot(\theta_1/2),\dots,\cot(\theta_n/2)$.

Answers

1. 8, -1, -1; linearly independent eigenvectors are

$$\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

3. All eigenvalues are 1. Every non-zero vector is an eigenvector.

4. $\begin{bmatrix} 0 & 1/3 & 1/3 \\ 1/3 & 2/9 & -1/9 \\ 1/3 & -7/9 & -1/9 \end{bmatrix}$. 5. $A^{50}=\begin{bmatrix} 1 & 0 & 0 \\ 25 & 1 & 0 \\ 25 & 0 & 1 \end{bmatrix}$.

§ 3. Minimal Polynomial and minimal Equation of a matrix.

Suppose $f(x)$ is a polynomial in the indeterminate x and A is a square matrix of order n. If $f(A)=O$, then we say that the polynomial $f(x)$ *annihilates* the matrix A. We know that every matrix satisfies its characteristic equation. Also the characteristic polynomial of a matrix A is a zon-zero polynomial *i.e.* a polynomial in which the coefficients of various terms are not all zero. Note that in $|A-xI|$, the coefficient of x^n is $(-1)^n$ which is not zero. Therefore at least the characteristic polynomial of A is a non-zero polynomial that annihilates A. Thus the set of those non-zero polynomials which annihilate A is not empty.

Monic polynomial. Definition. A polynomial in x in which the coefficient of the highest power of x is unity is called a monic polynomial. The coefficient of the highest power of x is also called the leading coefficient of the polynomial. Thus x^3-2x^2+5x+5 is a monic polynomial of degree 3 over the field of real numbers. In this polynomial the leading coefficient is 1.

Among those non-zero polynomials which annihilate a matrix A, the polynomial which is monic and which is of the lowest degree is of special interest. It is called the *minimal polynomial* of the matrix A.

Minimal polynomial of a matrix. Definition. *The monic polynomial of lowest degree that annihilates a matrix* A *is called the minimal polynomial of* A. *Also if $f(x)$ is the minimal polynomial of* A, *the equation $f(x)=0$ is called the minimal equation of the matrix* A. **(Punjab 1971)**

If A is of order n, then its characteristic polynomial is of degree n. Since the characteristic polynomial of A annihilates A, therefore the minimal polynomial of A cannot be of degree greater than n. Its degree must be less than or equal to n.

Theorem 1. *The minimal polynomial of a matrix is unique.*

Proof. Suppose the minimal polynomial of a matrix **A** is of degree r. Then no non-zero polynomial of degree less than r can annihilate A. Let $f(x)=x^r+a_1x^{r-1}+a_2x^{r-2}+\ldots+a_{r-1}\,x+a_r$ and $g(x)=x^r+b_1x^{r-1}+b_2x^{r-2}+\ldots+b_{r-1}\,x+b_r$ be two minimal polynomials of A. Then both $f(x)$ and $g(x)$ annihilate A. Therefore we have $f(\mathbf{A})=\mathbf{O}$ and $g(\mathbf{A})=\mathbf{O}$. These give

$$\mathbf{A}^r+a_1\mathbf{A}^{r-1}+\ldots+a_{r-1}\mathbf{A}+a_r\mathbf{I}=\mathbf{O}, \qquad \ldots(1)$$

and $$\mathbf{A}^r+b_1\mathbf{A}^{r-1}+\ldots+b_{r-1}\mathbf{A}+b_r\mathbf{I}=\mathbf{O}. \qquad \ldots(2)$$

Subtracting (1) from (2), we get

$$(b_1-a_1)\,\mathbf{A}^{r-1}+\ldots+(b_r-a_r)\,\mathbf{I}=\mathbf{O}. \qquad \ldots(3)$$

From (3), we see that the polynomial $(b_1-a_1)\,x^{r-1}+\ldots+(b_r-a_r)$ also annihilates A. Since its degree is less than r, therefore it must be a zero polynomial. This gives $b_1-a_1=0, b_2-a_2=0,\ldots,b_r-a_r=0$. Thus $a_1=b_1,\ldots,a_r=b_r$. Therefore $f(x)=g(x)$ and thus the minimal polynomial of A is unique.

Theorem 2. *The minimal polynomial of a matrix is a divisor of every polynomial that annihilates this matrix.*

(Nagarjuna 1990, Punjab 71)

Proof. Suppose $m(x)$ is the minimal polynomial of a matrix **A. Let $h(x)$** be any polynomial that annihilates A. Since $m(x)$ and $h(x)$ are two polynomials, therefore by the division algorithm there exist two polynomials $q(x)$ and $r(x)$ such that

$$h(x)=m(x)\,q(x)+r(x), \qquad \ldots(1)$$

where either $r(x)$ is a zero polynomial or its degree is less than the degree of $m(x)$. Putting $x=\mathbf{A}$ on both sides of (1), we get

$$h(\mathbf{A})=m(\mathbf{A})\,q(\mathbf{A})+r(\mathbf{A})$$

$$\Rightarrow \mathbf{O}=\mathbf{O}\,q(\mathbf{A})+r(\mathbf{A}) \quad [\because \text{ both } m(x) \text{ and } h(x) \text{ annihilate A}]$$

$$\Rightarrow r(\mathbf{A})=\mathbf{O}.$$

Thus $r(x)$ is a polynomial which also annihilates A. If $r(x)\neq 0$, then it is a non-zero polynomial of degree smaller than the degree of the minimal polynomial $m(x)$ and thus we arrive at a contradiction that $m(x)$ is minimal polynomial of A. Therefore $r(x)$ must be a zero polynomial. Then (1) gives

$$h(x)=m(x)\,q(x) \Rightarrow m(x) \text{ is a divisor of } h(x).$$

Corollary 1. *The minimal polynomial of a matrix is a divisor of the characteristic polynomial of that matrix.*

(Nagarjuna 1977; Andhra 90)

Proof. Suppose $f(x)$ is the characteristic polynomial of a matrix A. Then $f(\mathbf{A})=\mathbf{O}$ by Cayley-Hamilton theorem. Thus $f(x)$ annihilates A. If $m(x)$ is the minimal polynomial of A, then by the above theorem we see that $m(x)$ must be a divisor of $f(x)$.

Corollary 2. *Every root of the Minimal equation of a matrix is also a characteristic root of the matrix.*

Proof. Suppose $f(x)$ is the characteristic polynomial of a matrix A and $m(x)$ is its minimal polynomial. Then $m(x)$ is a divisor of $f(x)$. Therefore there exists a polynomial $q(x)$ such that

$$f(x)=m(x)\,q(x). \qquad \ldots(1)$$

Suppose λ is a root of the equation $m(x)=0$. Then $m(\lambda)=0$. Putting $x=\lambda$ on both sides of (1) we get $f(\lambda)=m(\lambda)g(\lambda)=0\;q(\lambda)=0$. Therefore λ is also a root of $f(x)=0$. Thus λ is also a characteristic root of A.

Derogatory and Non-derogatory Matrices. Definition. *An n-rowed matrix is said to be derogatory or non-derogatory, according as the degree of its minimal equation is less than or equal to n.*

Thus a matrix is non-derogatory if the degree of its minimal polynomial is equal to the degree of its characteristic polynomial.

Theorem 3. *Every root of the characteristic equation of a matrix is also a root of the minimal equation of the matrix.*

(Nagarjuna 1990)

Proof. Suppose $m(x)$ is the minimal polynomial of a matrix A. Then $m(\mathbf{A})=\mathbf{O}$. We know that if λ is a characteristic root of a matrix A and $g(x)$ is any polynomial, then $g(\lambda)$ is a characteristic root of the matrix $g(\mathbf{A})$.

Now suppose λ is a charateristic root of A. Taking $m(x)$ in place of $g(x)$, we see that $m(\lambda)$ is a characteristic root of $m(\mathbf{A})$. But $m(\mathbf{A})$ is a null matrix and each characteristic root of a null matrix is zero. Therefore, we get $m(\lambda)=0 \Rightarrow \lambda$ is a root of the equation $m(x)=0$. Hence every characteristic root of a matrix A is also a root of the minimal polynomial of that matrix.

Note. From theorem 3 and corollary 2 to theorem 2, we conclude that λ is a characteristic root of a matrix if and only if it is a root of the minimal equation of that matrix. Thus if $f(x)$ is the characteristic polynomial of a matrix A and $m(x)$ is the minimal polynomial of A, then both $f(x)$ and $m(x)$ have the same roots though they occur in greater multiplicity in $f(x)$.

Theorem 4. *If the roots of the characteristic equation of a matrix are all distinct, then the matrix is non-derogatory.*

Proof. Suppose A is a matrix of order n whose n characteristic roots are all distinct. We know that each characteristic root of A is also a root of the minimal polynomial of A. Therefore in this case the minimal polynomial of A will be of degree n.

Consequently the matrix will be a non-derogatory matrix. In this case the characteristic equation of A will also give us the minimal equation of A provided we make its leading coefficient unity.

Theorem 5. *The minimal polynomial of an $n \times n$ matrix A is $(-1)^n \mid \mathbf{A}-\lambda \mathbf{I} \mid / g(\lambda)$, where the monic polynomial $g(\lambda)$ is the H.C.F. of the minors of order $n-1$ in $\mid \mathbf{A}-\lambda \mathbf{I} \mid$.*

Proof. We know that $\mid \mathbf{A}-\lambda \mathbf{I} \mid$ can be expressed as a linear combination of the minors of its any row. Since $g(\lambda)$ is a divisor of every minor of order $(n-1)$ of $\mid \mathbf{A}-\lambda \mathbf{I} \mid$, therefore it is also a divisor of $\mid \mathbf{A}-\lambda \mathbf{I} \mid$. So let

$$\mid \mathbf{A}-\lambda \mathbf{I} \mid = (-1)^n g(\lambda)\, h(\lambda), \qquad \ldots(1)$$

where $h(\lambda)$ is some monic polynomial. We claim that $h(\lambda)$ is the minimal polynomial of A.

Each element of Adj. $(\mathbf{A}-\lambda \mathbf{I})$ is numerically equal to some minor of order $n-1$ of $\mid \mathbf{A}-\lambda \mathbf{I} \mid$ Let $\mathbf{B}(\lambda)$ be the matrix obtained on dividing each element of Adj. $(\mathbf{A}-\lambda \mathbf{I})$ by $g(\lambda)$. Then we have

$$\text{Adj}\,(\mathbf{A}-\lambda \mathbf{I}) = g(\lambda)\, \mathbf{B}(\lambda). \qquad \ldots(2)$$

Here $\mathbf{B}(\lambda)$ is an $n \times n$ matrix and is such that its elements are polynomials in λ having no factor (other than a constant) in common.

Pre-multiplying both sides of (2) by $\mathbf{A}-\lambda \mathbf{I}$, we get

$$(\mathbf{A}-\lambda \mathbf{I})\, \text{Adj}\,(\mathbf{A}-\lambda \mathbf{I}) = g(\lambda)(\mathbf{A}-\lambda \mathbf{I})\, \mathbf{B}(\lambda). \qquad \ldots(3)$$

Since $(\mathbf{A}-\lambda \mathbf{I})\, \text{Adj}\,(\mathbf{A}-\lambda \mathbf{I}) = \mid \mathbf{A}-\lambda \mathbf{I} \mid \mathbf{I}$, therefore from (3), we get

$$\mid \mathbf{A}-\lambda \mathbf{I} \mid \mathbf{I} = g(\lambda)(\mathbf{A}-\lambda \mathbf{I})\, \mathbf{B}(\lambda)$$

or $\quad (-1)^n g(\lambda)\, h(\lambda)\, \mathbf{I} = g(\lambda)(\mathbf{A}-\lambda \mathbf{I})\, \mathbf{B}(\lambda)$. [from (1)]

Since $g(\lambda) \neq 0$, therefore cancelling $g(\lambda)$ from both sides, we get

$$(-1)^n h(\lambda)\, \mathbf{I} = (\mathbf{A}-\lambda \mathbf{I})\, \mathbf{B}(\lambda). \qquad \ldots(4)$$

Putting $\lambda = \mathbf{A}$ on both sides of (4), we see that $h(\mathbf{A}) = \mathbf{O}$ Thus the polynomial $h(\lambda)$ annihilates A.

Let $m(\lambda)$ be the minimal polynomial of A. Then

$$m(\mathbf{A}) = \mathbf{O}$$
$$\Rightarrow m(\mathbf{A}) - m(\lambda)\, \mathbf{I} = -m(\lambda)\, \mathbf{I}$$
$$\Rightarrow m(\mathbf{A}) - m(\lambda \mathbf{I}) = -m(\lambda)\, \mathbf{I}$$
$$\Rightarrow (\mathbf{A}-\lambda \mathbf{I})\, \mathbf{L}(\lambda) = -m(\lambda)\, \mathbf{I} \qquad \ldots(5)$$

where $\mathbf{L}(\lambda)$ is a matrix polynomial.

Pre-multiplying both sides of (5) by Adj $(\mathbf{A}-\lambda \mathbf{I})$, we get

$$\text{Adj}(\mathbf{A}-\lambda \mathbf{I})\,(\mathbf{A}-\lambda \mathbf{I})\, \mathbf{L}(\lambda) = -m(\lambda)\, \text{Adj}\,(\mathbf{A}-\lambda \mathbf{I})$$

or $\quad \mid \mathbf{A}-\lambda \mathbf{I} \mid \mathbf{I}\, \mathbf{L}(\lambda) = -m(\lambda)\, \text{Adj}\,(\mathbf{A}-\lambda \mathbf{I})$

or $\quad (-1)^n g(\lambda) h(\lambda) \mathbf{L}(\lambda) = -m(\lambda) g(\lambda)\, \mathbf{B}(\lambda)$ [from (1) and (2)]

or $\quad (-1)^n h(\lambda)\, \mathbf{L}(\lambda) = -m(\lambda)\, \mathbf{B}(\lambda). \qquad \ldots(6)$

[$\because \; g(\lambda) \neq 0$]

From (6), we see that $h(\lambda)$ is a factor of each element of the matrix $m(\lambda)\, \mathbf{B}(\lambda)$. But the elements of $\mathbf{B}(\lambda)$ have no factor in common. Therefore $h(\lambda)$ must be a divisor of $m(\lambda)$. But $h(\lambda)$ annihilates A and $m(\lambda)$ is the minimal polynomial of A. Therefore $m(\lambda)$ is also a divisor of $h(\lambda)$. Since both $h(\lambda)$ and $m(\lambda)$ are monic polynomials, therefore we must have $m(\lambda) = h(\lambda)$. This proves the theorem.

Solved Examples

Ex. 1. *Show that the matrix*

$$\mathbf{A} = \begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix} \textit{ is derogatory.}$$

Solution. We have

$$\mid \mathbf{A}-\lambda \mathbf{I} \mid = \begin{vmatrix} 7-\lambda & 4 & -1 \\ 4 & 7-\lambda & -1 \\ -4 & -4 & 4-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} 7-\lambda & 4 & -1 \\ 4 & 7-\lambda & -1 \\ 0 & 3-\lambda & 3-\lambda \end{vmatrix}, \text{ by } R_3+R_2$$

$$= (3-\lambda) \begin{vmatrix} 7-\lambda & 4 & -1 \\ 4 & 7-\lambda & -1 \\ 0 & 1 & 1 \end{vmatrix}$$

$$= (3-\lambda) \begin{vmatrix} 7-\lambda & 4 & -5 \\ 4 & 7-\lambda & \lambda-8 \\ 0 & 1 & 0 \end{vmatrix}, \text{ by } C_3-C_2$$

$$= -(3-\lambda) \begin{vmatrix} 7-\lambda & -5 \\ 4 & \lambda-8 \end{vmatrix}, \text{ expanding along third row}$$

$$= -(3-\lambda) \begin{vmatrix} 3-\lambda & 3-\lambda \\ 4 & \lambda-8 \end{vmatrix}, \text{ by } R_1-R_2$$

$$= -(3-\lambda)^2 \begin{vmatrix} 1 & 1 \\ 4 & \lambda-8 \end{vmatrix} = -(3-\lambda)^2 (\lambda-12).$$

Therefore the roots of the equation $\mid \mathbf{A}-\lambda \mathbf{I} \mid = 0$ are $\lambda = 3, 3, 12$. These are the characteristic roots of A.

Let us now find the minimal polynomial of A. We know that each characteristic root of A is also a root of its minimal polynomial. So if $m(x)$ is the minimal polynomial of A, then both $x-3$ and $x-12$ are factors of $m(x)$. Let us try whether the polynomial

$h(x)=(x-3)(x-12)=x^2-15x+36$ annihilates **A** or not.

We have $\mathbf{A}^2=\begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix}$.

$$\therefore\ \mathbf{A}^2-15\mathbf{A}+36\mathbf{I}=\begin{bmatrix} 69 & 60 & -15 \\ 60 & 69 & -15 \\ -60 & -60 & 24 \end{bmatrix}$$

$$-15\begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}+\begin{bmatrix} 36 & 0 & 0 \\ 0 & 36 & 0 \\ 0 & 0 & 36 \end{bmatrix}$$

$$=\begin{bmatrix} 105 & 60 & -15 \\ 60 & 105 & -15 \\ -60 & -60 & 60 \end{bmatrix}-\begin{bmatrix} 105 & 60 & -15 \\ 60 & 105 & -15 \\ -60 & -60 & 60 \end{bmatrix}=\mathbf{O}.$$

$\therefore$ $h(x)$ annihilates **A**. Thus $h(x)$ is the monic polynomial of lowest degree which annihilates **A**. Hence $h(x)$ is the minimal polynomial of **A**. Since its degree is less than the order of **A**, therefore **A** is derogatory.

Note. In order to find the minimal polynomial of a matrix **A**, we should not forget that each characteristic root of **A** must also be a root of the minimal polynomial. We should try to find the monic polynomial of lowest degree which annihilates **A** and which has also all the characteristic roots of **A** as its roots.

Ex. 2. *Show that the matrix*

$$\mathbf{A}=\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \text{ is derogatory.}$$

Solution. We have

$$|\mathbf{A}-\lambda\mathbf{I}|=\begin{vmatrix} 1-\lambda & 0 & 0 \\ 1 & -1-\lambda & 0 \\ 1 & 0 & -1-\lambda \end{vmatrix}=(1-\lambda)(1+\lambda)^2.$$

$\therefore$ the characteristic roots of **A** are $\lambda=1, -1, -1$.

Now both $(x-1)$ and $(x+1)$ must be factors of the minimal polynomial of **A**. Let us see whether $h(x)=(x-1)(x+1)=x^2-1$ annihilates **A** or not. We have

$$\mathbf{A}^2=\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

$$=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}=\mathbf{I}.$$

Therefore $\mathbf{A}^2-\mathbf{I}=\mathbf{O}$. Thus $h(x)$ annihilates **A**. Therefore $h(x)$ is the minimal polynomial of **A**. Therefore **A** is derogatory because degree of $h(x)$ is less than 3.

Ex. 3. *Show that every unit matrix of order* $\geqslant 2$ *is derogatory.*

Solution. Let **I** be a unit matrix of order n where $n \geqslant 2$. We see that the polynomial $m(x)=x-1$ annihilates **I**. Therefore $x-1$ is the minimal polynomial of **I**. Since degree of $x-1$ is 1 which is less than n, therefore **I** is derogatory.

Ex. 4 *Find the minimal polynomial of the* $n \times n$ *matrix* **A** *each of whose elements is* 1.

Solution. Let **A** be the $n \times n$ matrix each of whose elements is 1. Then $\mathbf{A}^2=n\mathbf{A}$. Therefore the polynomial x^2-nx annihilates **A**. Now the polynomial $x+a$ cannot annihilate **A** whatever may be the value of the scalar a. Therefore x^2-nx is the monic polynomial of lowest degree which annihilates **A**. Hence x^2-nx is the minimal polynomial of **A**.

Ex. 5. *If* **A** *and* **B** *be* $n \times n$ *matrices and* **B** *be non-singular, then show that* **A** *and* $\mathbf{B}^{-1}\mathbf{AB}$ *have the same minimal polynomial.*

Solution. First we shall show that a monic polynomial $f(x)$ annihilates **A** if and only if it annihilates $\mathbf{B}^{-1}\mathbf{AB}$. We have

$$(\mathbf{B}^{-1}\mathbf{AB})^2=\mathbf{B}^{-1}\mathbf{AB}\,\mathbf{B}^{-1}\mathbf{AB}=\mathbf{B}^{-1}\mathbf{A}^2\mathbf{B}.$$

Proceeding in this way we can show that $(\mathbf{B}^{-1}\mathbf{AB})^k=\mathbf{B}^{-1}\mathbf{A}^k\mathbf{B}$, where k is any positive integer.

Let $f(x)=x^r+a_1x^{r-1}+\ldots+a_{r-1}x+a_r$.

Then $f(\mathbf{A})=\mathbf{A}^r+a_1\mathbf{A}^{r-1}+\ldots+a_{r-1}\mathbf{A}+a_r\mathbf{I}$.

Also
$$\begin{aligned} f(\mathbf{B}^{-1}\mathbf{AB}) &= (\mathbf{B}^{-1}\mathbf{AB})^r+\ldots+a_{r-1}(\mathbf{B}^{-1}\mathbf{AB})+a_r\mathbf{I} \\ &= \mathbf{B}^{-1}\mathbf{A}^r\mathbf{B}+\ldots+a_{r-1}(\mathbf{B}^{-1}\mathbf{AB})+a_r\mathbf{B}^{-1}\mathbf{B} \\ &= \mathbf{B}^{-1}(\mathbf{A}^r+a_1\mathbf{A}^{r-1}+\ldots+a_{r-1}\mathbf{A}+a_r\mathbf{I})\mathbf{B} \\ &= \mathbf{B}^{-1}f(\mathbf{A})\mathbf{B}. \end{aligned}$$

Since **B** is non-singular, therefore

$$\mathbf{B}^{-1}f(\mathbf{A})\mathbf{B}=\mathbf{O} \text{ if and only if } f(\mathbf{A})=\mathbf{O}.$$

Thus $f(x)$ annihilates **A** if and only if it annihilates $\mathbf{B}^{-1}\mathbf{AB}$.

Therefore if $f(x)$ is the polynomial of lowest degree that annihilates **A** then it is also the polynomial of lowest degree that annihilates $\mathbf{B}^{-1}\mathbf{AB}$ and conversely.

Hence **A** and $\mathbf{B}^{-1}\mathbf{AB}$ have the same minimal polynomial.

Ex. 6. *A square matrix* **A** *is said to be idempotent if* $\mathbf{A}^2=\mathbf{A}$. *Show that if* **A** *is idempotent, then all eigenvalues of* **A** *are equal to* 1 *or* 0.

Solution. We have $A^2=A \Rightarrow A^2-A=O$

$\Rightarrow$ A satisfies the polynomial $\lambda^2-\lambda$

$\Rightarrow \lambda^2-\lambda$ annihilates A

$\Rightarrow$ the minimal polynomial $m(\lambda)$ of A divides $\lambda(\lambda-1)$

$\Rightarrow m(\lambda)=\lambda,\ \lambda-1$ or $\lambda(\lambda-1)$

$\Rightarrow \lambda=0$ or 1 are the only roots of $m(\lambda)$.

Now we know that each eigenvalue of A is also a root of its minimal polynomial. Hence all eigenvalues of A are equal to 1 or 0.

Note. Null matrix is the only matrix whose minimal polynomial is λ. Unit matrix is the only matrix whose minimal polynomial is $\lambda-1$. Therefore if $A^2=A$ and $A \neq O$ and $A \neq I$, then the minimal polynomial of A is $\lambda^2-\lambda$.

Exercises

1. Find the minimal polynomial of the matrix

$$\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$$ and show that it is derogatory.

Ans. x^2-3x+2.

2. Show that the minimal polynomial of a non-zero idempotent matrix C $(\neq I)$ is $\lambda^2-\lambda$.

3. Determine the minimal and characteristic equations of the following matrices :

$$\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}, \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

4. Show that the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$$ is non-derogatory.

9

Orthogonal Vectors

§ 1. Inner product of two vectors.

Definition. *Let* **X** *and* **Y** *be two complex n-vectors written as column vectors. Suppose* $X=\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix},\ Y=\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}$.

Then the inner product of **X** *and* **Y**, *denoted by* (X, Y) *is defined as*

$$(X, Y)=\bar{x}_1 y_1+\bar{x}_2 y_2+\dots+\bar{x}_n y_n.$$

Here $\bar{x}_1$ etc. denotes the conjugate complex of the complex number x_1 etc.

For practical purposes we generally identify a 1×1 matrix with its single element. Thus if $[a]_{1\times1}$ is a 1×1 matrix, then we shall simply regard it as the scalar a. With the help of this identification, the inner product of the vectors X and Y may be conveniently defined as

$$(X, Y)=X^\theta Y.$$

Note that X^θ is a $1\times n$ matrix and Y is a $n\times1$ matrix. Therefore $X^\theta Y$ is a 1×1 matrix and it has been taken equal to its element.

If X and Y are real n-vectors written as column vectors, then their inner product is defined as

$$(X, Y)=X^T Y=x_1 y_1+x_2 y_2+\dots+x_n y_n.$$

If X and Y are complex n-vectors, written as row vectors, then their inner product is defined as

$$(X, Y)=X Y^\theta=x_1 \bar{y}_1+x_2 \bar{y}_2+\dots+x_n \bar{y}_n.$$

Note. Generally in our discussion of inner products we shall take the n-vectors X, Y etc. as column vectors unless otherwise stated.

§ 2. Properties of inner Product.

Theorem. *Suppose* **X, Y, Z** *are any three complex n-vectors and k is any complex number, then* (*i*) (X, X) $\geqslant$ 0, *and* (X, X)=0 *if and only if* X=**O**.

(*ii*) (X, Y)=$\overline{(\text{Y, X})}$.

(*iii*) (X, Y+Z)=(X, Y)+(X, Z).

(*iv*) (X+Y, Z)=(X, Z)+(Y, Z).

(*v*) (X, *k*Y)=*k* (X, Y). (*vi*) (*k*X, Y)=$\bar{k}$ (X, Y).

Proof. Let X=$[x_1,\ldots, x_n]^T$, Y=$[y_1,\ldots, y_n]^T$, Z=$[z_1,\ldots, z_n]^T$.

(i) We have (X, X)=$X^\theta X=\bar{x}_1x_1+\bar{x}_2x_2+\ldots+\bar{x}_nx_n$

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

[$\because$ if z is a complex number, then $|z| \geqslant 0$].

Also (X, X)=0 if and only if $|x_1|^2+\ldots+|x_n|^2=0$

or if and only if $|x_1|^2=0,\ldots, |x_n|^2=0$

or if and only if $x_1=0,\ldots, x_n=0$

or if and only if X=**O**.

(ii) We have $\overline{(\text{Y, X})}=\overline{(Y^\theta X)}$

$=\overline{[(Y^\theta X)^T]}$ [$\because$ $Y^\theta X$ is a 1×1 matrix $\Rightarrow (Y^\theta X)^T=Y^\theta X$]

$=(Y^\theta X)^\theta=X^\theta (Y^\theta)^\theta=X^\theta Y=(X, Y)$.

(iii) $(X, Y+Z)=X^\theta (Y+Z)=X^\theta Y+X^\theta Z=(X, Y)+(X, Z)$

(iv) $(X+Y, Z)=(X+Y)^\theta Z=(X^\theta+Y^\theta) Z$

$=X^\theta Z+Y^\theta Z=(X, Z)+(Y, Z)$.

(v) $(X, kY)=X^\theta (kY)=k (X^\theta Y)=k (X, Y)$.

(vi) $(kX, Y)=(kX)^\theta Y=(\bar{k} X^\theta) Y=\bar{k} (X^\theta Y)=\bar{k} (X, Y)$.

Note 1. If X_1, X_2, X_3, X_4 are any four complex *n*-vectors and a_1, a_2, a_3, a_4 are any four complex numbers, then it can be easily seen that $(a_1X_1+a_2X_2, a_3X_3+a_4X_4)$

$=\bar{a}_1a_3 (X_1, X_3)+\bar{a}_1a_4 (X_1, X_4)+\bar{a}_2a_3 (X_2, X_3)+\bar{a}_2a_4 (X_2, X_4)$.

Note 2. If X, Y are any two real *n*-vectors and k is any real number, then (X, Y)=(Y, X) and (*k*X, Y)=*k* (X, Y) etc.

§ 3. Length of a vector. Definition. *Let X be a complex n-vector. Then the positive square root of the inner product of* X *and* X, *i.e. of* $X^\theta X$, *is called the length of* X.

The length of a vector X is sometimes also called the **norm** of the vector X and is denoted by $\|X\|$. Thus if X=$[x_1, x_2,\ldots, x_n]^T$, then $\|X\|=\sqrt{(X, X)}=\sqrt{(X^\theta X)}=\sqrt{(|x_1|^2+|x_2|^2+\ldots+|x_n|^2)}$.

It is obvious that the length of a vector is zero if and only if the vector is a zero vector.

If X is a real *n*-vector, then $\|X\|=\sqrt{(x_1^2+\ldots+x_n^2)}$.

Unit vector. Definition. *A vector* X *is said to be a unit vector if* $\|X\|=1$. A unit vector is sometimes also called a **normal vector.**

§ 4. Orthogonal vectors. Definition. *Let* X *and* Y *be two complex n-vectors; then* X *is said to be orthogonal to* Y *if*

(X, Y)=0 *i.e. if* $X^\theta Y=0$.

The relation of orthogonality in the set of all complex *n*-vectors is **symmetric.** We have X is orthogonal to Y $\Rightarrow$ (X, Y)=0 $\Rightarrow \overline{(\text{X, Y})}=\bar{0} \Rightarrow$ (Y, X)=0 $\Rightarrow$ Y is orthogonal to X.

So without any ambiguity we can say that two complex *n*-vectors X and Y are orthogonal if and only if (X, Y)=0.

Note 1. If X is orthogonal to Y, then every scalar multiple of X is orthogonal to every scalar multiple of Y. Let *a*, *b* be any two scalars. Then $(aX, bY)=\bar{a}b\,(X, Y)=\bar{a}b0=0$, since (X, Y) ... Thus *a*X and *b*Y are also orthogonal vectors.

Note 2. The zero vector is the only vector which is orthogonal to itself. We have X is orthogonal to X $\Rightarrow$ (X, X)=0 $\Rightarrow$ X=**O**.

Note 3. Two real *n*-vectors

$X=[x_1, x_2,\ldots, x_n]^T$, $Y=[y_1, y_2,\ldots, y_n]^T$

are orthogonal if and only if (X, Y)=0 *i.e.* if and only if $X^TY=0$ *i.e.* if and only if $x_1y_1+\ldots+x_ny_n=0$.

Orthogonal set. Definition. *A set S of complex n-vectors* $X_1, X_2,\ldots, X_k$ *is said to be an orthogonal set if any two distinct vectors in S are orthogonal.*

Orthonormal set. Definition. *A set S of complex n-vectors* $X_1, X_2,\ldots, X_k$ *is said to be an orthonormal set if*

(*i*) *each vector in S is a unit vector.*

(*ii*) *any two distinct vectors in S are orthogonal.*

(Nagarjuna 1980)

Kronecker delta. *The symbol* δ_{ij} *is said to be Kronecker delta if*

$\delta_{ij}=0$ *when* $i\neq j$

and $\delta_{ij}=1$ *when* $i=j$.

In terms of Kronecker delta the unit matrix I_n can also be written as $I_n=[\delta_{ij}]_{n\times n}$.

In terms of Kronecker delta an orthonormal set may be defined as follows :

A set S of complex n-vectors $X_1, X_2,\ldots, X_k$ *is said to be an orthonormal set if*

$(X_i, X_j)=\delta_{ij}$, $i=1, 2,\ldots, k$, $j=1, 2,\ldots, k$.

§ 5. Properties of Orthogonal sets.

Theorem 1. *Every orthogonal set of non-zero vectors is linearly independent.*

Proof. Let $S=\{X_1, X_2, ..., X_k\}$ be an orthogonal set of non-zero complex n-vectors. Then to prove that S is linearly independent. Let $c_1, c_2, ..., c_k$ be scalars such that

$$c_1X_1+c_2X_2+...+c_kX_k=O. \qquad ...(1)$$

Let $1 \leqslant m \leqslant k$. Then forming inner products of both sides of (1) with the vector X_m, we get

$$(X_m, c_1X_1+c_2X_2+...+c_kX_k)=(X_m, O)$$

or $c_1(X_m, X_1)+c_2(X_m, X_2)+...+c_k(X_m, X_k)=0$ [$\because (X_m, O)=0$]

or $c_m(X_m, X_m)=0$ [$\because$ any two distinct vectors of S are orthogonal]

or $c_m=0$, since $X_m \neq O \Rightarrow (X_m, X_m) \neq 0$.

Thus $c_m=0$, where $m=1, 2, ..., k$. In this way the relation (1) implies that $c_1=0, ..., c_k=0$. Therefore the vectors $X_1, ..., X_k$ are linearly independent.

Corollary. *If n non-zero n-vectors form an orthogonal set, then they constitute a basis of* V_n.

Proof. Let $S=\{X_1, ..., X_n\}$ be an orthogonal set of n non-zero n-vectors. Then S is linearly independent. Since dimension of V_n is n and S is a linearly independent subset of V_n containing n vectors, therefore S is a basis of V_n.

Theorem 2. *Every orthonormal set of vectors is linearly independent.*

If X is a vector belonging to an orthonormal set, then

$$(X, X)=1 \neq 0.$$

So the proof of theorem 1 will serve the purpose with a slight change of words here and there.

Corollary. *If an orthonormal set S contains n complex n-vectors, then S is a basis of* V_n.

Orthogonal basis. Definition. *If an orthogonal set S is a basis of* V_n, *then it is called an orthogonal basis of* V_n.

Orthonormal basis. Definition. *If an orthonormal set S is a basis of* V_n, *then it is called an orthonormal basis of* V_n.

Theorem 3. *If* $S=\{X_1, ..., X_k\}$ *is an orthogonal set of non-zero complex n-vectors and Y is any complex n-vector, then*

$$Z=Y-\left\{\frac{(X_1, Y)}{(X_1, X_1)}X_1+\frac{(X_2, Y)}{(X_2, X_2)}X_2+...+\frac{(X_k, Y)}{(X_k, X_k)}X_k\right\}$$

is orthogonal to each of the vectors $X_1, X_2, ..., X_k$.

Proof. Let $1 \leqslant m \leqslant k$. Then

$$(X_m, Z)=(X_m, Y)-\left\{\frac{(X_1, Y)}{(X_1, X_1)}(X_m, X_1)+\frac{(X_2, Y)}{(X_2, X_2)}(X_m, X_2)+...+\frac{(X_k, Y)}{(X_k, X_k)}(X_m, X_k)\right\}$$

$$=(X_m, Y)-\frac{(X_m, Y)}{(X_m, X_m)}(X_m, X_m),$$

since any two distinct vectors in S are orthogonal

$$=(X_m, Y)-(X_m, Y)=0.$$

$\therefore$ $(X_m, Z)=0$ for every $1 \leqslant m \leqslant k$.

Hence Z is orthogonal to each of the vectors belonging to S.

Gram Schmidt orthogonalization process.

Theorem 4. *We can always construct an orthogonal basis of the vector space* V_n *with the help of any given basis.*

Proof. Since the complex n-vector space V_n is of finite dimension n, therefore it definitely possesses a basis. Let $S=\{X_1, X_2, ..., X_n\}$ be a basis of V_n. We shall now give a process to construct an orthogonal basis $\{Y_1, Y_2, ..., Y_n\}$ of V_n with the help of the basis S. This process is known as **Gram-Schmidt Orthogonalization process.**

The main idea behind this construction is that we shall construct an orthogonal set $\{Y_1, ..., Y_n\}$ of non-zero vectors in such a way that each Y_j, $1 \leqslant j \leqslant n$ will be expressed as a linear combination of $X_1, ..., X_j$.

Let $Y_1=X_1$. Then $Y_1 \neq O$, since $X_1 \neq O$. Also Y_1 is a linear combination of X_1.

Let $Y_2=X_2-\frac{(Y_1, X_2)}{(Y_1, Y_1)}Y_1$. Then Y_2 is orthogonal to Y_1 as can be easily seen. The vector Y_2 is not zero because otherwise the set $\{X_2, Y_1\}$ or $\{X_2, X_1\}$ will become linearly dependent while it is linearly independent, being a subset of a linearly independent set S. Also Y_2 is a linear combination of X_1, X_2 because $Y_1=X_1$.

Now suppose that we have constructed an orthogonal set $\{Y_1, .., Y_k\}$ of k (where $k < n$) non-zero vectors such that each Y_j $(j=1, ..., k)$ is a linear combination of $X_1, ..., X_j$.

Let
$$Y_{k+1}=X_{k+1}-\left\{\frac{(Y_1, X_{k+1})}{(Y_1, Y_1)}Y_1+\frac{(Y_2, X_{k+1})}{(Y_2, Y_2)}Y_2+...+\frac{(Y_k, X_{k+1})}{(Y_k, Y_k)}Y_k\right\}.$$

Then by theorem 3, Y_{k+1} is orthogonal to each of the vectors $Y_1, ..., Y_k$. Therefore the set $\{Y_1, ..., Y_k, Y_{k+1}\}$ is orthogonal.

Suppose $Y_{k+1}=O$. Then X_{k+1} is a linear combination of $Y_1,\ldots,Y_k$. But according to our assumption each Y_j $(j=1,\ldots,k)$ is a linear combination of $X_1,\ldots,X_j$. Therefore X_{k+1} is a linear combination of $X_1,\ldots,X_k$. This is not possible because $X_1,\ldots,X_k,X_{k+1}$ are linearly independent. Therefore we must have $Y_{k+1}\neq O$. Also Y_{k+1} is a linear combination of $X_1, X_2,\ldots,X_k$ since each Y_j $(j=1,\ldots,k)$ is a linear combination of $X_1,\ldots,X_j$.

Thus we have been able to construct an orthogonal set $\{Y_1,\ldots,Y_k,Y_{k+1}\}$ containing $k+1$ non-zero vectors such that each Y_j $(j=1,\ldots,k+1)$ is a linear combination of $X_1,\ldots,X_j$. Therefore by mathematical induction we can construct an orthogonal set $\{Y_1,\ldots,Y_n\}$ containing n non-zero vectors such that each Y_j $(j=1,\ldots,n)$ is a linear combination of $X_1,\ldots,X_j$. Since an orthogonal set of non-zero vectors is linearly independent, therefore $\{Y_1,\ldots,Y_n\}$ is a linearly independent subset of V_n containing n vectors. Hence it is an orthogonal basis of V_n.

Corollary. *We can always construct an orthonormal basis of the vector space V_n from a given basis.*

Proof. Let $S=\{X_1,\ldots,X_n\}$ be a given basis of V_n. Let $\{Y_1,\ldots,Y_n\}$ be an orthogonal basis of V_n constructed from S with the help of Gram Schmidt orthogonalization process.

Let $Z_1=\dfrac{Y_1}{\|Y_1\|}$, $Z_2=\dfrac{Y_2}{\|Y_2\|},\ldots, Z_n=\dfrac{Y_n}{\|Y_n\|}$.

Then $\{Z_1, Z_2,\ldots,Z_n\}$ is an orthonormal set of n vectors and so is an orthonormal basis of V_n.

Note. If V_n is the vector space over the field of real numbers and $\{X_1,\ldots,X_n\}$ is any basis of V_n, then we can construct an orthonormal basis of V_n from this basis. In this basis obviously all the vectors will be real vectors.

Theorem 5. *If X_1 be a non-zero n-vector, then there exists an orthonormal basis of V_n having X_1 as a member.*

Proof. Since X_1 is a non-zero vector, therefore the set $\{X_1\}$ is a linearly independent subset of V_n. So we can extend it to form a basis of V_n. Let $S=\{X_1, X_2,\ldots,X_n\}$ be a basis of V_n having X_1 as a member. By Gram-Schmidt orthogonalization process we can find an orthogonal basis $\{Y_1, Y_2,\ldots,Y_n\}$ of V_n such that $Y_1=X_1$.

§ 6. Unitary and Orthogonal Matrices.

Unitary Matrix. Definition. *A square matrix P with complex elements is said to be unitary if $P^\theta P=I$.*

If P is a unitary matrix, then $P^\theta P=I$

$\Rightarrow |P^\theta P|=|I| \Rightarrow |P^\theta|\,|P|=1$

$\Rightarrow |P|\neq 0 \Rightarrow$ P is invertible.

Also then $P^\theta P=I \Rightarrow P^\theta=P^{-1}$ which in turn implies $PP^\theta=I$.

Thus P is a unitary matrix if and only if $P^\theta P=I=PP^\theta$ *i.e.* if and only if $P^\theta=P^{-1}$.

Orthogonal Matrix.

Definition. *A square matix P with real elements is said to be orthogonal if $P^T P=I$.*

As in the case of a unitary matrix, it can be easily seen that a real matrix P is orthogonal if and only if $P^T P=I=PP^T$ *i.e.*, if and only if $P^T=P^{-1}$.

Properties of Orthogonal and Unitary Matrices.

Theorem 1. *A real matrix is unitary if and only if it is orthogonal.*

Proof. Suppose A is a real matrix. Then $A^\theta=A^T$. If A is unitary, then $A^\theta A=I \Rightarrow A^T A=I \Rightarrow$ A is orthogonal.

Conversely if A is orthogonal, then $A^T A=I \Rightarrow A^\theta A=I \Rightarrow$ A is unitary.

Theorem 2.

(*a*) *If P is unitary so are P^T, $\bar{P}$, P^θ and P^{-1}.*

(*b*) *If P and Q are unitary so is PQ.*

(*c*) *If P is unitary, then | P | is of unit modulus.*

(*d*) *Any two eigenvectors corresponding to the distinct eigenvalues of a unitary matrix are orthogonal.*

Proof. (a) P is unitary $\Rightarrow P^\theta P=I \Rightarrow (P^\theta P)^T=I^T$

$$\Rightarrow P^T (P^\theta)^T=I \Rightarrow P^T [(\overline{P^T})]^T=I$$

$$\Rightarrow P^T (P^T)^\theta=I \Rightarrow P^T \text{ is unitary.}$$

Further P is unitary $\Rightarrow PP^\theta=I \Rightarrow (\overline{PP^\theta})=\bar{I}$

$$\Rightarrow \bar{P}\,(\overline{P^\theta})=I \Rightarrow \bar{P}\,\overline{[(\bar{P})^T]}=I$$

$$\Rightarrow \bar{P}\,(\bar{P})^\theta=I \Rightarrow \bar{P} \text{ is unitary.}$$

Also P is unitary $\Rightarrow P^\theta P=I \Rightarrow P^\theta (P^\theta)^\theta=I \Rightarrow P^\theta$ is unitary.

Finally P is unitary $\Rightarrow P^\theta P=I$

$$\Rightarrow (P^\theta P)^{-1}=I^{-1} \Rightarrow P^{-1} (P^\theta)^{-1}=I$$

$$\Rightarrow P^{-1} (P^{-1})^\theta=I \Rightarrow P^{-1} \text{ is unitary.}$$

(b) Suppose P and Q are unitary matrices. Then $P^\theta P=I=P\,P^\theta$ and $Q^\theta Q=I=QQ^\theta$.

To prove that **PQ** is also unitary. We have
$(PQ)^\theta (PQ)=Q^\theta P^\theta PQ=Q^\theta (P^\theta P)\, Q=Q^\theta I\, Q=Q^\theta Q=I.$

$\therefore$ **PQ** is unitary.

(c) We have **P** is unitary $\Rightarrow P^\theta P=I$

$\Rightarrow |P^\theta P|=|I| \Rightarrow |P^\theta|.|P|=1$

$\Rightarrow |\overline{(P^T)}|.|P|=1 \Rightarrow \overline{|P^T|}.|P|=1$

$\Rightarrow \overline{|P|}.|P|=1 \Rightarrow |P|$ is of unit modulus.

(d) Let **A** be a unitary matrix and X_1, X_2 be two eigenvectors of **A** corresponding to the eigenvalues λ_1, λ_2 of **A** where $\lambda_1 \neq \lambda_2$.

Then $\quad AX_1=\lambda_1 X_1 \quad \ldots(1)$

and $\quad AX_2=\lambda_2 X_2. \quad \ldots(2)$

Taking conjugate transpose of (2), we get $(AX_2)^\theta=(\lambda_2 X_2)^\theta$

or $\quad X_2^\theta A^\theta=\bar{\lambda}_2 X_2^\theta. \quad \ldots(3)$

Post-multiplying both sides of (3) by AX_1, we get

$X_2^\theta A^\theta AX_1=\bar{\lambda}_2 X_2^\theta AX_1$

or $\quad X_2^\theta X_1=\bar{\lambda}_2 X_2^\theta \lambda_1 X_1 \quad [\because A^\theta A=I \text{ and } AX_1=\lambda_1 X_1]$

or $\quad X_2^\theta X_1=\bar{\lambda}_2 \lambda_1 X_2^\theta X_1$

or $\quad (1-\bar{\lambda}_2\lambda_1)\, X_2^\theta X_1=O. \quad \ldots(4)$

But eigenvalues of a unitary matrix are of unit modulus.

Therefore $\bar{\lambda}_2 \lambda_2=1$. So from (4), we get

$$\left(1-\frac{\lambda_1}{\lambda_2}\right) X_2^\theta X_1=O$$

$$\Rightarrow \left(\frac{\lambda_2-\lambda_1}{\lambda_2}\right) X_2^\theta X_1=O$$

$$\Rightarrow X_2^\theta X_1=O \qquad [\because \lambda_2 \neq \lambda_1 \Rightarrow \lambda_2-\lambda_1 \neq 0]$$

$\Rightarrow X_1$ and X_2 are orthogonal vectors.

Theorem 3.

(*a*) *If* **P** *is orthogonal so are* P^T *and* P^{-1}.

(*b*) *If* **P** *and* **Q** *are orthogonal so is* **PQ**.

(*c*) *If* **P** *is orthogonal, then* $|P|=\pm 1$. **(Madurai 1985)**

Proof. For proof proceed exactly on the same lines as in theorem 2. We shall prove part (c) only.

(c) **P** is orthogonal $\Rightarrow P^T P=I$

$\Rightarrow |P^T P|=|I| \Rightarrow |P^T|.|P|=1$

$\Rightarrow |P|.|P|=1 \Rightarrow (|P|)^2=1 \Rightarrow |P|=\pm 1.$

Definition. *An orthogonal matrix* **P** *is said to be proper if* $|P|=1$, *improper if* $|P|=-1$.

Obviously, P^{-1} is proper or improper according as **P** is proper or improper. Moreover, if **P** and **Q** are both proper or both improper, then **PQ** is proper but if one of **P**, **Q** is proper and one improper, then **PQ** is improper.

Theorem 4. Orthogonal group. *The set of all orthogonal matrices of the same order is a group with respect to the operation of multiplication.*

Proof. Let M be the set of all orthogonal matrices of the same order n. We shall prove that M is a group with respect to multiplication of matrices.

Closure Property. Let **A**, **B** be two orthogonal matrices of the same order n. Then $A^T A=AA^T=I$, $B^T B=I=BB^T$.

We have $\quad (AB)^T (AB)=(B^T A^T)(AB)$
$\qquad =B^T (A^T A)\, B=B^T I\, B=B^T B=I.$

Therefore **AB** is also an orthogonal matrix of order n. Thus if **A**, **B** are in M, then **AB** is also in M.

Associativity. We know that matrix multiplication is associative.

Existence of Identity. Let **I** be unit matrix of order n. Then $I^T I=I \Rightarrow I$ is orthogonal. Thus **I** is also in M.

Existence of Inverse. Let **A** be an orthogonal matrix of order n. Then $\quad A^T A=I \Rightarrow (A^T A)^{-1}=I^{-1}$

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

$\Rightarrow A^{-1}$ is orthogonal $\Rightarrow A^{-1}$ is in M.

Hence the theorem.

Theorem 5. *A square matrix is unitary if and only if its columns (rows) form an orthogonal set of vectors.*

Proof. Let **A** be a square matrix of order n. Let $C_1, C_2, \ldots, C_n$ denote the column vectors of **A**. Then $A=[C_1, C_2, \ldots, C_n]$.

Now $\quad A^\theta A=[C_1, C_2, \ldots, C_n]^\theta\, [C_1, C_2, \ldots, C_n]$

$$=\begin{bmatrix} C_1^\theta \\ C_2^\theta \\ \vdots \\ C_n^\theta \end{bmatrix} [C_1, C_2, \ldots, C_n]$$

$$=\begin{bmatrix} C_1^\theta C_1 & C_1^\theta C_2 & \cdots & C_1^\theta C_n \\ \vdots & \vdots & & \vdots \\ C_n^\theta C_1 & C_n^\theta C_2 & \cdots & C_n^\theta C_n \end{bmatrix}$$

$$=\begin{bmatrix}(C_1, C_1) & (C_1, C_2) & \cdots & (C_1, C_n)\\ \vdots & \vdots & & \vdots \\ (C_n, C_1) & (C_n, C_2) & \cdots & (C_n, C_n)\end{bmatrix}$$

$$=[(C_i, C_j)]_{n\times n}.$$

Therefore $A^\theta A=I$ if and only if

(i) $(C_i, C_i)=1,\ i=1, 2,\ldots, n$

(ii) $(C_i, C_j)=0, i=1,\ldots, n, j=1,\ldots, n$ and $i\neq j$

i.e., if and only if the vectors $C_1, C_2,\ldots,C_n$ form an orthonormal set.

In order to show that row vectors of A form an orthonormal set we write

$$A=\begin{bmatrix}R_1\\ R_2\\ \vdots\\ R_n\end{bmatrix}$$ where $R_1, R_2,\ldots,R_n$ are row vectors of A.

Then we use the condition

$$AA^\theta=I.$$

Theorem 6. *A square matrix is orthogonal if and only if its columns (rows) form an orthonormal set of vectors.*

For proof proceed as in Theorem 5.

Theorem 7. *The columns of a square matrix form an orthonormal set if and only if the rows form an orthonormal set.*

It is just a corollary to theorems 5 and 6.

Theorem 8. *If the order of the rows (or columns) of a unitary (real orthogonal) matrix is changed, then the resulting matrix is also unitary (real orthogonal).*

Theorem 9. *Let X_1 be any unit n-vector. Then there exists a unitary matrix U having X_1 as its first column.*

Proof. Let X_1 be any unit n-vector. Then $X_1\neq O$. Therefore $\{X_1\}$ is a linearly independent subset of V_n. So $\{X_1\}$ can be extended to form a basis of V_n. Let $\{X_1, X_2,\ldots,X_n\}$ be a basis of V_n having X_1 as a member. By Gram-Schmidt orthogonalization process we can find an orthogonal basis $\{Y_1,\ldots,Y_n\}$ of V_n such that $Y_1=X_1$. Let

$$Z_i=Y_i/\|Y_i\|,\ i=1,\ldots,n.$$

Then $\{Z_1, Z_2,\ldots,Z_n\}$ is an orthonormal basis of V_n with $Z_1=Y_1/\|Y_1\|=X_1/\|X_1\|=X_1$, since X_1 is a unit vector $\Rightarrow \|X_1\|=1$.

Now let U be the matrix whose columns are $Z_1, Z_2,\ldots,Z_n$ *i.e.*

$$U=[Z_1, Z_2,\ldots, Z_n].$$

Since $Z_1, Z_2,\ldots, Z_n$ form an orthonormal set of vectors, therefore U is a unitary matrix. The first column of U is $Z_1=Y_1=X_1$.

Theorem 10. *Let X_1 be any real n-vector. Then there exists an orthonormal matrix P having X_1 as its first column.*

The proof of theorem 9 will hold good in this case.

Solved Examples

Ex. 1. *Show that the matrix* $A=\begin{bmatrix}a+ic & -b+id\\ b+id & a-ic\end{bmatrix}$ *is unitary if and only if* $a^2+b^2+c^2+d^2=1$.

Solution. We have

$$AA^\theta=\begin{bmatrix}a+ic & -b+id\\ b+id & a-ic\end{bmatrix}\begin{bmatrix}a-ic & b-id\\ -b-id & a+ic\end{bmatrix}$$

$$=\begin{bmatrix}a^2+b^2+c^2+d^2 & 0\\ 0 & a^2+b^2+c^2+d^2\end{bmatrix}.$$

$\therefore$ $AA^\theta=I$ if and only if $a^2+b^2+c^2+d^2=1$

i.e. A is unitary if and only if $a^2+b^2+c^2+d^2=1$.

Ex. 2. *Show that the matrix*

$$A=\begin{bmatrix}0 & 2m & n\\ l & m & -n\\ l & -m & n\end{bmatrix},\ l=\frac{1}{\sqrt2},\ m=\frac{1}{\sqrt6},\ n=\frac{1}{\sqrt3}$$

is orthogonal.

Solution. Let C_1, C_2, C_3 be the column vectors of A. Then

$$C_1=\begin{bmatrix}0\\ l\\ l\end{bmatrix},\ C_2=\begin{bmatrix}2m\\ m\\ -m\end{bmatrix},\ C_3=\begin{bmatrix}n\\ -n\\ n\end{bmatrix}.$$

We have $(C_1, C_1)=0+l^2+l^2=2l^2=2.\frac12=1,$

$(C_2, C_2)=4m^2+m^2+m^2=6m^2=6.\frac16=1,$

$(C_3, C_3)=n^2+n^2+n^2=3n^2=3.\frac13=1.$

Also $(C_1, C_2)=C_1^TC_2=0.2m+l.m+l.(-m)=0$

$(C_2, C_3)=2m.n+m.(-n)+(-m).n=0,$

$(C_3, C_1)=0.n+l.(-n)+l.n=0.$

Thus the columns of A form an orthonormal set of vectors. Therefore A is an orthogonal matrix.

Ex. 3. A *is a unitary matrix such that the elements of its first column other than the first element, are all zero; show that every element of the first row, other than the first, is also zero.*

Solution. Let a_{11} be the first element of the first column of **A**. Since **A** is unitary, therefore the first column vector of **A** is a unit vector. But each element of the first column of **A** other than the first element is zero. Therefore we must have

$$a_{11}\bar{a}_{11}=1 \Rightarrow |a_{11}|^2=1.$$

Let $[a_{11}, a_{12}, ..., a_{1n}]$ be the first row vector of **A**. Then it is also a unit vector. Therefore

$$|a_{11}|^2+|a_{12}|^2+...+|a_{1n}|^2=1$$
$$\Rightarrow |a_{12}|^2+...+|a_{1n}|^2=0 \qquad [\because \; |a_{11}|^2=1]$$
$$\Rightarrow a_{12}=0, ..., a_{1n}=0$$

$\Rightarrow$ each element of the first row of **A** other than the first element is zero.

Ex. 4. *If* $X_1, X_2, ..., X_r$ *is any orthonormal set,* $r < n$, *of real column n-vectors, show that there exists an orthogonal matrix*

$$[X_1, X_2, ..., X_r, X_{r+1}, ..., X_n].$$

Solution. Since $\{X_1, X_2, ..., X_r\}$ is an orthonormal set, therefore it is a linearly independent subset of V_n. So it can be extended to form a basis of V_n. Let $\{X_1, ..., X_r, Y_{r+1}, ..., Y_n\}$ be a basis of V_n. Applying Gram-Schmidt process to this basis, we shall get an orthonormal basis of V_n. In this process the vectors $X_1, ..., X_r$ will remain unchanged. Thus we shall get an orthonormal basis $\{X_1, ..., X_r, X_{r+1}, ..., X_n\}$ of V_n.

Consequently the matrix $P=[X_1, ..., X_r, X_{r+1}, ..., X_n]$ will be orthogonal.

Ex. 5. *Show that if* **A** *is Hermitian and* **P** *orthogonal, then* $P^{-1}AP$ *is symmetric.*

Solution. It is given that **A** is symmetric. Therefore $A^T=A$. Also **P** is orthogonal. Therefore $P^{-1}=P^T$. We have

$$(P^{-1}AP)^T=(P^TAP)^T=P^T A^T (P^T)^T$$
$$=P^T AP=P^{-1}AP.$$

Therefore $P^{-1}AP$ is symmetric.

Ex. 6. *Show that if* **A** *is Hermitian and* **P** *unitary, then* $P^{-1}AP$ *is Hermitian.*

Solution. Since **P** is unitary, therefore $P^{-1}=P^\theta$. Also **A** is Hermitian implies $A^\theta=A$. We have

$$(P^{-1}AP)^\theta=(P^\theta AP)^\theta=P^\theta A^\theta (P^\theta)^\theta=P^\theta AP=P^{-1}AP.$$

Therefore $P^{-1}AP$ is Hermitian.

Ex. 7. *Prove that the eigenvalues of* A^θ *are the conjugates of the eigenvalues of* **A**. *If* k_1 *and* k_2 *are distinct eigenvalues of* **A**, *prove that any eigenvector of* **A** *corresponding to* k_1 *is orthogonal to any eigenvector of* A^θ *corresponding to* $\bar{k}_2$.

Solution. The first part has already been proved.

Let X_1 be an eigenvector of **A** corresponding to its eigenvalue k_1 and let X_2 be an eigenvector of A^θ corresponding to its eigenvalue $\bar{k}_2$. Then

$$AX_1=k_1X_1 \qquad ...(1)$$

and
$$A^\theta X_2=\bar{k}_2X_2. \qquad ...(2)$$

Taking conjugate transpose of both sides of (1), we get

$$X_1^\theta A^\theta=\bar{k}_1X_1^\theta. \qquad ...(3)$$

Post-multiplying both sides of (3) by X_2, we get

$$X_1^\theta A^\theta X_2=\bar{k}_1X_1^\theta X_2$$

or $\quad X_1^\theta \bar{k}_2X_2=\bar{k}_1X_1^\theta X_2 \qquad$ [from (2)]

or $\quad \bar{k}_2X_1^\theta X_2=\bar{k}_1X_1^\theta X_2$

or $\quad (\bar{k}_2-\bar{k}_1)X_1^\theta X_2=O$

or $\quad X_1^\theta X_2=O$, since $k_1\neq k_2 \Rightarrow \bar{k}_1\neq\bar{k}_2 \Rightarrow \bar{k}_2-\bar{k}_1\neq 0$

$\therefore$ X_1 and X_2 are orthogonal.

10
Similarity of Matrices

§ 1. Similarity of Matrices. Definition. *Let* A *and* B *be square matrices of order n. Then* B *is said to be similar to* A *if there exists a non-singular matrix* P *such that*

$$B=P^{-1}AP. \qquad \textbf{(Nagarjuna 1980)}$$

Theorem 1. *Similarity of matrices is an equivalence relation.* **(Madras 1980)**

Proof. If A and B are two $n \times n$ matrices, then B is said to be similar to A if there exists an $n \times n$ invertible matrix P such that $B=P^{-1}AP$.

Reflexivity. We are to prove that every matrix is similar to itself. Let A be any matrix of order n. We can write $A=I^{-1}AI$, where I is unit matrix of order n. Therefore A is similar to A.

Symmetry. Let A be similar to B. Then to prove that B is also similar to A. Since A is similar to B, therefore there exists an $n \times n$ non-singular matrix P such that

$$\begin{aligned} & A=P^{-1}BP \\ \Rightarrow\ & PAP^{-1}=P\,(P^{-1}BP)\,P^{-1} \\ \Rightarrow\ & PAP^{-1}=B \\ \Rightarrow\ & B=PAP^{-1} \\ \Rightarrow\ & B=(P^{-1})^{-1}AP^{-1} \end{aligned}$$

[$\because$ P is invertible means P^{-1} is invertible and $(P^{-1})^{-1}=P$]

$\Rightarrow$ B is similar to A.

Transitivity. Let A be similar to B and B be similar to C. Then to prove that A is similar to C. Since A is similar to B and B is similar to C, therefore

$$A=P^{-1}BP \text{ and } B=Q^{-1}CQ,$$

where P and Q are invertible $n \times n$ matrices. We have

$$\begin{aligned} A&=P^{-1}BP=P^{-1}\,(Q^{-1}CQ)\,P \\ &=(P^{-1}Q^{-1})\,C\,(QP)=(QP)^{-1}\,C\,(QP) \end{aligned}$$

[$\because$ P and Q are invertible means QP is invertible and $(QP)^{-1}=P^{-1}Q^{-1}$]

$\therefore$ A is similar to C.

Hence similarity of matrices is an equivalence relation in the set of all $n \times n$ matrices over a given field.

Theorem 2. *Similar matrices have the same determinant.*

Proof. Suppose A and B are similar matrices. Then there exists an invertible matrix P such that $B=P^{-1}AP$.

$$\begin{aligned} \therefore\ \det B&=\det(P^{-1}AP)=(\det P^{-1})(\det A)(\det P) \\ &=(\det P^{-1})(\det P)(\det A)=(\det P^{-1}P)(\det A) \\ &=(\det I)(\det A)=1\,(\det A)=\det A. \end{aligned}$$

Theorem 3. *Similar matrices have the same characteristic polynomial and hence the same eigenvalues. If* X *is an eigenvector of* A *corresponding to the eigenvalue* λ, *then* $P^{-1}X$ *is an eigenvector of* B *corresponding to the eigenvalue* λ *where*

$$B=P^{-1}AP. \qquad \textbf{[Madurai 1985]}$$

Proof. Suppose A and B are similar matrices. Then there exists an invertible matrix P such that $B=P^{-1}AP$. We have

$$\begin{aligned} B-xI&=P^{-1}AP-xI \\ &=P^{-1}AP-P^{-1}(xI)P \qquad [\because P^{-1}(xI)P=xP^{-1}P=xI] \\ &=P^{-1}(A-xI)P. \end{aligned}$$

$$\begin{aligned} \therefore\ \det(B-xI)&=\det P^{-1}\det(A-xI)\det P \\ &=\det P^{-1}.\det P.\det(A-xI)=\det(P^{-1}P).\det(A-xI) \\ &=\det I.\det(A-xI)=1.\det(A-xI)=\det(A-xI). \end{aligned}$$

Thus the matrices A and B have the same characteristic polynomial and so they have the same eigenvalues.

If λ is an eigenvalue of A and X is a corresponding eigenvector, then $AX=\lambda X$, and hence

$$B\,(P^{-1}X)=(P^{-1}AP)\,P^{-1}X=P^{-1}AX=P^{-1}(\lambda X)=\lambda\,(P^{-1}X).$$

$\therefore$ $P^{-1}X$ is an eigenvector of B corresponding to its eigen value λ.

Corollary. *If* A *is similar to a diagonal matrix* D, *the diagonal elements of* D *are the eigenvalues of* A.

Proof. We know that similar matrices have same eigenvalues. Therefore A and D have the same eigenvalues. But the eigenvalues of the diagonal matrix D are its diagonal elements. Hence the eigenvalues of A are the diagonal elements of D.

§ 2. Diagonalizable matrix. Definition. *A matrix* A *is said to be diagonalizable if it is similar to a diagonal matrix.*

Thus a matrix A is diagonalizable if there exists an invertible matrix P such that $P^{-1}AP=D$ where D is a diagonal matrix. Also the matrix P is then said to diagonalize A or transform A to diagonal form.

Theorem 1. *An $n \times n$ matrix is diagonalizable if and only if it possesses n linearly independent eigenvectors.*

(Banaras 1988; I.C.S. 89)

Proof. Suppose A is diagonalizable. Then A is similar to a diagonal matrix $D = \text{dia.}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Therefore there exists an invertible matrix $P = [X_1, X_2, \ldots, X_n]$ such that

$$P^{-1} AP = D$$

i.e., $\quad AP = PD$

i.e., $\quad A[X_1, X_2, \ldots, X_n] = [X_1, X_2, \ldots, X_n] \text{ dia. } [\lambda_1, \lambda_2, \ldots, \lambda_n]$

i.e., $\quad [AX_1, AX_2, \ldots, AX_n] = [\lambda_1 X_1, \lambda_2 X_2, \ldots, \lambda_n X_n]$

i.e., $\quad AX_1 = \lambda_1 X_1, AX_2 = \lambda_2 X_2, \ldots, AX_n = \lambda_n X_n.$

Therefore $X_1, X_2, \ldots, X_n$ are eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ respectively. Since the matrix P is non-singular, therefore its column vectors $X_1, X_2, \ldots, X_n$ are linearly independent. Hence A possesses n linearly independent eigenvectors.

Conversely, suppose that A possesses n linearly independent eigenvectors $X_1, X_2, \ldots, X_n$ and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the corresponding eigenvalues. Then $AX_1 = \lambda_1 X_1, AX_2 = \lambda_2 X_2, \ldots, AX_n = \lambda_n X_n$.

Let $P = [X_1, X_2, \ldots, X_n]$ and $D = \text{dia.}[\lambda_1, \lambda_2, \ldots, \lambda_n]$.

Then $AP = A[X_1, X_2, \ldots, X_n] = [AX_1, AX_2, \ldots, AX_n]$

$$= [\lambda_1 X_1, \lambda_2 X_2, \ldots, \lambda_n X_n]$$

$$= [X_1, X_2, \ldots, X_n] \text{ dia. } [\lambda_1, \lambda_2, \ldots, \lambda_n] = PD.$$

Since the column vectors $X_1, X_2, \ldots, X_n$ of the matrix P are linearly independent, therefore X is invertible and X^{-1} exists. Therefore $\quad AP = PD \Rightarrow P^{-1} AP = P^{-1} PD$

$$\Rightarrow P^{-1} AP = D$$

$\Rightarrow$ A is similar to a diagonal matrix D

$\Rightarrow$ A is diagonalizable.

Remark. In the proof of the above theorem we have shown that if A is diagonalizable and P diagonalizes A, then

$$P^{-1} AP = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} = D$$

if and only if the j^{th} column of P is an eigenvector of A corresponding to the eigenvalue λ_j of A, $(j = 1, 2, \ldots, n)$. The diagonal elements of D are the eigenvalues of A and they occur in the same order as is the order of their corresponding eigenvectors in the column vectors of P.

Theorem 2. *If the eigenvalues of an $n \times n$ matrix are all distinct then it is always similar to a diagonal matrix.*

(Banaras 1968)

Proof. Let A be a square matrix of order n and suppose it has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. We know that eigenvectors of a matrix corresponding to distinct eigenvalues are linearly independent. Therefore A has n linearly independent eigenvectors and so it is similar to a diagonal matrix

$$D = \text{dia.}[\lambda_1, \lambda_2, \ldots, \lambda_n].$$

Corollary. *Two $n \times n$ matrices with the same set of n distinct eigenvalues are similar.*

Proof. Suppose A and B are two $n \times n$ matrices with the same set of n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let $D = \text{dia.}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Then both A and B are similar to D. Now A is similar to D and D is similar to B implies that A is similar to B.

Note that the relation of similarity is transitive.

Theorem 3. *The necessary and sufficient condition for a square matrix to be similar to a diagonal matrix is that the geometric multiplicity of each of its eigenvalues coincides with the algebraic multiplicity.*

Proof. The condition is necessary. Suppose A is similar to a diagonal matrix $D = \text{diag.}[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Then $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A and there exists a non-singular matrix P such that

$$P^{-1} AP = D.$$

Let α be an eigenvalue of A of algebraic multiplicity k. Then exactly k among $\lambda_1, \lambda_2, \ldots, \lambda_n$ are equal to α.

Let $m = \text{rank}(A - \alpha I)$. Then the system of equations

$$(A - \alpha I) X = O$$

have $n - m$ linearly independent solutions and so $n - m$ will be the geometric multiplicity of α. We are to prove that $k = n - m$. We know that the rank of a matrix does not change on multiplication by a non-singular matrix. Therefore

$$\text{rank}(A - \alpha I) = \text{rank}[P^{-1}(A - \alpha I) P] = \text{rank}[P^{-1} AP - \alpha I]$$

$$= \text{rank}[D - \alpha I] = \text{rank dia.}[\lambda_1 - \alpha, \lambda_2 - \alpha, \ldots, \lambda_n - \alpha]$$

$= n - k$, since exactly k elements of

dia. $[\lambda_1 - \alpha, \lambda_2 - \alpha, \ldots, \lambda_n - \alpha]$ are equal to zero.

Thus $\text{rank}(A - \alpha I) = m = n - k$. Therefore $k = n - m$.

Thus there are exactly k linearly independent eigenvectors corresponding to the eigenvalue α.

The condition is sufficient. Suppose that the geometric multiplicity of each eigenvalue of **A** is equal to its algebraic multiplicity. Let $\lambda_1, \ldots, \lambda_p$ be the set of p distinct eigenvalues of **A** with respective multiplicities $r_1, \ldots, r_p$. We have

$$r_1 + \ldots + r_p = n.$$

To prove that **A** is diagonalizable.

Let

$$\left.\begin{matrix} \mathbf{C}_{11}, \mathbf{C}_{12}, \ldots, \mathbf{C}_{1r_1} \\ \ldots \quad \ldots \quad \ldots \quad \ldots \\ \ldots \\ \mathbf{C}_{p1}, \mathbf{C}_{p2}, \ldots, \mathbf{C}_{pr_p} \end{matrix}\right\} \qquad \ldots(1)$$

be linearly independent sets of eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_p$ respectively. We claim that the n vectors given in (1) are linearly independent. Let

$$(a_{11} \mathbf{C}_{11} + a_{12} \mathbf{C}_{12} + \ldots + a_{1r_1} \mathbf{C}_{1r_1}) + \ldots + (a_{p1} \mathbf{C}_{p1} + \ldots + a_{pr_p} \mathbf{C}_{pr_p}) = \mathbf{O}. \qquad \ldots(2)$$

The relation (3) may be written as

$$\mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_p = \mathbf{O}, \qquad \ldots(3)$$

where $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p$ denote the vectors written within brackets in (2) *i.e.*, $\mathbf{X}_1 = a_{11} \mathbf{C}_{11} + \ldots + a_{1r_1} \mathbf{C}_{1r_1}$, and so on.

Now $\mathbf{X}_1$ is a linear combination of eigenvectors of **A** corresponding to the eigenvalue λ_1. Therefore if $\mathbf{X}_1 \neq \mathbf{O}$, then $\mathbf{X}_1$ is also an eigenvector of **A** corresponding to the eigenvalue λ_1.

Similarly we can speak for $\mathbf{X}_2, \ldots, \mathbf{X}_p$.

In case some one of $\mathbf{X}_1, \ldots, \mathbf{X}_p$ is not zero, then the relation (3) implies that a system of eigenvectors of **A** corresponding to distinct eigenvalues of **A** is linearly dependent. But this is not possible. Hence each of the vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_p$ must be zero.

Since $\mathbf{C}_{11}, \mathbf{C}_{12}, \ldots, \mathbf{C}_{1r_1}$ is a set of linearly independent vectors, therefore $\mathbf{O} = \mathbf{X}_1 = a_{11} \mathbf{C}_{11} + \ldots + a_{1r_1} \mathbf{C}_{1r_1}$ implies that

$$a_{11} = 0, \ldots, a_{1r_1} = 0.$$

Similarly we can show that each of the scalars in relation (2) is zero. Therefore the n vectors given in (1) are linearly independent. Thus **A** has n linearly independent eigenvectors. So it is similar to a diagonal matrix.

Solved Examples

Ex. 1. *Show that the rank of every matrix similar to* **A** *is the same as that of* **A**.

Solution. Let **B** be a matrix similar to **A**. Then there exists a non-singular matrix **P** such that $\mathbf{B} = \mathbf{P}^{-1} \mathbf{AP}$. We know that the rank of a matrix does not change on multiplication by a non-singular matrix. Therefore

$$\text{rank } (\mathbf{P}^{-1} \mathbf{AP}) = \text{rank } \mathbf{A} \Rightarrow \text{rank } \mathbf{B} = \text{rank } \mathbf{A}.$$

Ex. 2. *Let* **A** *and* **B** *be n-rowed square matrices and let* **A** *be non-singular. Show that the matrices* $\mathbf{A}^{-1}\mathbf{B}$ *and* $\mathbf{BA}^{-1}$ *have the same eigenvalues.*

Solution. We have $\mathbf{A}^{-1} (\mathbf{B} \mathbf{A}^{-1}) \mathbf{A} = \mathbf{A}^{-1} \mathbf{B}$.

Therefore $\mathbf{BA}^{-1}$ is similar to $\mathbf{A}^{-1}\mathbf{B}$. But similar matrices have the same eigenvalues. Therefore $\mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{BA}^{-1}$ have the same eigenvalues.

Ex. 3. *If* **U** *be a unitary matrix such that* $\mathbf{U}^{\theta}\mathbf{AU} = \textit{diag } [\lambda_1, \ldots, \lambda_n]$, *show that* $\lambda_1, \ldots, \lambda_n$ *are the eigenvalues of* **A**.

Solution. Let $\text{diag } [\lambda_1, \ldots, \lambda_n] = \mathbf{D}$. Since **U** is unitary, therefore $\mathbf{U}^{\theta} = \mathbf{U}^{-1}$. So

$$\mathbf{U}^{\theta}\mathbf{AU} = \mathbf{D} \Rightarrow \mathbf{U}^{-1} \mathbf{AU} = \mathbf{D}.$$

Thus **A** is similar to the diagonal matrix **D**. But similar matrices have the same eigenvalues and eigenvalues of **D** are its diagonal elements. Therefore $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of **A**.

Ex. 4. *If* **A** *and* **B** *are non-singular matrices of order n, show that the matrices* **AB** *and* **BA** *are similar.*

Solution. Since **A** is non-singular, therefore $\mathbf{A}^{-1}$ exists. We have $\mathbf{A}^{-1} (\mathbf{AB}) \mathbf{A} = \mathbf{BA}$.

Therefore **AB** and **BA** are similar matrices.

Ex. 5. **A** *and* **B** *are two* $n \times n$ *matrices with the same set of n distinct eigenvalues. Show that there exist two matrices* **P** *and* **Q** *(one of them non-singular) such that*

$$\mathbf{A} = \mathbf{PQ}, \ \mathbf{B} = \mathbf{QP}.$$

Solution. Since **A** and **B** have the same set of n distinct eigenvalues, therefore they are similar. So there exists a non-singular matrix **P** such that

$$\mathbf{P}^{-1} \mathbf{AP} = \mathbf{B}. \qquad \ldots(1)$$

Let $\mathbf{P}^{-1}\mathbf{A} = \mathbf{Q}$. Then from (1), we get $\mathbf{QP} = \mathbf{B}$. Also

$$\mathbf{P}^{-1} \mathbf{A} = \mathbf{Q} \Rightarrow \mathbf{A} = \mathbf{PQ}.$$

Ex. 6. *Prove that if* **A** *is similar to a diagonal matrix, then* $\mathbf{A}^T$ *is similar to* **A**.

Solution. Suppose A is similar to a diagonal matrix D. Then there exists a non-singular matrix P such that

$$P^{-1}AP=D$$
$$\Rightarrow A=PDP^{-1}$$
$$\Rightarrow A^T=(PDP^{-1})^T=(P^{-1})^T D^T P^T$$
$$\Rightarrow A^T=(P^T)^{-1} D P^T \quad [\because D \text{ is diagonal} \Rightarrow D^T=D]$$
$$\Rightarrow A^T \text{ is similar to } D$$
$$\Rightarrow D \text{ is similar to } A^T.$$

Finally A is similar to D and D is similar to A^T implies that A is similar to A^T.

Nilpotent Matrix. Definition.

A non-zero matrix A is said to be nilpotent, if for some positive integer r, $A^r=O$.

Ex. 7. *Show that a non-zero matrix is nilpotent if and only if all its eigenvalues are equal to zero.*

Solution. Suppose $A \neq O$ and A is nilpotent. Then

$A^r=O$, for some positive integer r

$\Rightarrow$ the polynomial λ^r annihilates A

$\Rightarrow$ the minimal polynomial $m(\lambda)$ of A divides λ^r

$\Rightarrow$ $m(\lambda)$ is of the type λ^s, where s is some positive integer

$\Rightarrow$ 0 is the only root of $m(\lambda)$

$\Rightarrow$ 0 is the only eigenvalue of A

$\Rightarrow$ all eigenvalues of A are zero.

Conversely, each eigenvalue of $A=0$

$\Rightarrow$ characteristic equation of A is $\lambda^n=0$

$\Rightarrow$ $A^n=O$, since A satisfies its characteristic equation

$\Rightarrow$ A is nilpotent.

Ex. 8. *Prove that a non-zero nilpotent matrix cannot be similar to a diogonal matrix.*

Solution. Suppose A is a non-zero nilpotent matrix similar to a diagonal matrix D. Since A is a non-zero nilpotent matrix, therefore each eigenvalue of A is zero. But A and D have the same eigenvalues and the eigenvalues of D are its diagonal elements. Therefore D must be a zero matrix. Now A is similar to D implies that there exists a non-singular matrix P such that

$$P^{-1}AP=D$$
$$\Rightarrow P^{-1}AP=O \quad [\because D=O]$$
$$\Rightarrow P(P^{-1}AP)P^{-1}=POP^{-1}$$
$$\Rightarrow A=O.$$

But this contradicts the hypothesis that A is a non-zero matrix. So A cannot be similar to a diagonal matrix.

Ex. 9. *Show that the matrix*

$$A=\begin{bmatrix} -9 & 4 & 4 \\ -8 & 3 & 4 \\ -16 & 8 & 7 \end{bmatrix}$$

is diagonalizable. Also find the diagonal form and a diagonalizing matrix **P**.

Solution. The characteristic equation of A is

$$\begin{vmatrix} -9-\lambda & 4 & 4 \\ -8 & 3-\lambda & 4 \\ -16 & 8 & 7-\lambda \end{vmatrix}=0$$

or $$\begin{vmatrix} -1-\lambda & 4 & 4 \\ -1-\lambda & 3-\lambda & 4 \\ -1-\lambda & 8 & 7-\lambda \end{vmatrix}=0, \text{ applying } C_1+C_2+C_3$$

or $$-(1+\lambda)\begin{vmatrix} 1 & 4 & 4 \\ 1 & 3-\lambda & 4 \\ 1 & 8 & 7-\lambda \end{vmatrix}=0$$

or $$(1+\lambda)\begin{vmatrix} 1 & 4 & 4 \\ 0 & -1-\lambda & 0 \\ 0 & 4 & 3-\lambda \end{vmatrix}=0, \text{ applying } R_2-R_1,\ R_3-R_1$$

or $$(1+\lambda)(1+\lambda)(3-\lambda)=0.$$

The roots of this equation are $-1, -1, 3$.

$\therefore$ the eigenvalues of the matrix A are $-1, -1, 3$.

The eigenvectors X of A corresponding to the eigenvalue -1 are given by the equation $(A-(-1)I)X=O$ or $(A+I)X=O$

or $$\begin{bmatrix} -8 & 4 & 4 \\ -8 & 4 & 4 \\ -16 & 8 & 8 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

These equations are equivalent to the equations

$$\begin{bmatrix} -8 & 4 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ applying } R_2 \to R_2-R_1 \text{ and } R_3 \to R_3-2R_1.$$

The matrix of coefficients of these equations has rank 1. Therefore these equations have two linearly independent solutions. We see that these equations reduce to the single equation

$$-2x_1+x_2+x_3=0.$$

Obviously $X_1=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $X_2=\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ are two linearly independent solutions of this equation. Therefore X_1 and X_2 are two linearly independent eigenvectors of **A** corresponding to the eigenvalue -1. Thus the geometric multiplicity of the eigenvalue -1 is equal to its algebraic multiplicity.

Now the eigenvectors of **A** corresponding to the eigenvalue 3 are given by $(A-3I)\,X=O$

i.e.
$$\begin{bmatrix}-12 & 4 & 4\\ -8 & 0 & 4\\ -16 & 8 & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

or
$$\begin{bmatrix}-12 & 4 & 4\\ 4 & -4 & 0\\ -4 & 4 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}, \text{ applying } R_2\to R_2-R_1,\ R_3\to R_3-R_1$$

or
$$\begin{bmatrix}-12 & 4 & 4\\ 4 & -4 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}, \text{ applying } R_3\to R_3+R_2.$$

The matrix of coefficients of these equations has rank 2. Therefore these equations have $3-2=1$ linearly independent solution. These equations can be written as $-12x_1+4x_2+4x_3=0$, $4x_1-4x_2=0$. From these, we get $x_1=x_2=1$, say. Then $x_3=2$.

Therefore $X_3=\begin{bmatrix}1\\ 1\\ 2\end{bmatrix}$ is an eigenvector of **A** corresponding to the eigenvalue 3. The geometric multiplicity of the eigenvalue 3 is 1 and its algebraic multiplicity is also 1.

Since the geometric multiplicity of each eigenvalue of **A** is equal to its algebraic multiplicity, therefore **A** is similar to a diagonal matrix.

Let
$$P=\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 1\\ 1 & -1 & 2\end{bmatrix}$$

The columns of **P** are linearly independent eigenvectors of **A** corresponding to the eigenvalues $-1, -1, 3$ respectively. The matrix **P** will transform **A** to diagonal form **D** which is given by the relation

$$P^{-1}AP=\begin{bmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 3\end{bmatrix}=D.$$

Ex. 10. *Show that the matrix*

$$A=\begin{bmatrix}8 & -8 & -2\\ 4 & -3 & -2\\ 3 & -4 & 1\end{bmatrix}$$

is diagonalizable. Also find the transforming matrix and diagonal matrix.

Solution. The characteristic equation of **A** is

$$\begin{vmatrix}8-\lambda & -8 & -2\\ 4 & -3-\lambda & -2\\ 3 & -4 & 1-\lambda\end{vmatrix}=0$$

or
$$\begin{vmatrix}1-\lambda & -1+\lambda & -1+\lambda\\ 4 & -3-\lambda & -2\\ 3 & -4 & 1-\lambda\end{vmatrix}=0, \text{ applying } R_1-(R_2+R_3)$$

or
$$(1-\lambda)\begin{vmatrix}1 & -1 & -1\\ 4 & -3-\lambda & -2\\ 3 & -4 & 1-\lambda\end{vmatrix}=0,$$

or
$$(1-\lambda)\begin{vmatrix}1 & 0 & 0\\ 4 & 1-\lambda & 2\\ 3 & -1 & 4-\lambda\end{vmatrix}=0, \text{ applying } C_2+C_1,\ C_3+C_1$$

or $(1-\lambda)\,[(1-\lambda)\,(4-\lambda)+2]=0$

or $(1-\lambda)\,(\lambda^2-5\lambda+6)=0$ or $(1-\lambda)\,(\lambda-2)\,(\lambda-3)=0$

The roots of this equation are 1, 2, 3.

Since the eigenvalues of the matrix **A** are all distinct, therefore **A** is similar to a diagonal matrix. Since the algebraic multiplicity of each eigenvalue of **A** is 1, therefore there will be one and only one linearly independent eigenvector of **A** corresponding to each eigenvalue of **A**.

The eigenvectors **X** of **A** corresponding to the eigenvalue 1 are given by the equation $(A-1I)\,X=O$ or $(A-I)\,X=O$

or
$$\begin{bmatrix}7 & -8 & -2\\ 4 & -4 & -2\\ 3 & -4 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

or
$$\begin{bmatrix}7 & -8 & -2\\ -3 & 4 & 0\\ 3 & -4 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}, \text{ by } R_2\to R_2-R_1$$

or
$$\begin{bmatrix}7 & -8 & -2\\ -3 & 4 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}, \text{ by } R_3\to R_3+R_2.$$

The matrix of coefficients of these equations has rank 2. Therefore these equations have only one linearly independent solution as it should have been because the algebraic multiplicity of the eigenvalue 1 is 1. Note that the geometric multiplicity cannot exceed the algebraic multiplicity. The above equations can be written as $7x_1-8x_2-2x_3=0$, $-3x_1+4x_2=0$. From the last equation, we get $x_1=4$, $x_2=3$. Then the first gives $x_3=2$. Therefore

$\mathbf{X}_1=\begin{bmatrix}4\\3\\2\end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 1.

The eigenvectors **X** of A corresponding to the eigenvalue 2 are given by the equation $(\mathbf{A}-2\mathbf{I})\,\mathbf{X}=\mathbf{O}$

$$\text{or}\quad \begin{bmatrix}6 & -8 & -2\\ 4 & -5 & -2\\ 3 & -4 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix}6 & -8 & -2\\ -2 & 3 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix},\ \text{applying } R_2\to R_2-R_1,\ R_3\to R_3-\tfrac{1}{2}R_1.$$

These equations can be written as $6x_1-8x_2-2x_3=0$, $-2x_1+3x_2=0$. From these, we get $x_1=3$, $x_2=2$, $x_3=1$.

Therefore $\mathbf{X}_2=\begin{bmatrix}3\\2\\1\end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 2.

The eigenvectors **X** of A corresponding to the eigenvalue 3 are given by the equation $(\mathbf{A}-3\mathbf{I})\,\mathbf{X}=\mathbf{O}$

$$\text{or}\quad \begin{bmatrix}5 & -8 & -2\\ 4 & -6 & -2\\ 3 & -4 & -2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix}5 & -8 & -2\\ -1 & 2 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix},\ \text{applying } R_2\to R_2-R_1,\ R_3\to R_3+R_1-2R_2.$$

These equations can be written as

$$5x_1-8x_2-2x_3=0,\ -x_1+2x_2=0.$$

From these, we get $x_1=2$, $x_2=1$, $x_3=1$.

$\therefore\ \mathbf{X}_3=\begin{bmatrix}2\\1\\1\end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 3.

Let $\mathbf{P}=[\mathbf{X}_1\ \mathbf{X}_2\ \mathbf{X}_3]=\begin{bmatrix}4 & 3 & 2\\ 3 & 2 & 1\\ 2 & 1 & 1\end{bmatrix}$.

The columns of **P** are linearly independent eigenvectors of **A** corresponding to the eigenvalues 1, 2, 3 respectively. The matrix **P** will transform **A** to diagonal form **D** which is given by the relation

$$\mathbf{P}^{-1}\,\mathbf{A}\,\mathbf{P}=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}=\mathbf{D}.$$

Ex. 11. *Show that the matrix*

$$\mathbf{A}=\begin{bmatrix}1 & -6 & -4\\ 0 & 4 & 2\\ 0 & -6 & -3\end{bmatrix}$$

is similar to a diagonal matrix. Also find the transforming matrix and diagonal matrix.

Solution. The characteristic equation of A is

$$\begin{vmatrix}1-\lambda & -6 & -4\\ 0 & 4-\lambda & 2\\ 0 & -6 & -3-\lambda\end{vmatrix}=0$$

or $(1-\lambda)\,[(4-\lambda)(-3-\lambda)+12]=0$

or $(1-\lambda)\,[\lambda^2-\lambda]=0$ or $\lambda\,(1-\lambda)\,(\lambda-1)=0$.

$\therefore$ the eigenvalues of A are 0, 1, 1.

The eigenvectors **X** of A corresponding to the eigenvalue 1 are given by $(\mathbf{A}-\mathbf{I})\,\mathbf{X}=\mathbf{O}$

$$\text{or}\quad \begin{bmatrix}0 & -6 & -4\\ 0 & 3 & 2\\ 0 & -6 & -4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix}0 & -6 & -4\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix},\ \text{by } R_2\to R_2+\tfrac{1}{2}R_1,\ R_3\to R_3-R_1.$$

The coefficient matrix of these equations is of rank 1. So these equations have 2 linearly independent solutions.

These equations can be written as $0x_1-6x_2-4x_3=0$. We see that

$$\mathbf{X}_1=\begin{bmatrix}1\\-2\\3\end{bmatrix},\ \mathbf{X}_2=\begin{bmatrix}2\\-2\\3\end{bmatrix}$$

are two linearly independent solutions of these equations.

The eigenvectors of A corresponding to the eigenvalue 0 are given by $(\mathbf{A}-0\mathbf{I})\,\mathbf{X}=\mathbf{O}$

$$\text{or}\quad \begin{bmatrix}1 & -6 & -4\\ 0 & 4 & 2\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix}1 & -6 & -4\\0 & 4 & 2\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}, \text{ by } R_3\to R_3+\tfrac{3}{2}R_2.$$

These equations can be written as $x_1-6x_2-4x_3=0$, $4x_2+2x_3=0$.

From these, we get $x_2=1$, $x_3=-2$, $x_1=2$.

$\therefore\quad X_3=\begin{bmatrix}2\\1\\-2\end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 0.

Since the geometric multiplicity of each eigenvalue of A is equal to its algebraic multiplicity, therefore A is similar to a diagonal matrix.

Let $P=\begin{bmatrix}1 & 2 & 2\\-2 & -2 & 1\\3 & 3 & -2\end{bmatrix}$. The columns of P are linearly independent eigenvectors of A corresponding to the eigenvalues 1, 1, 0 respectively. The matrix P will transform A to diagonal form D given by $P^{-1}AP=D=\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix}$.

Ex. 12. *Show that the following matrices are not similar to diagonal matrices:*

$$(i)\ \begin{bmatrix}2 & 3 & 4\\0 & 2 & -1\\0 & 0 & 1\end{bmatrix}\quad (ii)\ \begin{bmatrix}2 & -1 & 1\\2 & 2 & -1\\1 & 2 & -1\end{bmatrix}$$

Solution. (*i*) Let $A=\begin{bmatrix}2 & 3 & 4\\0 & 2 & -1\\0 & 0 & 1\end{bmatrix}$.

The characteristic equation of A is

$$\begin{vmatrix}2-\lambda & 3 & 4\\0 & 2-\lambda & -1\\0 & 0 & 1-\lambda\end{vmatrix}=0$$

or $(2-\lambda)(2-\lambda)(1-\lambda)=0$.

$\therefore$ The eigenvalues of A are 2, 2, 1.

The eigenvectors X of A corresponding to the eigenvalue 2 are given by

$$(A-2I)\,X=O$$

$$\text{or}\quad \begin{bmatrix}0 & 3 & 4\\0 & 0 & -1\\0 & 0 & -1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix}0 & 3 & 4\\0 & 0 & -1\\0 & 0 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}, \text{ applying } R_3\to R_3+R_2.$$

The coefficient matrix of these equations is of rank 2. So these equations have only one linearly independent solution. Thus the geometric multiplicity of the eigenvalue 2 is one while its algebraic multiplicity is 2. Since the geometric multiplicity of this eigenvalue is not equal to its algebraic multiplicity therefore A is not similar to a diagonal matrix.

(ii) Let $A=\begin{bmatrix}2 & -1 & 1\\2 & 2 & -1\\1 & 2 & -1\end{bmatrix}$.

The characteristic equation of A is

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

$$\text{or}\quad \begin{vmatrix}2-\lambda & -1 & 0\\2 & 2-\lambda & 1-\lambda\\1 & 2 & 1-\lambda\end{vmatrix}=0, \text{ applying } C_3\to C_3+C_2$$

$$\text{or}\quad (1-\lambda)\begin{vmatrix}2-\lambda & -1 & 0\\2 & 2-\lambda & 1\\1 & 2 & 1\end{vmatrix}=0$$

$$\text{or}\quad (1-\lambda)\begin{vmatrix}2-\lambda & -1 & 0\\1 & -\lambda & 0\\1 & 2 & 1\end{vmatrix}=0, \text{ applying } R_2\to R_2-R_3$$

or $(1-\lambda)\,[-\lambda\,(2-\lambda)+1]=0$

or $(1-\lambda)\,(\lambda^2-2\lambda+1)=0$ or $(1-\lambda)^3=0$.

$\therefore$ the eigenvalues of A are 1, 1, 1 *i.e.* 1 is the only eigenvalue of A with algebraic multiplicity 3.

The eigenvectors X of A corresponding to the eigenvalue 1 are given by

$$(A-I)\,X=O$$

$$\text{or}\quad \begin{bmatrix}1 & -1 & 1\\2 & 1 & -1\\1 & 2 & -2\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$\text{or}\quad \begin{bmatrix}1 & -1 & 1\\0 & 3 & -3\\0 & 3 & -3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}, \text{ by } R_2\to R_2-2R_1,\ R_3\to R_3-R_1$$

or $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & -3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, by $R_3 \to R_3 - R_2$.

The coefficient matrix of these equations is of rank 2. So these equations have only one linearly independent solution. Thus the geometric multiplicity of the eigenvalue 1 is 1. Since the geometric multiplicity of this eigenvalue is not equal to its algebraic multiplicity, therefore **A** is not similar to a diagonal matrix.

Exercises

1. Show that each of the following matrices is similar to a diagonal matrix. Also in each case find the diagonal form **D** and a diagonalizing matrix **P**.

(a) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$ (b) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

(c) $\begin{bmatrix} 4 & 2 & -2 \\ -5 & 3 & 2 \\ -2 & 4 & 1 \end{bmatrix}$ (d) $\begin{bmatrix} -17 & 18 & -6 \\ -18 & 19 & -6 \\ -9 & 9 & 2 \end{bmatrix}$.

2. Show that the following matrices are not similar to diagonal matrices :

(a) $\begin{bmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{bmatrix}$ (b) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$.

3. Transform the matrix $\begin{bmatrix} 8 & -12 & 5 \\ 15 & -25 & 11 \\ 24 & -42 & 19 \end{bmatrix}$ into diagonal form.

(Punjab 1972)

Answers

1. (a) $\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 15 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{bmatrix}$

(b) $\mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} -1 & 1 & 2 \\ 0 & 2 & -1 \\ 2 & 0 & 1 \end{bmatrix}$

(c) $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$

(d) $\mathbf{D} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $\mathbf{P} = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & 0 \end{bmatrix}$.

§ 3. Orthogonally Similar Matrices.

Definition. *Let* **A** *and* **B** *be square matrices of order n. Then* **B** *is said to be orthogonally similar to* **A** *if there exists an orthogonal matrix* **P** *such that*

$$\mathbf{B} = \mathbf{P}^{-1} \mathbf{A} \mathbf{P}.$$

If **A** and **B** are orthogonally similar, then they are similar also. Further it can be easily shown that the *relation of being 'orthogonally similar' is an equivalence relation in the set of all $n \times n$ matrices over the field of complex numbers.*

Theorem 1. Orthogonal reduction of real symmetric matrices. *Every real symmetric matrix is orthogonally similar to a diagonal matrix with real elements.* **(Andhra 1990)**

Proof. We shall prove the theorem by induction on n, the order of the given matrix. If $n = 1$, the theorem is obviously true. Let us assume as our induction hypothesis that the theorem is true for all real symmetric matrices of order $n-1$. Then we shall show that it is also true for an $n \times n$ real symmetric matrix **A**. Let λ_1 be an eigenvalue of **A**. Since **A** is real symmetric, therefore λ_1 is real. Let $\mathbf{X}_1$ be any unit eigenvector of **A** corresponding to the eigenvalue λ_1, so that

$$\mathbf{A}\mathbf{X}_1 = \lambda_1 \mathbf{X}_1. \qquad \ldots(1)$$

As the matrix **A** is real and the number λ_1 is real, the column vector $\mathbf{X}_1$ is also real. Since $\mathbf{X}_1$ is a real unit vector, therefore there exists an orthogonal matrix **S** with $\mathbf{X}_1$ as first column.

Now consider the matrix $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$. Since $\mathbf{X}_1$ is the first column of **S**, therefore the first column of $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is

$$= \mathbf{S}^{-1}\mathbf{A}\mathbf{X}_1$$
$$= \mathbf{S}^{-1}\lambda_1\mathbf{X}_1 \qquad [\because \mathbf{A}\mathbf{X}_1 = \lambda_1\mathbf{X}_1 \text{ from (1)}]$$
$$= \lambda_1\mathbf{S}^{-1}\mathbf{X}_1.$$

But $\mathbf{S}^{-1}\mathbf{X}_1$ is the first column of $\mathbf{S}^{-1}\mathbf{S} = \mathbf{I}$. Therefore the first column of $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is $= [\lambda_1 \;\; 0 \;\; 0 \ldots 0]^T$.

Now $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is symmetric as shown ahead. Since **S** is orthogonal, therefore $\mathbf{S}^T = \mathbf{S}^{-1}$. So

$$(\mathbf{S}^{-1}\mathbf{A}\mathbf{S})^T = (\mathbf{S}^T\mathbf{A}\mathbf{S})^T = \mathbf{S}^T \mathbf{A}^T (\mathbf{S}^T)^T = \mathbf{S}^T\mathbf{A}\mathbf{S} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}.$$

Thus $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is symmetric. Therefore the first row of $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ is $= [\lambda_1 \;\; 0 \;\; 0 \ldots 0]$.

Therefore, $\mathbf{S}^{-1}\mathbf{A}\mathbf{S} = \begin{bmatrix} \lambda_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{A}_1 \end{bmatrix}$, $\qquad \ldots(2)$

where A_1 is a square matrix of order $n-1$.

Since $S^{-1} AS$ is a real symmetric matrix, therefore A_1 is also a real symmetric matrix of order $n-1$. So by our induction hypothesis there exists an orthogonal matrix Q of order $n-1$ such that

$$Q^{-1} A_1 Q = D_1, \quad \ldots(3)$$

where D_1 is a diagonal matrix of order $n-1$. Let $R=\begin{bmatrix} 1 & O \\ O & Q \end{bmatrix}$

be an $n \times n$ matrix. Obviously R is invertible and $R^{-1}=\begin{bmatrix} 1 & O \\ O & Q^{-1} \end{bmatrix}$.

Also $R^T=\begin{bmatrix} 1 & O \\ O & Q^T \end{bmatrix}=\begin{bmatrix} 1 & O \\ O & Q^{-1} \end{bmatrix}$ [$\because$ Q is orthogonal]
$=R^{-1}$.

Therefore R is orthogonal.

Since R and S are orthogonal matrices of the same order n, therefore SR is also an orthogonal matrix of order n. Let $SR=P$. Then $P^{-1}AP=(SR)^{-1} A (SR)=R^{-1} (S^{-1} AS)R$

$$=\begin{bmatrix} 1 & O \\ O & Q^{-1} \end{bmatrix}\begin{bmatrix} \lambda_1 & O \\ O & A_1 \end{bmatrix}\begin{bmatrix} 1 & O \\ O & Q \end{bmatrix} \text{ [from (2)]}$$

$$=\begin{bmatrix} \lambda_1 & O \\ O & Q^{-1}A_1 \end{bmatrix}\begin{bmatrix} 1 & O \\ O & Q \end{bmatrix}=\begin{bmatrix} \lambda_1 & O \\ O & Q^{-1}A_1Q \end{bmatrix}$$

$$=\begin{bmatrix} \lambda_1 & O \\ O & D_1 \end{bmatrix} \text{ [from (3)]}$$

$=D$ where D is a diagonal matrix.

Thus A is orthogonally similar to a diagonal matrix D. The diagonal elements of D will be eigenvalues of A which are all real.

The proof is now complete by induction.

Corollary. *A real symmetric matrix of order* n *has* n *mutually orthogonal real eigenvectors.*

Proof. Let A be a real symmetric matrix of order n. Then there exists an orthogonal matrix P such that $P^{-1} AP=D$, where D is a diagonal matrix. Each column vector of P is an eigenvector of A. Since P is an orthogonal matrix, therefore its column vectors, are mutually orthogonal real vectors. Thus A has n mutually orthogonal real eigenvectors.

We have just seen that if A is a real symmetric matrix, then we can always find an orthogonal matrix P such that $P^{-1} AP$ is a diagonal matrix. The following two theorems will enable us to develop a practical method to find such an orthogonal matrix P.

Theorem 2. *Any two eigenvectors corresponding to two distinct eigenvalues of a real symmetric matrix are orthogonal.*

Proof. Let X_1, X_2 be two eigenvectors corresponding to two distinct eigenvalues λ_1, λ_2 of a real symmetric matrix A. Then

$$AX_1=\lambda_1 X_1 \quad \ldots(1)$$

and $$AX_2=\lambda_2 X_2 \quad \ldots(2).$$

It should be noted that the numbers λ_1, λ_2 are real and X_1, X_2 are real vectors.

Now $\lambda_1 X_2^T X_1 = X_2^T (\lambda_1 X_1)$
$= X_2^T (AX_1)$ [from (1)]
$=(X_2^T A) X_1$
$=(X_2^T A^T) X_1$ [$\because$ A is symmetric $\Rightarrow A^T=A$]
$=(AX_2)^T X_1$
$=(\lambda_2 X_2)^T X_1$ [from (2)]
$=\lambda_2 X_2^T X_1$.

$\therefore$ $\lambda_1 X_2^T X_1=\lambda_2 X_2^T X_1$
$\Rightarrow (\lambda_1-\lambda_2) X_2^T X_1=O$
$\Rightarrow X_2^T X_1=O$ [$\because$ λ_1 and λ_2 are distinct $\Rightarrow \lambda_1-\lambda_2\neq 0$]
$\Rightarrow X_1$ and X_2 are orthogonal.

Theorem 3. *If* λ *occurs exactly* p *times as an eigenvalue of a real symmetric matrix* A, *then* A *has* p *but not more than* p *mutually orthogonal real eigenvectors corresponding to* λ.

Proof. Suppose A is a real symmetric matrix of order n. By theorem 1, there exists an orthogonal matrix P such that $P^{-1} AP$ is a diagonal matrix. Thus the matrix A is diagonalizable. Therefore if λ is an eigenvalue of A having algebraic multiplicity p, then the geometric multiplicity of λ is also p. So the system of equations $(A-\lambda I) X=O$ has p linearly independent solutions in the real vector space V_n. Therefore the set of vectors X such that $(A-\lambda I) X=O$, constitute a subspace of the real vector space V_n of dimension p. But every inner product space has an orthonormal basis and an orthonormal set is linearly independent. Therefore there are p but not more than p mutually orthogonal unit vectors in this subspace. These are p mutually orthogonal eigenvectors of A. They are not, of course, uniquely determined.

Working rule for orthogonal reduction of a real symmetric matrix. Suppose A is a real symmetric matrix. First we should find the eigenvalues of A. If λ is an eigenvalue of A having p as its algebraic multiplicity, then we shall be able to find an orthonormal set of p eigenvectors of A corresponding to this eigenvalue. We should repeat this process for each eigenvalue of A.

Since the eigenvectors corresponding to two distinct eigenvalues of a real symmetric matrix are mutually orthogonal, therefore the n eigenvectors found in this manner constitute an orthonormal set.

The matrix, **P**, having as its columns the members of the orthonormal set obtained above, is orthogonal and is such that $\mathbf{P}^{-1}\mathbf{AP}$ is a diagonal matrix.

Solved Examples

Ex. 1. *Find an orthogonal matrix that will diagonalize the real symmetric matrix* $\mathbf{A}=\begin{bmatrix}1 & 2 & 3\\ 2 & 4 & 6\\ 3 & 6 & 9\end{bmatrix}$.

Also write the resulting diagonal matrix.

Solution. The characteristic equation of A is

$$\begin{vmatrix}1-\lambda & 2 & 3\\ 2 & 4-\lambda & 6\\ 3 & 6 & 9-\lambda\end{vmatrix}=0$$

or $$\begin{vmatrix}-\lambda & 2 & 3\\ 2\lambda & 4-\lambda & 6\\ -\lambda & 6 & 9-\lambda\end{vmatrix}=0, \text{ applying } C_1+C_3-2C_2$$

or $$\lambda\begin{vmatrix}-1 & 2 & 3\\ 2 & 4-\lambda & 6\\ -1 & 6 & 9-\lambda\end{vmatrix}=0$$

or $$\lambda\begin{vmatrix}-1 & 2 & 3\\ 0 & 8-\lambda & 12\\ 0 & 4 & 6-\lambda\end{vmatrix}=0, \text{ by } R_2-2R_1,\ R_3-R_1$$

or $$\lambda\,[(8-\lambda)(6-\lambda)-48]=0$$

or $$\lambda\,(\lambda^2-14\lambda)=0$$

or $$\lambda^2\,(\lambda-14)=0.$$

$\therefore$ the eigenvalues of A are 0, 0, 14.

The eigenvalue 14 is of algebraic multiplicity 1. So there will be only one linearly independent eigenvector corresponding to this eigenvalue. The eigenvectors **X** corresponding to this eigenvalue are given by

$$(\mathbf{A}-14\mathbf{I})\,\mathbf{X}=\mathbf{O}$$

or $$\begin{bmatrix}-13 & 2 & 3\\ 2 & -10 & 6\\ 3 & 6 & -5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}$$

Since these equations have only one linearly independent solution, therefore the coefficient matrix is of rank 2 and its third row can be made zero by elementary row operations. So in order to find **X** it is sufficient to find x_1, x_2, x_3 satisfying the equations

$$-13x_1+2x_2+3x_3=0, \qquad \ldots(1)$$
$$2x_1-10x_2+6x_3=0. \qquad \ldots(2)$$

Multiplying the first equation by 2 and subtracting from the second, we get $28x_1-14x_2=0$ or $2x_1-x_2=0$. Let $x_1=1$, $x_2=2$. Then (1) gives $x_3=3$. Therefore $\mathbf{X}_1=[1\ \ 2\ \ 3]^T$ is an eigenvector of A corresponding to the eigenvalue 14.

The eigenvalue 0 is of algebraic multiplicity 2. So there will be two linearly independent eigenvectors corresponding to this eigenvalue. The eigenvectors **X** corresponding to this eigenvalue are given by

$$(\mathbf{A}-0\mathbf{I})\,\mathbf{X}=\mathbf{O}$$

or $$\begin{bmatrix}1 & 2 & 3\\ 2 & 4 & 6\\ 3 & 6 & 9\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

Since these equations have two linearly independent solutions, therefore the coefficient matrix is of rank 1 and its second and third rows can be made zero by elementary row operations. So it is sufficient to find two linearly independent orthogonal solutions of the equation

$$x_1+2x_2+3x_3=0. \qquad \ldots(3)$$

Obviously $x_1=0$, $x_2=3$, $x_3=-2$ is a solution.

$\therefore$ $\mathbf{X}_2=[0\ \ 3\ \ -2]^T$ is an eigenvector of A corresponding to the eigenvalue 0. Let $\mathbf{X}_3=[x\ \ y\ \ z]^T$ be another eigenvector of A corresponding to the eigenvalue 0 and let $\mathbf{X}_3$ be orthogonal to $\mathbf{X}_2$. Then

$x+2y+3z=0$ $\quad [\because\ \mathbf{X}_3$ satisfies (3)]

and $0+3y-2z=0$ $\quad [\because\ \mathbf{X}_2$ and $\mathbf{X}_3$ are orthogonal]

Obviously $y=2$, $z=3$, $x=-13$ is a solution.

$\therefore$ $\mathbf{X}_3=[-13\ \ 2\ \ 3]^T$.

Now let us normalize the vectors $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ *i.e.* let us find unit vectors $\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3$ which are scalar multiples of $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ respectively.

Length of the vector $\mathbf{X}_1=\sqrt{(1+4+9)}=\sqrt{14}$.

$\therefore$ $\mathbf{S}_1=\dfrac{1}{\sqrt{14}}\mathbf{X}_1=a\mathbf{X}_1$ where $a=\dfrac{1}{\sqrt{14}}$.

Similarly $S_2=\dfrac{1}{\sqrt{13}}X_2=b\,X_2$ where $b=\dfrac{1}{\sqrt{13}}$

and $$S_3=\frac{1}{\sqrt{182}}X_3=c\,X_3 \text{ where } c=\frac{1}{\sqrt{182}}.$$

Let $P=[S_1\ S_2\ S_3]=\begin{bmatrix} a & 0 & -13c \\ 2a & 3b & 2c \\ 3a & -2b & 3c \end{bmatrix}$.

Then P is an orthogonal matrix and

$$P^{-1}AP=\begin{bmatrix} 14 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence $P^{-1}=P^T$ since P is orthogonal.

The order of the columns of P determines the order in which the eigenvalues of A appear in the diagonal form of A.

Ex. 2. *Determine diagonal matrices orthogonally similar to the following real symmetric matrices, obtaining also the transforming matrices :*

(*i*) $A=\begin{bmatrix} 3 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 3 \end{bmatrix}$, (*ii*) $A=\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$

(*iii*) $A=\begin{bmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{bmatrix}$, (*iv*) $A=\begin{bmatrix} 7 & 0 & -2 \\ 0 & 5 & -2 \\ -2 & -2 & 6 \end{bmatrix}$

Solution. (i) The characteristic equation of A is

$$\begin{vmatrix} 3-\lambda & -1 & 1 \\ -1 & 5-\lambda & -1 \\ 1 & -1 & 3-\lambda \end{vmatrix}=0$$

or $$\begin{vmatrix} 3-\lambda & -1 & 1 \\ 3-\lambda & 5-\lambda & -1 \\ 3-\lambda & -1 & 3-\lambda \end{vmatrix}=0, \text{ by } C_1+C_2+C_3$$

or $$(3-\lambda)\begin{vmatrix} 1 & -1 & 1 \\ 1 & 5-\lambda & -1 \\ 1 & -1 & 3-\lambda \end{vmatrix}=0$$

or $$(3-\lambda)\begin{vmatrix} 1 & -1 & 1 \\ 0 & 6-\lambda & -2 \\ 0 & 0 & 2-\lambda \end{vmatrix}=0, \text{ by } R_2-R_1,\ R_3-R_1$$

or $$(3-\lambda)(6-\lambda)(2-\lambda)=0.$$

$\therefore$ the eigenvalues of A are 2, 3, 6 which are all distinct.

In order to find an eigenvector $X=[x_1\ x_2\ x_3]^T$ corresponding to the eigenvalue 2, it is sufficient to find x_1, x_2, x_3 satisfying the equations $x_1-x_2+x_3=0,\ \ x_1+3x_2-x_3=0.$ Obviously $x_1=1$, $x_2=0$, $x_3=-1$ is a solution of these equations. Therefore $X_1=[1\ \ 0\ \ -1]^T$ is an eigenvector of A corresponding to the eigenvalue 2.

To find an eigenvector of A corresponding to the eigenvalue 3, it is sufficient to find x_1, x_2, x_3 satisfying the equations

$$0x_1-x_2+x_3=0,$$
$$-x_1+2x_2-x_3=0.$$

Obviously $x_1=1$, $x_2=1$, $x_3=1$ is a solution of these equations.

$\therefore$ $X_2=\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 3.

An eigenvector corresponding to the eigenvalue 6 is found by solving

$$-3x_1-x_2+x_3=0 \quad \ldots(1)$$
$$-x_1-x_2-x_3=0. \quad \ldots(2)$$

Adding (1) and (2), we get $-4x_1-2x_2=0$ or $2x_1+x_2=0$.

$\therefore$ $x_1=1$, $x_2=-2$, $x_3=1$ is a solution.

$\therefore$ $X_3=\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 6.

The lengths of the vectors X_1, X_2, X_3 are $\sqrt{2}, \sqrt{3}, \sqrt{6}$ respectively. Therefore $\dfrac{1}{\sqrt{2}}X_1, \dfrac{1}{\sqrt{3}}X_2, \dfrac{1}{\sqrt{6}}X_3$ are unit vectors which are scalar multiples of X_1, X_2, X_3. So if P is the required orthogonal matrix that will diagonalize A, then

$$P=\left[\frac{1}{\sqrt{2}}X_1,\ \frac{1}{\sqrt{3}}X_2,\ \frac{1}{\sqrt{6}}X_3\right]=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

We have $P^{-1}=P^T$. Thus

$$P^{-1}AP=P^T AP=\begin{bmatrix}2 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 6\end{bmatrix}.$$

(ii) **The characteristic equation of A is**

$$\begin{vmatrix}6-\lambda & -2 & 2\\ -2 & 3-\lambda & -1\\ 2 & -1 & 3-\lambda\end{vmatrix}=0$$

or $$\begin{vmatrix}6-\lambda & -2 & 2\\ -2 & 3-\lambda & 2-\lambda\\ 2 & -1 & 2-\lambda\end{vmatrix}=0, \text{ by } C_3+C_2$$

or $$(2-\lambda)\begin{vmatrix}6-\lambda & -2 & 0\\ -2 & 3-\lambda & 1\\ 2 & -1 & 1\end{vmatrix}=0$$

or $$(2-\lambda)\begin{vmatrix}6-\lambda & -2 & 0\\ -4 & 4-\lambda & 0\\ 2 & -1 & 1\end{vmatrix}=0, \text{ by } R_2-R_3$$

or $(2-\lambda)\,[(6-\lambda)(4-\lambda)-8]=0$

or $(2-\lambda)(\lambda^2-10\lambda+16)=0$

or $(2-\lambda)(\lambda-2)(\lambda-8)=0.$

$\therefore$ the eigenvalues of A are 2, 2, 8.

In order to find an eigenvector $X=\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}$ corresponding to the eigenvalue 8, it is sufficient to find x_1, x_2, x_3 satisfying the equations

$$-2x_1-2x_2+2x_3=0, \qquad \text{...(1)}$$
$$-2x_1+5x_2-x_3=0. \qquad \text{...(2)}$$

Subtracting (2) from (1), we get

$$3x_2+3x_3=0.$$

$\therefore$ $x_2=1, x_3=-1, x_1=-2$ is a solution.

Thus $X_1=\begin{bmatrix}-2\\ 1\\ 1\end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 8.

The eigenvalue 2 is of algebraic multiplicity 2. So we are to find two mutually orthogonal eigenvectors corresponding to it. For this we should find two orthogonal solutions of the equation

$$4x_1-2x_2+2x_3=0$$

i.e. $$2x_1-x_2+x_3=0.$$

Obviously $x_1=0, x_2=1, x_3=1$ is a solution.

$\therefore$ $X_2=\begin{bmatrix}0\\ 1\\ 1\end{bmatrix}$ is an eigenvector of A corresponding to the eigenvalue 2. Let $X_3=\begin{bmatrix}x\\ y\\ z\end{bmatrix}$ be another eigenvector of A corresponding to the eigenvalue 2 and let X_3 be orthogonal to X_2.

Then $\quad 2x-y+z=0,$

and $\quad 0+y+z=0.$

Obviously $y=1, z=-1, x=1$ is a solution.

$$\therefore \quad X_3=\begin{bmatrix}1\\ 1\\ -1\end{bmatrix}.$$

Lengths of the vectors X_1, X_2, X_3 are $\sqrt{6}, \sqrt{2}, \sqrt{3}$ respectively.

$\therefore$ $P=\begin{bmatrix}-2/\sqrt{6} & 0 & 1/\sqrt{3}\\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3}\\ -1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3}\end{bmatrix}$ is the required orthogonal matrix that will diagonalize A. We have $P^{-1}=P^T$ and

$$P^{-1}AP=P^TAP=\begin{bmatrix}8 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix}$$

(iii) **The characteristic equation of A is**

$$\begin{vmatrix}7-\lambda & 4 & -4\\ 4 & -8-\lambda & -1\\ -4 & -1 & -8-\lambda\end{vmatrix}=0$$

or $$\begin{vmatrix}7-\lambda & 4 & 0\\ 4 & -8-\lambda & -9-\lambda\\ -4 & -1 & -9-\lambda\end{vmatrix}=0, \text{ by } C_3+C_2.$$

It can be easily seen that the eigenvalues of A are 9, −9, −9. $X_1=[4, 1, -1]^T$, is an eigenvector corresponding to $\lambda=9$. $X_2=[0, 1, 1]^T$, $X_3=[1, -2, 2]^T$ are two mutually orthogonal eigenvectors corresponding to $\lambda=-9$.

$P=\begin{bmatrix}\frac{4}{\sqrt{18}} & 0 & \frac{1}{3}\\ \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} & -\frac{2}{3}\\ -\frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} & \frac{2}{3}\end{bmatrix}$ is the required orthogonal matrix that will diagonalize **A**.

Also $\mathbf{P}^{-1}\mathbf{AP}=\mathbf{P}^T\mathbf{AP}=\text{diag}\,[9, -9, -9]$.

(iv) The characteristic equation of **A** is

$$\begin{vmatrix} 7-\lambda & 0 & -2 \\ 0 & 5-\lambda & -2 \\ -2 & -2 & 6-\lambda \end{vmatrix}=0$$

or $$\lambda^3-18\lambda^2+99\lambda-162=0.$$

$\therefore$ $$\lambda=3, 6, 9.$$

$\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ are eigenvectors of **A** corresponding to $\lambda=3, 6, 9$ respectively.

Let $\mathbf{P}=\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \end{bmatrix}$. Then **P** is the required orthogonal matrix and $\mathbf{P}^{-1}\,\mathbf{AP}=\text{diag}\,[3, 6, 9]$.

Ex. 3. *If* **P** *is a real orthogonal matrix and* **D** *a real diagonal matrix such that* $\mathbf{P}^{-1}\mathbf{AP}=\mathbf{D}$, *show that* **A** *is a real symmetric matrix.*

Solution. Since **P** is a real orthogonal matrix, therefore $\mathbf{P}^{-1}$ is also a real matrix and $\mathbf{P}^{-1}=\mathbf{P}^T$. We have

$$\mathbf{P}^{-1}\,\mathbf{AP}=\mathbf{D}$$
$$^*\Rightarrow\ \mathbf{A}=\mathbf{PDP}^{-1}=\mathbf{PDP}^T.$$

Since **P**, **D**, $\mathbf{P}^T$ are all real matrices, therefore **A** is also a real matrix. Also

$$\begin{aligned} \mathbf{A}^T &= (\mathbf{PDP}^T)^T=\mathbf{PD}^T\mathbf{P}^T \\ &= \mathbf{PDP}^T \quad [\because\ \mathbf{D} \text{ is a diagonal matrix} \Rightarrow \mathbf{D}^T=\mathbf{D}] \\ &= \mathbf{A}. \end{aligned}$$

$\therefore$ **A** is a symmetric matrix.

Exercises

1. If an $n\times n$ matrix **A** possesses a set of n real orthogonal eigenvectors $\mathbf{X}_1,\ldots, \mathbf{X}_n$, then **A** is orthogonally similar to a diagonal matrix.
2. If a matrix be orthogonally similar to a diagonal matrix, it must be symmetric.
3. Find orthogonal matrices that will diagonalize each of the following real symmetric matrices :

(a) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$ (b) $\begin{bmatrix} 1 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 3 \end{bmatrix}$.

*The symbol $\Rightarrow$ is read as 'implies'.

Answers

3. (a) $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$ (b) $\begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$.

§ 4. Unitarily Similar Matrices.

Definition. *Let* **A** *and* **B** *be square matrices of order* n. *Then* **B** *is said to be unitarily similar to* **A** *if there exists a unitary matrix* **P** *such that*

$$\mathbf{B}=\mathbf{P}^{-1}\,\mathbf{AP}.$$

If **A** and **B** are unitarily similar, then they are similar also.

Theorem 1. *The relation of being 'unitarily similar' is an equivalence relation in the set of all* $n\times n$ *matrices over the field of complex numbers.*

Proof. Reflexivity. If **A** is any $n\times n$ complex matrix, then $\mathbf{A}=\mathbf{I}^{-1}\mathbf{AI}$ where the identity matrix **I** is a unitary matrix. Therefore **A** is unitarily similar to **A**.

Symmetry. Let **A** be unitarily similar to **B**. Then

$\mathbf{A}=\mathbf{P}^{-1}\mathbf{BP}$, where **P** is a unitary matrix

$\Rightarrow\ \mathbf{PAP}^{-1}=\mathbf{B}$

$\Rightarrow\ (\mathbf{P}^{-1})^{-1}\,\mathbf{AP}^{-1}=\mathbf{B}$

$\Rightarrow$ **B** is unitarily similar to **A** since $\mathbf{P}^{-1}$ is also a unitary matrix.

Transitivity. Let **A** be unitarily similar to **B** and **B** be unitarily similar to **C**. Then $\mathbf{A}=\mathbf{P}^{-1}\mathbf{BP}$, $\mathbf{B}=\mathbf{Q}^{-1}\mathbf{CQ}$ where **P** and **Q** are unitary matrices. From these, we get

$$\mathbf{A}=\mathbf{P}^{-1}\,(\mathbf{Q}^{-1}\mathbf{CQ})\,\mathbf{P}=(\mathbf{QP})^{-1}\mathbf{C}\,(\mathbf{QP}).$$

If **P** and **Q** are unitary matrices, then **QP** is also a unitary matrix. So **A** is unitarily similar to **C**.

Theorem 2. Unitary Reduction of Hermitian Matrices. *Every Hermitian matrix is unitarily similar to a diagonal matrix.*

Proof. We shall prove the theorem by induction on n, the order of the given matrix. If $n=1$, the theorem is obviously true. Let us assume as our induction hypothesis that the theorem is true for all Hermitian matrices of order $n-1$. Then we shall show that it is also true for an $n\times n$ Hermitian matrix A.

Let λ_1 be an eigenvalue of A. Since A is a Hermitian matrix, therefore λ_1 is real. Let X_1 be any unit eigenvector of A corresponding to the eigenvalue λ_1, so that

$$AX_1=\lambda_1X_1. \quad \ldots(1)$$

It is possible to choose an orthonormal basis of the complex vector space V_n, having X_1 as a member. Therefore there exists a unitary matrix S with X_1 as its first column.

Now consider the matrix $S^{-1}AS$. Since X_1 is the first column of S, therefore the first column of $S^{-1}AS$ is

$$=S^{-1}AX_1$$
$$=S^{-1}\lambda_1X_1 \qquad [\because \ AX_1=\lambda_1X_1 \text{ from (1)}]$$
$$=\lambda_1S^{-1}X_1.$$

But $S^{-1}X_1$ is the first column of $S^{-1}S=I$. Therefore the first column of $S^{-1}AS$ is $=[\lambda_1 \ \ 0 \ \ 0\ldots\ldots 0]^T$.

Now S is unitary implies that $S^{-1}=S^\theta$. Therefore

$$(S^{-1}AS)^\theta=(S^\theta AS)^\theta=S^\theta A^\theta(S^\theta)^\theta$$
$$=S^\theta AS \qquad [\because \ A \text{ is Hermitian} \Rightarrow A^\theta=A]$$
$$=S^{-1}AS.$$

Thus $S^{-1}AS$ is a Hermitian matrix. Therefore the first row of $S^{-1}AS$ is $=[\lambda_1, \ \ 0, \ \ 0,\ldots, 0]$. Therefore

$$S^{-1}AS=\begin{bmatrix}\lambda_1 & O\\ O & A_1\end{bmatrix}, \quad \ldots(2)$$

where A_1 is a square matrix of order $n-1$.

Since $S^{-1}AS$ is a Hermitian matrix, therefore A_1 is also a Hermitian matrix of order $n-1$. So by our induction hypothesis there exists a unitary matrix Q of order $n-1$ such that

$$Q^{-1}A_1Q=D_1, \quad \ldots(3)$$

where D_1 is a diagonal matrix of order $n-1$.

Let $R=\begin{bmatrix}1 & O\\ O & Q\end{bmatrix}$ be an $n\times n$ matrix. Obviously R is invertible and $R^{-1}=\begin{bmatrix}1 & O\\ O & Q^{-1}\end{bmatrix}$. Also

$$R^\theta=\begin{bmatrix}1 & O\\ O & Q^\theta\end{bmatrix}=\begin{bmatrix}1 & O\\ O & Q^{-1}\end{bmatrix}, \text{ since Q is unitary}$$
$$=R^{-1}.$$ Therefore R is unitary.

Since R and S are unitary matrices of the same order n, therefore SR is also a unitary matrix of order n. Let $SR=P$. Then

$$P^{-1}AP=(SR)^{-1}A(SR)=R^{-1}(S^{-1}AS)R$$
$$=\begin{bmatrix}1 & O\\ O & Q^{-1}\end{bmatrix}\begin{bmatrix}\lambda_1 & O\\ O & A_1\end{bmatrix}\begin{bmatrix}1 & O\\ O & Q\end{bmatrix} \text{ [from (2)]}$$
$$=\begin{bmatrix}\lambda_1 & O\\ O & Q^{-1}A_1\end{bmatrix}\begin{bmatrix}1 & O\\ O & Q\end{bmatrix}=\begin{bmatrix}\lambda_1 & O\\ O & Q^{-1}A_1Q\end{bmatrix}$$
$$=\begin{bmatrix}\lambda_1 & O\\ O & D_1\end{bmatrix} \text{ [from (3)]}$$
$$=D, \text{ where D is a diagonal matrix.}$$

Thus A is unitarily similar to a diagonal matrix. The proof is now complete by induction.

Corollary. *An $n\times n$ Hermitian matrix* **H** *has n mutually orthogonal eigenvectors in the complex vector space V_n.*

Proof. Let H be a Hermitian matrix of order n. Then there exists a unitary matrix P such that $P^{-1}HP=D$, where D is a diagonal matrix. Each column vector of P is an eigenvector of H. Since P is a unitary matrix, therefore its column vectors are mutually orthogonal vectors. Hence H has n mutually orthogonal eigenvectors in the complex vector space V_n.

The following two theorems will enable us to develop a practical method to find a unitary matrix P that will diagonalize a given Hermitian matrix A.

Theorem 3. *Any two eigenvectors corresponding to two distinct eigenvalues of a Hermitian matrix are orthogonal.*

Proof. Let X_1, X_2 be two eigenvectors corresponding to two distinct eigenvalues λ_1, λ_2 of a Hermitian matrix H. Then

$$AX_1=\lambda_1X_1 \quad \ldots(1)$$
and
$$AX_2=\lambda_2X_2. \quad \ldots(2)$$

Since A is Hermitian therefore λ_1, λ_2 are real.

Now $\lambda_1X_2^\theta X_1=X_2^\theta(\lambda_1X_1)=X_2^\theta(AX_1)$ [from (1)]

$$=(X_2^\theta A)X_1$$
$$=(X_2^\theta A^\theta)X_1 \qquad [\because \ A \text{ is Hermitian} \Rightarrow A^\theta=A]$$
$$=(AX_2)^\theta X_1$$
$$=(\lambda_2X_2)^\theta X_1 \qquad \text{[from (2)]}$$
$$=\bar{\lambda}_2X_2^\theta X_1=\lambda_2X_2^\theta X_1 \qquad [\because \ \lambda_2 \text{ is real}]$$

$\therefore \quad \lambda_1X_2^\theta X_1=\lambda_2X_2^\theta X_1$

$\Rightarrow \quad (\lambda_1-\lambda_2)X_2^\theta X_1=O$

$\Rightarrow X_2^\theta X_1 = O$ $[\because \lambda_1 \neq \lambda_2 \Rightarrow \lambda_1 - \lambda_2 \neq 0]$

$\Rightarrow X_1$ and X_2 are orthogonal.

Theorem 4. *If λ occurs exactly p times as an eigenvalue of a Hermitian matrix A then A has p but not more than p mutually orthogonal eigenvectors corresponding to λ.*

Proof. Proceed as in theorem 3 of § 3.

Solved Examples

Ex. 1. *Determine the diagonal matrix unitarily similar to the Hermitian matrix* $A = \begin{bmatrix} 2 & 1-2i \\ 1+2i & -2 \end{bmatrix}$, *obtaining also the transformation matrix.*

Solution. The characteristic equation of A is

$$\begin{vmatrix} 2-\lambda & 1-2i \\ 1+2i & -2-\lambda \end{vmatrix} = 0$$

or $(\lambda-2)(\lambda+2)-(1+2i)(1-2i)=0$

or $\lambda^2-4-(1+4)=0$ or $\lambda^2-9=0$.

$\therefore$ the eigenvalues of A are $-3, 3$.

The eigenvector $X = \begin{bmatrix} x \\ y \end{bmatrix}$ corresponding to the eigenvalue -3 is given by

$$\begin{bmatrix} 5 & 1-2i \\ 1+2i & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

i.e. $5x+(1-2i)\,y=0$, and $(1+2i)\,x+y=0$.

Obviously $x=1-2i$, $y=-5$ is a solution.

$\therefore X_1 = \begin{bmatrix} 1-2i \\ -5 \end{bmatrix}$ is an eigenvector corresponding to $\lambda=-3$.

Corresponding to $\lambda=3$, the eigenvector is given by

$$\begin{bmatrix} -1 & 1-2i \\ 1+2i & -5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

i.e. $-x+(1-2i)\,y=0$ and $(1+2i)\,x-5y=0$.

Obviously $x=5$, $y=1+2i$ is a solution.

$\therefore X_2 = \begin{bmatrix} 5 \\ 1+2i \end{bmatrix}$ is an eigenvector corresponding to $\lambda=3$.

Length of the vector $X_1 = \sqrt{(|1-2i|^2+|-5|^2)} = \sqrt{30}$

Length of the vector $X_2 = \sqrt{(|5|^2+|1+2i|^2)} = \sqrt{30}$.

$\therefore$ the unitary matrix P that will transform A to diagonal form is

$$P = \left[\frac{1}{\sqrt{30}} X_1, \frac{1}{\sqrt{30}} X_2\right] = \begin{bmatrix} (1-2i)/\sqrt{30} & 5/\sqrt{30} \\ -5/\sqrt{30} & (1+2i)/\sqrt{30} \end{bmatrix}.$$

Also $P^{-1}=P^\theta$ and $P^{-1}AP = \begin{bmatrix} -3 & 0 \\ 0 & 3 \end{bmatrix} = \text{diag.}\,[-3, 3]$.

Ex. 2. *Show that if P is unitary and $P^{-1}AP$ is a real diagonal matrix, then A is Hermitian.*

Solution. Since P is unitary, therefore $P^{-1}=P^\theta$.

Let $P^{-1}AP=D$ where D is a real diagonal matrix. Then

$$A = PDP^{-1} = PDP^\theta.$$

$\therefore A^\theta = (PDP^\theta)^\theta = PD^\theta P^\theta = PDP^\theta$ $[\because$ D is real $\Rightarrow D^\theta = D]$
$= A$.

$\therefore$ A is Hermitian.

Exercises

1. If an $n \times n$ matrix A possesses a set of orthogonal eigenvectors $X_1, \ldots, X_n$, then it is unitarily similar to a diagonal matrix.
2. Find a unitary matrix that will diagonalize the Hermitian matrix

$$\begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}.$$

Answers

2. $\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & i/\sqrt{2} \end{bmatrix}$.

§ 5. Normal Matrices.

Normal Matrix. Definition. *A matrix A is said to be normal if* $AA^\theta = A^\theta A$.

Theorem 1. *Prove that Hermitian, real symmetric, unitary, real orthogonal, skew-Hermitian, and real skew-symmetric matrices are normal.*

Proof. (i) Let A be a Hermitian or a real symmetric matrix. Then $A^\theta = A$. Therefore $AA^\theta = A^\theta A$ and thus A is normal.

(ii) Let A be a unitary or real orthogonal matrix. Then $A^\theta A = I = AA^\theta$. Therefore A is normal.

(iii) Let A be a skew-Hermitian or a real skew-symmetric matrix. Then $A^\theta = -A$. Therefore $AA^\theta = A(-A) = -A^2$ and
$A^\theta A = (-A)A = -A^2$.

Thus $AA^\theta = A^\theta A$ and so A is normal.

Theorem 2 *Prove that any diagonal matrix over the complex field is normal.*

Proof. Let D be a diagonal matrix over the complex field and let $D = \text{diag}\,[d_1, d_2, \ldots, d_n]$.

Then $D^\theta = \text{diag.}\ [\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_n]$.

We have $DD^\theta = \text{diag.}\ [d_1\bar{d}_1, d_2\bar{d}_2, \ldots, d_n\bar{d}_n]$

$= \text{diag.}\ [\bar{d}_1 d_1, \bar{d}_2 d_2, \ldots, \bar{d}_n d_n]$

$= D^\theta D$. Therefore D is normal.

Theorem 3. *Every matrix unitarily similar to a normal matrix is normal.*

Proof. Suppose B is unitarily similar to a normal matrix A. Then $B = P^{-1} AP$ where P is a unitary matrix.

To prove that B is a normal matrix.

Since P is unitary, therefore $P^{-1} = P^\theta$.

$\therefore$ $B = P^\theta AP$, and $B^\theta = (P^\theta AP)^\theta = P^\theta A^\theta P$.

Now $BB^\theta = (P^\theta AP)\,(P^\theta A^\theta P) = P^\theta A\,(PP^\theta)\,A^\theta P$

$= P^\theta AIA^\theta P$ $\quad [\because$ P is unitary $\Rightarrow PP^\theta = I]$

$= P^\theta AA^\theta P$

$= P^\theta A^\theta AP$ $\quad [\because$ A is normal $\Rightarrow AA^\theta = A^\theta A]$

$= P^\theta A^\theta IAP$

$= P^\theta A^\theta PP^\theta AP$ $\quad [\because \ PP^\theta = I]$

$= (P^\theta A^\theta P)\,(P^\theta AP) = B^\theta B$.

$\therefore$ B is normal.

Theorem 4. *If X is an eigenvector of a normal matrix A corresponding to an eigenvalue λ, then X is also an eigenvector of A^θ; the corresponding eigenvalue being $\bar{\lambda}$.*

Proof. Let A be a normal matrix and X be an eigenvector of A corresponding to the eigenvalue λ.

Then $AX = \lambda X$ *i.e.* $(A - \lambda I)\,X = O$.

Let $(A^\theta - \bar{\lambda} I)\,X = Y$. Then $Y^\theta = X^\theta\,(A^\theta - \bar{\lambda} I)^\theta = X^\theta\,(A - \lambda I)$.

Also A is normal implies that $A^\theta A = AA^\theta$. Then it can be easily seen that $A - \lambda I$ and $A^\theta - \bar{\lambda} I$ commute.

Now $Y^\theta Y = X^\theta\,(A - \lambda I)\,(A^\theta - \bar{\lambda} I)\,X$

$= X^\theta\,(A^\theta - \bar{\lambda} I)\,(A - \lambda I)\,X = X^\theta\,(A^\theta - \bar{\lambda} I)\,O = O$.

Therefore $Y = O \Rightarrow (A^\theta - \bar{\lambda} I)\,X = O \Rightarrow A^\theta X = \bar{\lambda} X$

$\Rightarrow$ X is an eigenvector of A^θ corresponding to the eigenvalue $\bar{\lambda}$.

Theorem 5. *Characteristic vectors corresponding to distinct characteristic values of a normal matrix are orthogonal.*

Proof. Let A be a normal matrix and X_1, X_2 be characteristic vectors of A corresponding to the characteristic values λ_1, λ_2 respectively where $\lambda_1 \neq \lambda_2$. Then $AX_1 = \lambda_1 X_1$ and $AX_2 = \lambda_2 X_2$.

Since A is normal, therefore X_2 will be a characteristic vector of A^θ corresponding to the characteristic value $\bar{\lambda}_2$. So $A^\theta X_2 = \bar{\lambda}_2 X_2$.

Now $\lambda_1 X_2^\theta X_1 = X_2^\theta \lambda_1 X_1 = X_2^\theta AX_1 = X_2^\theta (A^\theta)^\theta X_1$

$= (A^\theta X_2)^\theta X_1 = (\bar{\lambda}_2 X_2)^\theta X_1 = \lambda_2 X_2^\theta X_1$.

$\therefore$ $(\lambda_1 - \lambda_2)\, X_2^\theta X_1 = O$

$\Rightarrow$ $X_2^\theta X_1 = O$, since $\lambda_1 - \lambda_2 \neq 0$

$\Rightarrow$ X_1 and X_2 are orthogonal vectors.

Corollary. *Eigenvectors corresponding to distinct eigenvalues of Hermitian, real symmetric, unitary, real orthogonal, skew-Hermitian, and real skew-symmetric matrices are orthogonal.*

Proof. As proved in theorem 1, allthese matrices are normal matrices. So the result follows from the above theorem.

Theorem 6. *A triangular matrix is normal if and only if it is diagonal.*

Proof. We know that a diagonal matrix is normal. So if a triangular matrix is diagonal, then it is normal.

Now it remains to prove that every normal triangular matrix is diagonal. We shall prove the result by induction on n, the order of the given matrix. If $n = 1$, the result is obviously true. Let us assume as our induction hypothesis that the result is true for matrices of order $n-1$. Then we shall prove that it is also true for an $n \times n$ matrix A. Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix}$$

be a normal triangular matrix. The element in the first row and first column of AA^θ is $a_{11}\,\bar{a}_{11} + a_{12}\bar{a}_{12} + \ldots + a_{1n}\,\bar{a}_{1n}$ and the corresponding element of $A^\theta A$ is $\bar{a}_{11}\,a_{11}$.

But $AA^\theta = A^\theta A$, since A is normal. Therefore we must have

$a_{11}\,\bar{a}_{11} + a_{12}\,\bar{a}_{12} + \ldots + a_{1n}\,\bar{a}_{1n} = \bar{a}_{11}\,a_{11}$

$\Rightarrow$ $|a_{12}|^2 + |a_{13}|^2 + \ldots + |a_{1n}|^2 = 0$

$\Rightarrow$ $a_{12} = 0, a_{13} = 0, \ldots, a_{1n} = 0$ $\quad [\because \ |a_{1i}|^2 \geqslant 0]$

$\therefore$ $A = \begin{bmatrix} a_{11} & O \\ O & A_1 \end{bmatrix}$, where A_1 is a matrix of order $n-1$.

Since A is triangular, therefore A_1 is also triangular.

Also $A^\theta = \begin{bmatrix} \bar{a}_{11} & O \\ O & A_1^\theta \end{bmatrix}$.

$\therefore \quad AA^\theta = \begin{bmatrix} a_{11}\bar{a}_{11} & O \\ O & A_1A_1^\theta \end{bmatrix}$ and $A^\theta A = \begin{bmatrix} \bar{a}_{11}a_{11} & O \\ O & A_1^\theta A_1 \end{bmatrix}$.

So $AA^\theta = A^\theta A \Rightarrow A_1A_1^\theta = A_1^\theta A_1 \Rightarrow A_1$ is normal.

Thus A_1 is a normal triangular matrix of order $n-1$. So by inductive hypothesis A_1 is diagonal and hence A is diagonal and the proof is complete by induction.

Theorem 7. Triangularization theorem or Jacobi's theorem. *Every square matrix is unitarily similar to a triangular matrix.*

Proof. We shall prove the theorem by induction on n, the order of the given matrix. If $n=1$, the theorem is obviously true. Let us assume as our induction hypothesis that the theorem is true for all matrices of order $n-1$. Then we shall show that it is also true for all matrices of order n.

Let A be a square matrix of order n. Let λ_1 be an eigenvalue of A. Let X_1 be any unit eigenvector of A corresponding to the eigenvalue λ_1, so that

$$AX_1 = \lambda_1 X_1. \qquad \ldots(1)$$

It is possible to choose an orthonormal basis of the complex vector space V_n having X_1 as a member. Therefore there exists a unitary matrix S with X_1 as its first column.

Now consider the matrix $S^{-1}AS$. Since X_1 is the first column of S, therefore the first column of $S^{-1}AS$ is $= S^{-1}AX_1 = S^{-1}\lambda_1 X_1$ [from (1)] $= \lambda_1 S^{-1}X_1$.

But $S^{-1}X_1$ is the first column of $S^{-1}S = I$. Therefore the first column of $S^{-1}AS$ is

$$= \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\therefore \quad S^{-1}AS = \begin{bmatrix} \lambda_1 & B_1 \\ O & A_1 \end{bmatrix}$, where B_1 is an $1 \times (n-1)$ matrix and A_1 is a square matrix of order $n-1$.

By our induction hypothesis there exists a unitary matrix Q of order $n-1$ such that

$$Q^{-1}A_1Q = C, \qquad \ldots(2)$$

where C is a triangular matrix of order $n-1$.

Let $R = \begin{bmatrix} 1 & O \\ O & Q \end{bmatrix}$ be an $n \times n$ matrix. Obviously R is invertible and $R^{-1} = \begin{bmatrix} 1 & O \\ O & Q^{-1} \end{bmatrix}$. Also $R^\theta = \begin{bmatrix} 1 & O \\ O & Q^\theta \end{bmatrix} = \begin{bmatrix} 1 & O \\ O & Q^{-1} \end{bmatrix} = R^{-1}$.

Therefore R is unitary.

Since R and S are unitary matrices of the same order n, therefore SR is also a unitary matrix of order n. Let $SR = P$. Then

$$P^{-1}AP = (SR)^{-1}A(SR) = R^{-1}(S^{-1}AS)R$$

$$= \begin{bmatrix} 1 & O \\ O & Q^{-1} \end{bmatrix} \begin{bmatrix} \lambda_1 & B_1 \\ O & A_1 \end{bmatrix} \begin{bmatrix} 1 & O \\ O & Q \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & B_1 \\ O & Q^{-1}A_1 \end{bmatrix} \begin{bmatrix} 1 & O \\ O & Q \end{bmatrix} = \begin{bmatrix} \lambda_1 & B_1Q \\ O & Q^{-1}A_1Q \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & B_1Q \\ O & C \end{bmatrix} \text{ [from (2)]}$$

$= T$ where T is a triangular matrix.

Thus A is unitarily similar to a triangular matrix T. The proof is now complete by induction.

Note. Similar matrices have same eigenvalues and the eigenvalues of a triangular matrix are precisely its diagonal elements. Therefore the diagonal elements of T are precisely the eigenvalues of A.

Theorem 8. *A matrix A is unitarily similar to a diagonal matrix if and only if it is normal.*

Proof. Suppose A is unitarily similar to a diagonal matrix D. Then $P^{-1}AP = D$ where P is unitary.

To prove that A is normal.

We have $A = PDP^{-1} = PDP^\theta$, since P is unitary $\Rightarrow P^\theta = P^{-1}$.

$\therefore \quad A^\theta = (PDP^\theta)^\theta = PD^\theta P^\theta$.

Now $\quad AA^\theta = (PDP^\theta)(PD^\theta P^\theta) = PDD^\theta P^\theta$

$= PD^\theta DP^\theta \quad$ [$\because$ D is diagonal $\Rightarrow$ D is normal]

$= PD^\theta P^\theta PDP^\theta = A^\theta A$.

So A is normal.

Conversely suppose that A is normal. By Jacobi's theorem there exists a unitary matrix P such that $P^{-1}AP = T$ where T is a triangular matrix.

Since P is unitary and A is normal, therefore $P^{-1}AP$ is normal [See theorem 3]

Now T is a normal triangular matrix.

Therefore T is a diagonal matrix. [See theorem 6]

Thus $P^{-1}AP$ = a diagonal matrix.

Therefore A is unitarily similar to a diagonal matrix.

Corollary. *An $n \times n$ matrix A has n mutually orthogonal eigen-vectors if and only if it is normal.*

Solved Examples

Ex. 1. *Prove that a square matrix A is normal if and only if it can be expressed as $B+iC$, where B and C are commutative Hermitian matrices.*

Solution. We know that every square matrix A can be uniquely expressed as $A=B+iC$ where B and C are Hermitian matrices.

We have $A=B+iC$

$$\Rightarrow A^\theta=(B+iC)^\theta=B^\theta+\bar{i}C^\theta$$

$$=B-iC \qquad [\because \text{ B and C are Hermitian}]$$

$\therefore$ $AA^\theta=(B+iC)(B-iC)=B^2-iBC+iCB+C^2$

and $A^\theta A=(B-iC)(B+iC)=B^2+iBC-iCB+C^2$.

$\therefore$ $AA^\theta=A^\theta A$ if and only if $-iBC+iCB=iBC-iCB$

i.e. if and only if $2iBC=2iCB$

i.e. if and only if $BC=CB$.

Thus A is normal if and only if B and C commute.

Ex. 2. *If A is normal and non-singular, prove that so also is* A^{-1}.

Solution. Suppose A is normal, and non-singular. Then $AA^\theta=A^\theta A$ and A^θ is also non-singular. We have

$(A^{-1})(A^{-1})^\theta=A^{-1}(A^\theta)^{-1}=(A^\theta A)^{-1}=(AA^\theta)^{-1}=(A^\theta)^{-1}A^{-1}=(A^{-1})^\theta A^{-1}$.

Therefore A^{-1} is normal.

Ex. 3. *If A is normal, then show that A is similar to* A^T.

Solution. Since A is normal therefore there exists a unitary matrix P such that $P^{-1}AP=D$ where D is a diagonal matrix. Now $D=P^{-1}AP \Rightarrow D^T=(P^{-1}AP)^T \Rightarrow D=P^TA^T(P^{-1})^T \Rightarrow D=P^T A^T (P^T)^{-1}$
$\Rightarrow A^T$ is similar to D. Thus A^T is similar to D and D is similar to A. Therefore A^T is similar to A.

Ex. 4. *Let A be a normal matrix. Show that (i) if all the characteristic roots of A are real, then A is Hermitian; (ii) if all the characteristic roots of A are of modulus 1, then A is unitary.*

Solution. (i) Suppose A is a normal matrix having all its characteristic roots real. Since A is normal, therefore there exists a unitary matrix P such that $P^{-1}AP=D$ where D is a diagonal matrix.

Since D and A are similar matrices, therefore they have same eigenvalues. So the diagonal elements of D are all real and D is real diagonal matrix.

Now $P^{-1}AP=D \Rightarrow A=PDP^{-1}=PDP^\theta$.

$\therefore$ $A^\theta=(PDP^\theta)^\theta=PD^\theta P^\theta$

$=PDP^\theta$ $[\because$ D is real diagonal matrix $\Rightarrow D^\theta=D]$

$=A$.

$\therefore$ A is Hermitian.

(ii) Suppose A is a normal matrix having all its eigenvalues of unit modulus. Since A is normal, therefore there exists a unitary matrix P such that

$$P^{-1}AP=D=\text{diag}\,[d_1, d_2, \ldots, d_n]$$

Since A and D are similar matrices, therefore they have same eigenvalues. But eigenvalues of D are precisely its diagonal elements. Therefore $|d_i|=1, i=1, 2, \ldots, n$.

Now $P^{-1}AP=D \Rightarrow A=PDP^{-1}=PDP^\theta$.

$\therefore$ $AA^\theta=(PDP^\theta)(PDP^\theta)^\theta=PDP^\theta PD^\theta P^\theta$

$=PDID^\theta P^\theta=PDD^\theta P^\theta$

$=P.\text{diag}\,[d_1\bar{d}_1, d_2\bar{d}_2, \ldots, d_n\bar{d}_n]\,P^\theta$

$=P.\text{diag}\,[\,|d_1|^2, |d_2|^2, \ldots, |d_n|^2]\,P^\theta$

$=P.\text{diag}\,[1, 1, \ldots, 1]\,P^\theta=PIP^\theta=PP^\theta=I$.

$\therefore$ A is unitary.

Ex. 5. *If A, B are square matrices each of order n and I is the corresponding unit matrix, show that the equation*

$$AB-BA=I$$

can never hold. (I.C.S. 1986)

Solution. Let us suppose that $AB-BA=I$.

Then $AB-BA=I$

$\Rightarrow$ trace $(AB-BA)$ = trace I

$\Rightarrow$ trace (AB) − trace $(BA)=n$

[$\because$ trace I = sum of the elements of I lying along its principal diagonal $=n$]

$\Rightarrow 0=n,$ $[\because \text{tr}\,(AB)=\text{tr}\,(BA)]$

which is not possible because n is a positive integer.

Hence our assumption that $AB-BA=I$ is wrong and so the equation $AB-BA=I$ can never hold.

Ex. 6. *Prove that the trace of a matrix is equal to the sum of its characteristic roots.* (I. A. S. 1982)

Solution. Let A be a square matrix of order n. By Jacobi's theorem there exists a unitary matrix P such that

$$P^{-1}AP=T,$$

where T is a triangular matrix.

Since the matrices A and **T are similar**, therefore they **have** the same characteristic roots. Also the characteristic roots of T are just the diagonal elements of T because T is a triangular matrix.

Now trace $(P^{-1}AP) = \text{tr}\{(P^{-1}A)\,P\}$

$= \text{tr}\{P\,(P^{-1}A)\}$ [$\because$ tr $(AB) = $ tr (BA)]

$= \text{tr}\{(P\,P^{-1})\,A\} = \text{tr}\,(IA) = \text{tr}\,A.$

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

= sum of the elements of T lying along its principal diagonal

= sum of the characteristic roots of T

= sum of the characteristic roots of A because A and T have the same characteristic roots.

Hence the trace of a matrix is equal to the sum of its characteristic roots.

Exercises

1. If A and B are normal, and if A and B^θ commute, then show that AB and BA are normal.
2. Let A be a normal matrix. Show that if all the eigenvalues of A are zero or purely imaginary, then A is skew-Hermitian.

---

11 Quadratic Forms

§ 1. Quadratic Forms. Definition. *An expression of the form*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j,$$

where a_{ij}'s are elements of a field F, is called a quadratic form in the n variables $x_1, x_2, \ldots\ldots, x_n$ over a field F. **(Jabalpur 1970)**

Real Quadratic Form. Definition. *An expression of the form*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j,$$

where a_{ij}'s are all real numbers, is called a real quadratic form in the n variables $x_1, x_2, \ldots, x_n$. For example,

(i) $2x^2+7xy+5y^2$ is a real quadratic form in the two variables x and y.

(ii) $2x^2-y^2+2z^2-2yz-4zx+6xy$ is a real quadratic form in the three variables x, y and z.

(iii) $x_1^2-2x_2^2+4x_3^2-4x_4^2-2x_1x_2+3x_1x_4+4x_2x_3-5x_3x_4$ is a real quadratic form in the four variables x_1, x_2, x_3 and x_4.

Theorem. *Every quadratic form over a field F in n variables $x_1, x_2, \ldots, x_n$ can be expressed in the form* **X'BX** *where*

$$\mathbf{X} = [x_1, x_2, \ldots\ldots, x_n]^T$$

is a column vector and **B** *is a symmetric matrix of order n over the field F.* **(Jabalpur 1970)**

Proof. Let $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$, ...(1)

be a quadratic form over the field F in the n variables $x_1, x_2, \ldots, x_n$.

In (1) it is assumed that $x_ix_j = x_jx_i$. Then the total coefficient of x_ix_j in (1) is $a_{ij}+a_{ji}$. Let us assign half of this coefficient to x_{ij} and half to x_{ji}. Thus we define another set of scalars b_{ij}, such that $b_{ii}=a_{ii}$ and $b_{ij}=b_{ji}=\frac{1}{2}(a_{ij}+a_{ji})$, $i\neq j$. Then we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j=\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j.$$

Let $\mathbf{B}=[b_{ij}]_{n\times n}$. Then $\mathbf{B}$ is a symmetric matrix of order n over the field F since $b_{ij}=b_{ji}$.

Let $\mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ \dots \\ x_n \end{bmatrix}$. Then $\mathbf{X}^T$ or $\mathbf{X}'=[x_1, x_2, \dots, x_n]$.

Now $\mathbf{X}^T\mathbf{B}\mathbf{X}$ is a matrix of the type 1×1. It can be easily seen that the single element of this matrix is $\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j$. If we identify a 1×1 matrix with its single element *i.e.* if we regard a 1×1 matrix equal to its single element, then we have

$$\mathbf{X}^T\mathbf{B}\mathbf{X}=\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}x_ix_j=\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j.$$

Hence the result.

Matrix of a quadratic form. Definition. *If* $\phi=\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$ *is a quadratic form in n variables* $x_1, x_2, \dots, x_n$, *then there exists a unique symmetric matrix* $\mathbf{B}$ *of order n such that* $\phi=\mathbf{X}^T\mathbf{B}\mathbf{X}$ *where* $\mathbf{X}=[x_1, x_2, \dots, x_n]^T$. *The symmetric matrix* $\mathbf{B}$ *is called the matrix of the quadratic form* $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$.

Since every quadratic form can always be so written that matrix of its coefficients is a symmetric matrix, therefore we shall be considering quadratic forms which are so adjusted that the coefficient matrix is symmetric.

Quadratic form corresponding to a symmetric matrix. Let $\mathbf{A}=[a_{ij}]_{n\times n}$ be a symmetric matrix over the field F and let

$$\mathbf{X}=[x_1, x_2, \dots, x_n]^T$$

be a column vector. Then $\mathbf{X}^T\mathbf{A}\mathbf{X}$ determines a unique quadratic form

$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j$ in n variables $x_1, x_2, \dots, x_n$ over the field F.

Thus we have seen that there exists a one-to-one correspondence between the set of all quadratic forms in n variables over a field F and the set of all n-rowed symmetric matrices over F.

Solved Examples

Ex. 1. *Write down the matrix of each of the following quadratic forms and verify that they can be written as matrix products* $\mathbf{X}^T\mathbf{A}\mathbf{X}$:

(*i*) $x_1^2-18x_1x_2+5x_2^2$.

(*ii*) $x_1^2+2x_2^2-5x_3^2-x_1x_2+4x_2x_3-3x_3x_1$.

Solution. (i) The given quadratic form can be written as

$$x_1x_1-9x_1x_2-9x_2x_1+5x_2x_2.$$

Let $\mathbf{A}$ be the matrix of this quadratic form. Then

$$\mathbf{A}=\begin{bmatrix} 1 & -9 \\ -9 & 5 \end{bmatrix}.$$

Let $\mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then $\mathbf{X}'=[x_1 \quad x_2]$.

We have $\mathbf{X}'\mathbf{A}=[x_1 \quad x_2]\begin{bmatrix} 1 & -9 \\ -9 & 5 \end{bmatrix}=[x_1-9x_2 \quad -9x_1+5x_2]$.

$$\therefore \quad \mathbf{X}'\mathbf{A}\mathbf{X}=[x_1-9x_2 \quad -9x_1+5x_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$=x_1\,(x_1-9x_2)+x_2\,(-9x_1+5x_2)$$

$$=x_1^2-9x_1x_2-9x_2x_1+5x_2^2$$

$$=x_1^2-18x_1x_2+5x_2^2.$$

(ii) The given quadratic form can be written as

$$x_1x_1-\tfrac{1}{2}x_1x_2-\tfrac{3}{2}x_1x_3-\tfrac{1}{2}x_2x_1+2x_2x_2+2x_2x_3 -\tfrac{3}{2}x_3x_1+2x_3x_2-5x_3x_3.$$

Let $\mathbf{A}$ be the matrix of this quadratic form. Then

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

Obviously $\mathbf{A}$ is a symmetric matrix.

Let $\mathbf{X}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Then $\mathbf{X}'=[x_1 \quad x_2 \quad x_3]$.

We have $\mathbf{X}'\mathbf{A}=[x_1 \quad x_2 \quad x_3]\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & 2 & 2 \\ -\frac{3}{2} & 2 & -5 \end{bmatrix}$

$$=[x_1-\tfrac{1}{2}x_2-\tfrac{3}{2}x_3 \quad -\tfrac{1}{2}x_1+2x_2+2x_3 \quad -\tfrac{3}{2}x_1+2x_2-5x_3]$$

$\therefore$ **X'AX**

$$=\left[x_1-\tfrac{1}{2}x_2-\tfrac{3}{2}x_3 \quad -\tfrac{1}{2}x_1+2x_2+2x_3 \quad -\tfrac{3}{2}x_1+2x_2-5x_3\right]\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$$

$$=x_1\left(x_1-\tfrac{1}{2}x_2-\tfrac{3}{2}x_3\right)+x_2\left(-\tfrac{1}{2}x_1+2x_2+2x_3\right)+x_3\left(-\tfrac{3}{2}x_1+2x_2-5x_3\right)$$
$$=x_1^2-\tfrac{1}{2}x_1x_2-\tfrac{3}{2}x_1x_3-\tfrac{1}{2}x_2x_1+2x_2^2+2x_2x_3-\tfrac{3}{2}x_3x_1+2x_3x_2-5x_3^2$$
$$=x_1^2+2x_2^2-5x_3^2-x_1x_2+4x_2x_3-3x_3x_1.$$

Ex. 2. *Obtain the matrices corresponding to the following quadratic forms*

(i) $x^2+2y^2+3z^2+4xy+5yz+6zx.$ **(Agra 1970)**

(ii) $ax^2+by^2+cz^2+2fyz+2gzx+2hxy.$

(iii) $a_{11}x_1^2+a_{22}x_2^2+a_{33}x_3^2+2a_{12}x_1x_2+2a_{23}x_2x_3+2a_{31}x_3x_1.$

(iv) $x_1^2-2x_2^2+4x_3^2-4x_4^2-2x_1x_2+3x_1x_4+4x_2x_3-5x_3x_4.$

(v) $x_1x_2+x_2x_3+x_3x_1+x_1x_4+x_2x_4+x_3x_4.$

(vi) $x_1^2-2x_2x_3-x_3x_4.$

(vii) $d_1x_1^2+d_2x_2^2+d_3x_3^2+d_4x_4^2+d_5x_5^2.$

Solution. (i) The given quadratic form can be written as
$$x^2+2xy+3xz+2yx+2y^2+\tfrac{5}{2}yz+3zx+\tfrac{5}{2}zy+3z^2.$$

$\therefore$ if **A** is the matrix of this quadratic form, then

$$\mathbf{A}=\begin{bmatrix}1&2&3\\2&2&\frac{5}{2}\\3&\frac{5}{2}&3\end{bmatrix},$$ which is a symmetric matrix of order 3.

(ii) The given quadratic form can be written as
$$ax^2+hxy+gxz+hyx+by^2+fyz+gzx+fzy+cz^2.$$

$\therefore$ if **A** is the matrix of this quadratic form, then
$$\mathbf{A}=\begin{bmatrix}a&h&g\\h&b&f\\g&f&c\end{bmatrix}$$

(iii) The given quadratic form can be written as
$$a_{11}x_1^2+a_{12}x_1x_2+a_{31}x_1x_3+a_{12}x_2x_1+a_{22}x_2^2$$
$$+a_{23}x_2x_3+a_{31}x_3x_1+a_{23}x_3x_2+a_{33}x_3^2.$$

$\therefore$ if **A** is the matrix of this quadratic form, then
$$\mathbf{A}=\begin{bmatrix}a_{11}&a_{12}&a_{31}\\a_{12}&a_{22}&a_{23}\\a_{31}&a_{23}&a_{33}\end{bmatrix}.$$

(iv) The given quadratic form can be written as
$$x_1^2-x_1x_2+0x_1x_3+\tfrac{3}{2}x_1x_4-x_2x_1-2x_2^2+2x_2x_3+0x_2x_4+0x_3x_1$$
$$+2x_3x_2+4x_3^2-\tfrac{5}{2}x_3x_4+\tfrac{3}{2}x_4x_1+0x_4x_2-\tfrac{5}{2}x_4x_3-4x_4^2.$$

$\therefore$ if **A** is the matrix of this quadratic form, then
$$\mathbf{A}=\begin{bmatrix}1&-1&0&\frac{3}{2}\\-1&-2&2&0\\0&2&4&-\frac{5}{2}\\\frac{3}{2}&0&-\frac{5}{2}&-4\end{bmatrix}.$$

(v) The given quadratic form can be written as
$$0x_1^2+\tfrac{1}{2}x_1x_2+\tfrac{1}{2}x_1x_3+\tfrac{1}{2}x_1x_4+\tfrac{1}{2}x_2x_1+0x_2^2+\tfrac{1}{2}x_2x_3+\tfrac{1}{2}x_2x_4+\tfrac{1}{2}x_3x_1$$
$$+\tfrac{1}{2}x_3x_2+0x_3^2+\tfrac{1}{2}x_3x_4+\tfrac{1}{2}x_4x_1+\tfrac{1}{2}x_4x_2+\tfrac{1}{2}x_4x_3+0x_4^2.$$

$\therefore$ if **A** is the matrix of this quadratic form, then
$$\mathbf{A}=\begin{bmatrix}0&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&0&\frac{1}{2}\\\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&0\end{bmatrix}$$

(vi) The given quadratic form can be written as
$$x_1^2+0x_1x_2+0x_1x_3+0x_1x_4+0x_2x_1+0x_2^2-x_2x_3+0x_2x_4+0x_3x_1-x_3x_2$$
$$+0x_3^2-\tfrac{1}{2}x_3x_4+0x_4x_1+0x_4x_2-\tfrac{1}{2}x_4x_3+0x_4^2.$$

$\therefore$ if **A** is the matrix of the given quadratic form then
$$\mathbf{A}=\begin{bmatrix}1&0&0&0\\0&0&-1&0\\0&-1&0&-\frac{1}{2}\\0&0&-\frac{1}{2}&0\end{bmatrix}.$$

(vii) If **A** is the matrix of the given quadratic form, then obviously **A** is a diagonal matrix and $\mathbf{A}=\text{diag.}\,[d_1, d_2, d_3, d_4, d_5]$.

Ex. 3. *Write down the quadratic forms corresponding to the following symmetric matrices :*

(i) $\begin{bmatrix}1&2&3\\2&0&3\\3&3&1\end{bmatrix}$ (ii) diag. $[\lambda_1, \lambda_2, \ldots, \lambda_n]$.

(Agra 1970)

Solution. (i) Let $\mathbf{X}=[x_1\ x_2\ x_3]^T$ and **A** denote the given symmetric matrix. Then $\mathbf{X}^T\mathbf{AX}$ is the quadratic form corresponding to this matrix. We have

$$\mathbf{X}^T\mathbf{A}=[x_1\ x_2\ x_3]\begin{bmatrix}1&2&3\\2&0&3\\3&3&1\end{bmatrix}$$
$$=[x_1+2x_2+3x_3 \quad 2x_1+3x_3 \quad 3x_1+3x_2+x_3].$$

$$\therefore\ \mathbf{X}^T\mathbf{AX}=x_1(x_1+2x_2+3x_3)+x_2(2x_1+3x_3)+x_3(3x_1+3x_2+x_3)$$
$$=x_1^2+x_3^2+4x_1x_2+6x_1x_3+6x_2x_3.$$

(ii) The required quadratic form is
$$\lambda_1x_1^2+\lambda_2x_2^2+\ldots+\lambda_nx_n^2$$

Exercises

1. Obtain the matrices corresponding to the following quadratic forms

   (i) $ax^2+2hxy+by^2$. (ii) $x_1^2+5x_2^2-7x_3^2$.

   (iii) $2x_1x_2+6x_1x_3-4x_2x_3$.

2. Write down the quadratic forms corresponding to the following matrices

   (i) $\begin{bmatrix} 2 & 1 & 5 \\ 1 & 3 & -2 \\ 5 & -2 & 4 \end{bmatrix}$ (ii) $\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{bmatrix}$

   (iii) $\begin{bmatrix} 0 & a & b & c \\ a & 0 & l & m \\ b & l & 0 & p \\ c & m & p & 0 \end{bmatrix}$.

Answers

1. (i) $\begin{bmatrix} a & h \\ h & b \end{bmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -7 \end{bmatrix}$ (iii) $\begin{bmatrix} 0 & 1 & 3 \\ 1 & 0 & -2 \\ 3 & -2 & 0 \end{bmatrix}$.

2. (i) $2x_1^2+3x_2^2+4x_3^2+2x_1x_2+10x_1x_3-4x_2x_3$

   (ii) $2x_2^2+4x_3^2+6x_4^2+2x_1x_2+4x_1x_3+6x_1x_4$
   $+6x_2x_3+8x_2x_4+10x_3x_4$

   (iii) $2ax_1x_2+2bx_1x_3+2cx_1x_4+2lx_2x_3+2mx_2x_4+2px_3x_4$.

§ 2. Linear Transformations. Suppose V_n is the vector space of all ordered n-tuples of the elements of a field F and let the vectors in V_n be written as column vectors. Let **P** be a matrix of order n over the field F. If $\mathbf{Y}=[y_1, y_2, \ldots, y_n]^T$ is a vector in V_n, then **Y** is a matrix of the type $n \times 1$. Obviously **PY** is a matrix of the type $n \times 1$. Thus **PY** is also a vector in V_n. Let

$$\mathbf{PY}=\mathbf{X}=[x_1, x_2, \ldots, x_n]^T.$$

The relation $\mathbf{PY}=\mathbf{X}$ thus gives a mapping from V_n into V_n.

Since $\mathbf{P}(a\mathbf{Y}_1+b\mathbf{Y}_2)=a(\mathbf{PY}_1)+b(\mathbf{PY}_2)$, therefore this mapping is a linear transformation. If the matrix **P** is non-singular, then the linear transformation is also said to be non-singular. Also if the matrix **P** is non-singular, then the mapping $\mathbf{PY}=\mathbf{X}$ is one-one onto as shown below :

Mapping P is one-one. We have $\mathbf{PY}_1=\mathbf{PY}_2$

$$\Rightarrow \mathbf{P}^{-1}(\mathbf{PY}_1)=\mathbf{P}^{-1}(\mathbf{PY}_2) \Rightarrow \mathbf{Y}_1=\mathbf{Y}_2.$$

Therefore the mapping **P** is one-one.

Mapping P is onto. Let **Z** be any vector in V_n. Then $\mathbf{P}^{-1}\mathbf{Z}$ is also a vector of V_n and we have $\mathbf{P}(\mathbf{P}^{-1}\mathbf{Z})=(\mathbf{PP}^{-1})\mathbf{Z}=\mathbf{IZ}=\mathbf{Z}$. Therefore the mapping **P** is onto.

If the linear transformation $\mathbf{PY}=\mathbf{X}$ is non-singular, then $\mathbf{PY}=\mathbf{O}$ if and only if $\mathbf{Y}=\mathbf{O}$. If $\mathbf{Y}=\mathbf{O}$, then obviously $\mathbf{PY}=\mathbf{O}$. Conversely, $\mathbf{PY}=\mathbf{O} \Rightarrow \mathbf{P}^{-1}(\mathbf{PY})=\mathbf{P}^{-1}\mathbf{O} \Rightarrow \mathbf{Y}=\mathbf{O}$.

§ 3. Congruence of matrices. Definition. *A square matrix* **B** *of order n over a field F is said to be congruent to another square matrix* **A** *of order n over F, if there exists a non-singular matrix* **P** *over F such that*

$$\mathbf{B}=\mathbf{P}^T\mathbf{AP}.$$

Theorem 1. *The relation of 'congruence of matrices' is an equivalence relation in the set of all $n \times n$ matrices over a field F.* **(Punjab 1967)**

Proof. Reflexivity. Let **A** be any $n \times n$ matrix over a field F. Then $\mathbf{A}=\mathbf{I}^T\mathbf{AI}$, where **I** is unit matrix of order n over F. Since **I** is non-singular, therefore **A** is congruent to itself.

Symmetry. Suppose **A** is congruent to **B**. Then $\mathbf{A}=\mathbf{P}'\mathbf{BP}$, where **P** is non-singular.

$$\therefore \quad (\mathbf{P}')^{-1}\mathbf{AP}^{-1}=(\mathbf{P}')^{-1}\mathbf{P}'\mathbf{BPP}^{-1}=\mathbf{B}$$

$$\Rightarrow (\mathbf{P}^{-1})'\mathbf{AP}^{-1}=\mathbf{B} \qquad [\because \ (\mathbf{P}')^{-1}=(\mathbf{P}^{-1})']$$

$\Rightarrow$ **B** is congruent to **A**.

Transitivity. Suppose **A** is congruent to **B** and **B** is congruent to **C**. Then $\mathbf{A}=\mathbf{P}'\mathbf{BP}$, $\mathbf{B}=\mathbf{Q}'\mathbf{CQ}$, where **P** and **Q** are non-singular. Therefore $\mathbf{A}=\mathbf{P}'(\mathbf{Q}'\mathbf{CQ})\mathbf{P}=(\mathbf{P}'\mathbf{Q}')\mathbf{CQP}=(\mathbf{QP})'\mathbf{C}(\mathbf{QP})$. Since **QP** is also a non-singular matrix, therefore **A** is congruent to **C**.

Thus the relation of 'congruence of matrices' is reflexive, symmetric and transitive. So it is an equivalence relation.

Theorem 2. *Every matrix congruent to a symmetric matrix is a symmetric matrix.*

Proof. Let a matrix **B** be congruent to a symmetric matrix **A**. Then there exists a non-singular matrix **P** such that $\mathbf{B}=\mathbf{P}'\mathbf{AP}$. We have

$$\mathbf{B}'=(\mathbf{P}'\mathbf{AP})'=\mathbf{P}'\mathbf{A}'(\mathbf{P}')'=\mathbf{P}'\mathbf{A}'\mathbf{P}$$

$$=\mathbf{P}'\mathbf{AP} \qquad [\because \ \mathbf{A} \text{ is symmetric} \Rightarrow \mathbf{A}'=\mathbf{A}]$$

$$=\mathbf{B}.$$

$\therefore$ **B** is also a symmetric matrix.

Congruence operation on a square matrix or Congruence transformations of a square matrix.

A congruence operation of a square matrix is an operation of any one of the following three types :

(i) Interchange of the ith and the jth row as well as of the ith and the jth column. Both should be applied simultaneously. Thus the operation $R_i \leftrightarrow R_j$ followed by $C_i \leftrightarrow C_j$ is a congruence operation.

(ii) Multiplication of the ith row as well as the ith column by a non-zero number c *i.e.* $R_i \to cR_i$ followed by $C_i \to cC_i$.

(iii) $R_i \to R_i + kR_j$ followed by $C_i \to C_i + kC_j$.

Now we shall show that *each congruent transformation of a matrix consists of a pair of elementary transformations, one row and the other column, such that of the corresponding elementary matrices each is the transpose of the other.*

(*a*) Let $\mathbf{E}^*, \mathbf{E}$ be the elementary matrices corresponding to the elementary transformations $R_i \leftrightarrow R_j$ and $C_i \leftrightarrow C_j$ respectively. Then $\mathbf{E}^* = \mathbf{E} = \mathbf{E}'$.

(*b*) Let $\mathbf{E}^*, \mathbf{E}$ be the elementary matrices corresponding to the elementary transformations $R_i \to cR_i$ and $C_i \to cC_i$ respectively where $c \neq 0$. Then $\mathbf{E}^* = \mathbf{E} = \mathbf{E}'$.

(*c*) Let $\mathbf{E}^*, \mathbf{E}$ be the elementary matrices corresponding to the elementary transformations $R_i \to R_i + kR_j$ and $C_i \to C_i + kC_j$ respectively. Then $\mathbf{E}^* = \mathbf{E}'$.

Now we know that every elementary row (column) transformation of a matrix can be brought about by pre-multiplication (post-multiplication) with the corresponding elementary matrix. Therefore if a matrix **B** has been obtained from **A** by a finite chain of congruent operations applied on **A**, then there exist elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_s$ such that

$$\mathbf{B} = \mathbf{E}'_s \ldots \mathbf{E}'_2 \mathbf{E}'_1 \mathbf{A} \mathbf{E}_1 \mathbf{E}_2 \ldots \mathbf{E}_s$$
$$= (\mathbf{E}_1 \mathbf{E}_2 \ldots \mathbf{E}_s)' \mathbf{A} (\mathbf{E}_1 \mathbf{E}_2 \ldots \mathbf{E}_s)$$
$$= \mathbf{P}'\mathbf{A}\mathbf{P}, \text{ where } \mathbf{P} = \mathbf{E}_1 \mathbf{E}_2 \ldots \mathbf{E}_s \text{ is a non-singular matrix.}$$

Therefore **B** is congruent to **A**. *Thus every matrix* **B** *obtained from any given matrix* **A** *by subjecting* **A** *to a finite chain of congruent operations is congruent to* **B**.

The converse is also true. If **B** is congruent to **A**, then

$$\mathbf{B} = \mathbf{P}' \mathbf{A} \mathbf{P}$$

where **P** is a non-singular matrix. Now every non-singular matrix can be expressed as the product of elementary matrices. Therefore we can write $\mathbf{P} = \mathbf{E}_1 \mathbf{E}_2 \ldots \mathbf{E}_s$ where $\mathbf{E}_1, \ldots, \mathbf{E}_s$ are elementary matrices. Then $\mathbf{B} = \mathbf{E}'_s \ldots \mathbf{E}'_2 \mathbf{E}'_1 \mathbf{A} \mathbf{E}_1 \mathbf{E}_2 \ldots \mathbf{E}_s$. Therefore **B** is obtained from **A** by a finite chain of congruent operations applied on **A**.

§ 4. Congruence of Quadratic Forms or Equivalence of Quadratic Forms.

Definition. *Two quadratic forms* $\mathbf{X}^T \mathbf{A} \mathbf{X}$ *and* $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ *over a field F are said to be congruent or equivalent over F if their respective matrices* **A** *and* **B** *are congruent over F. Thus* $\mathbf{X}^T \mathbf{A} \mathbf{X}$ *is equivalent to* $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ *if there exists a non-singular matrix* **P** *over F such that*

$$\mathbf{P}^T \mathbf{A} \mathbf{P} = \mathbf{B}.$$

Since congruence of matrices is an equivalence relation, therefore equivalence of quadratic forms is also an equivalence relation.

Equivalence of Real Quadratic Forms.

Definition. *Two real quadratic forms* $\mathbf{X}^T \mathbf{A} \mathbf{X}$ *and* $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ *are said to be real equivalent, orthogonally equivalent, or complex equivalent according as there exists a non-singular real, orthogonal, or a non-singular complex matrix* **P** *such that*

$$\mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P}.$$

§ 5. The linear transformation of a quadratic form.

Consider a quadratic form

$$\mathbf{X}^T \mathbf{A} \mathbf{X} \qquad \ldots(1)$$

and a non-singular linear transformation

$$\mathbf{X} = \mathbf{P} \mathbf{Y} \qquad \ldots(2)$$

so that **P** is a non-singular matrix.

Putting $\mathbf{X} = \mathbf{P}\mathbf{Y}$ in (1), we get

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = (\mathbf{P}\mathbf{Y})^T \mathbf{A} (\mathbf{P}\mathbf{Y}) = \mathbf{Y}^T \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{Y}$$
$$= \mathbf{Y}^T \mathbf{B} \mathbf{Y}, \text{ where } \mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P}.$$

Since **B** is congruent to a symmetric matrix **A**, therefore **B** is also a symmetric matrix. Thus $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ is a quadratic form. It is called a linear transform of the form $\mathbf{X}^T \mathbf{A} \mathbf{X}$ by the non-singular matrix **P**. The matrix of the quadratic form $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ is

$$\mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P}.$$

Thus the quadratic form $\mathbf{Y}^T \mathbf{B} \mathbf{Y}$ is congruent to $\mathbf{X}^T \mathbf{A} \mathbf{X}$.

Theorem. *The ranges of values of two congruent quadratic forms are the same.*

Proof. Let $\phi = \mathbf{X}'\mathbf{A}\mathbf{X}$ and $\psi = \mathbf{Y}'\mathbf{B}\mathbf{Y}$ be two congruent quadratic forms. Then there exists a non-singular matrix **P** such that

$$\mathbf{B} = \mathbf{P}'\mathbf{A}\mathbf{P}.$$

Consider the linear transformation $\mathbf{X} = \mathbf{P}\mathbf{Y}$.

Let $\phi = p$ when $\mathbf{X} = \mathbf{X}_1$. Then $p = \mathbf{X}_1'\mathbf{A}\mathbf{X}_1$. The value of ψ when $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}_1$ is

$=(P^{-1}X_1)' \, B \, (P^{-1}X_1)=X_1' \, (P^{-1})' \, P'APP^{-1}X_1$
$=X_1' \, (P')^{-1} \, P' \, AX_1=X_1'AX_1=p.$

Thus each value of ϕ is equal to some value of ψ.

Conversely let $\psi=q$ when $Y=Y_1$. Then $q=Y_1'BY_1$. The value of ϕ when $X=PY_1$ is

$$=(PY_1)' \, A \, (PY_1)=Y_1'P' \, APY_1$$
$$=Y_1'BY_1=q.$$

Thus each value of ψ is equal to some value of ϕ.

Hence ϕ and ψ have the same ranges of values.

Corollary. *If the quadratic form* $Y'BY$ *is a linear transform of the quadratic form* $X'AX$ *by a non-singular matrix* P, *then the two forms are congruent and so have the same ranges of values.*

§ 6. Congruent reduction of a symmetric matrix.

Theorem. *If* A *be any n-rowed non-zero symmetric matrix of rank r, over a field F, then there exists an n-rowed non-singular matrix* P *over F such that*

$$P'AP=\begin{bmatrix} A_1 & O \\ O & O \end{bmatrix}$$

where, A_1 *is a non-singular diagonal matrix of order r over F and each* O *is a null matrix of suitable size.*

Or

Every symmetric matrix of rank r is congruent to a diagonal matrix, r, of whose diagonal elements only are non-zero.

Proof. We shall prove the theorem by induction on n, the order of the given matrix. If $n=1$, the theorem is obviously true. Let us suppose that the theorem is true for all symmetric matrices of order $n-1$. Then we shall show that it is also true for an $n \times n$ symmetric matrix A.

Let $A=\left[a_{ij} \right]_{n\times n}$ be a symmetric matrix of rank r over a field F. First we shall show that there exists a matrix $B=\left[b_{ij} \right]_{n\times n}$ over F congruent to A such that $b_{11}\neq 0$.

Case 1. If $a_{11}\neq 0$, then we take $B=A$.

Case 2. If $a_{11}=0$, but some diagonal element of A, say, $a_{ii}\neq 0$, then applying the congruent operation $R_i \leftrightarrow R_1$, $C_i \leftrightarrow C_1$ to A, we obtain a matrix B congruent to A such that

$$b_{11}=a_{ii}\neq 0.$$

Case 3. Suppose that each diagonal element of A is 0. Since

A is a non-zero matrix, let a_{ij} be a non-zero element of A. Then

$$a_{ij}=a_{ji}\neq 0.$$

Applying the congruent operation $R_i \to R_i+R_j$, $C_i \to C_i+C_j$ to A, we obtain a matrix $D=\left[d_{ij} \right]_{n\times n}$ congruent to A such that

$$d_{ii}=a_{ij}+a_{ji}=2a_{ij}\neq 0.$$

Now applying the congruent operation $R_i \leftrightarrow R_1$, $C_i \leftrightarrow C_1$ to D we obtain a matrix

$$B=\left[b_{ij} \right]_{n\times n}$$

congruent to D and, therefore, also congruent to A such that

$$b_{11}=d_{ii}\neq 0.$$

Thus there always exists a matrix

$$B=\left[b_{ij} \right]_{n\times n}$$

congruent to a symmetric matrix, such that the leading element of B is not zero. Since B is congruent to a symmetric matrix, therefore B itself is a symmetric matrix. Since $b_{11}\neq 0$, therefore all elements in the first row and first column of B, except the leading element, can be made 0 by suitable congruent operations. We thus have a matrix

$$C=\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B_1 & \\ 0 & & & \end{bmatrix}$$

congruent to B and, therefore, also congruent to A such that B_1 is a square matrix of order $n-1$. Since C is congruent to a symmetric matrix A, therefore C is also a symmetric matrix and consequently B_1 is also a symmetric matrix. Thus B_1 is a symmetric matrix of order $n-1$. Therefore by our induction hypothesis it can be reduced to a diagonal matrix by congruent operations. If the congruent operations applied to B_1 for this purpose be applied to C, they will not affect the first row and the first column of C. So C can be reduced to a diagonal matrix by congruent operations. Thus A is congruent to a diagonal matrix, say diag $[\lambda_1, \lambda_2, \ldots, \lambda_k, 0, 0, \ldots, 0]$. Thus there exists a non-singular matrix P such that

$$P'AP=\text{diag } [\lambda_1, \ldots, \lambda_k, 0, \ldots, 0].$$

Since rank $A=r$ and the rank of a matrix does not change on multiplication by a non-singular matrix, therefore rank of the matrix $P'AP = \text{diag. } [\lambda_1, \ldots, \lambda_k, 0, \ldots, 0]$ is also r. So precisely r elements of diag. $[\lambda_1, \ldots, \lambda_k, 0, \ldots, 0]$ are non-zero. Therefore $k=r$ and thus $P'AP = \text{diag. } [\lambda_1, \ldots, \lambda_r, 0, \ldots, 0]$.

Thus **A** can be reduced to diagonal form by congruent operations.

The proof is now complete by induction.

Corollary. *Corresponding to every quadratic form* $X'AX$ *over a field F, there exists a non-singular linear transformation*

$$X = PY$$

over F, such that the form $X'AX$ *transforms to a sum of r, square terms*

$$\lambda_1 y_1^2 + \ldots + \lambda_r y_r^2,$$

where $\lambda_1, \ldots, \lambda_r$ *belong to the field F and r is the rank of the matrix* **A**.

Rank of a quadratic form. Definition. *Let* $X'AX$ *be a quadratic form over a field F. The rank of the matrix* **A** *is called the rank of the quadratic form* $X'AX$. **(Nagarjuna 1980)**

If $X'AX$ is a quadratic form of rank r, then there exists a non-singular matrix **P** which will reduce the form $X'AX$ to a sum of r square terms.

Working Rule for Numerical problems.

We should transform the given symmetric matrix **A** to diagonal form by applying congruent operations. Then the application of corresponding column operations to the unit matrix I_n will give us a non-singular matrix **P** such that

$$P'AP = \text{a diagonal matrix.}$$

The whole process will be clear from the following examples.

Ex. 1. *Determine a non-singular matrix* **P** *such that* $P'AP$ *is a diagonal matrix, where*

$$A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}.$$

Interpet the result in terms of quadratic forms.

Solution. We write $A = IAI$

i.e. $$\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We shall reduce **A** to diagonal form by applying congruent operations. On the right hand side **IAI**, the corresponding row operations will be applied on prefactor **I** and the column operations will be applied on the post-factor **I**. There is no need of actually applying row operations on pre-factor **I** because at any stage the matrix obtained by applying row operations on pre-factor **I** will be the transpose of the matrix obtained by applying column operations on post-factor **I**.

Performing the congruent operations

$R_2 \to R_2 + \frac{1}{3}R_1$, $C_2 \to C_2 + \frac{1}{3}C_1$ and $R_3 \to R_3 - \frac{1}{3}R_1$, $C_3 \to C_3 - \frac{1}{3}C_1$,

we have

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & \frac{7}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{7}{3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -\frac{1}{3} & 0 & 1 \end{bmatrix} A \begin{bmatrix} 1 & \frac{1}{3} & -\frac{1}{3} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now performing the congruent operation

$R_3 \to R_3 + \frac{1}{7}R_2$, $C_3 \to C_3 + \frac{1}{7}C_2$, we get

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & \frac{7}{3} & 0 \\ 0 & 0 & \frac{16}{7} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{3} & 1 & 0 \\ -\frac{2}{7} & \frac{1}{7} & 1 \end{bmatrix} A \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus we obtain a non-singular matrix

$$P = \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix} \text{ such that } P'AP = \text{diag. } [6, \tfrac{7}{3}, \tfrac{16}{7}].$$

The quadratic form corresponding to the matrix **A** is

$$X'AX = 6x_1^2 + 3x_2^2 + 3x_3^2 - 4x_1x_2 - 2x_2x_3 + x_3x_1. \qquad \ldots(1)$$

The non-singular transformation corresponding to the matrix **P** is given by $X = PY$ *i.e.*

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{3} & -\frac{2}{7} \\ 0 & 1 & \frac{1}{7} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

which is equivalent to

$$\left.\begin{aligned} x_1 &= y_1 + \tfrac{1}{3}y_2 - \tfrac{2}{7}y_3 \\ x_2 &= y_2 + \tfrac{1}{7}y_3 \\ x_3 &= y_3. \end{aligned}\right\} \qquad \ldots(2)$$

The transformation (2) will reduce the quadratic form (1) to the diagonal form $Y'P'APY = 6y_1^2 + \frac{7}{3}y_2^2 + \frac{16}{7}y_3^2$.

The rank of the quadratic form $X'AX$ is 3. So it has been reduced to a form which is a sum of three squares.

Ex. 2. *Determine a non-singular matrix* **P** *such that* $P'AP$ *is a diagonal matrix, where*

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

Solution. We have

$$\begin{bmatrix}0 & 1 & 2\\ 1 & 0 & 3\\ 2 & 3 & 0\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing the row operation $R_1 \to R_1+R_2$, we get

$$\begin{bmatrix}1 & 1 & 5\\ 1 & 0 & 3\\ 2 & 3 & 0\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Now performing the corresponding column operation $C_1 \to C_1+C_2$, we get

$$\begin{bmatrix}2 & 1 & 5\\ 1 & 0 & 3\\ 5 & 3 & 0\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing the row-operations $R_2 \to R_2-\frac{1}{2}R_1$, $R_3 \to R_3-\frac{5}{2}R_1$, we get

$$\begin{bmatrix}2 & 1 & 5\\ 0 & -\frac{1}{2} & \frac{1}{2}\\ 0 & \frac{1}{2} & -\frac{25}{2}\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\ -\frac{1}{2} & \frac{1}{2} & 0\\ -\frac{5}{2} & -\frac{5}{2} & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Now performing the corresponding column operations, $C_2 \to C_2-\frac{1}{2}C_1$, $C_3 \to C_3-\frac{5}{2}C_1$, we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & -\frac{1}{2} & \frac{1}{2}\\ 0 & \frac{1}{2} & -\frac{25}{2}\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\ -\frac{1}{2} & \frac{1}{2} & 0\\ -\frac{5}{2} & -\frac{5}{2} & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & -\frac{1}{2} & -\frac{5}{2}\\ 1 & \frac{1}{2} & -\frac{5}{2}\\ 0 & 0 & 1\end{bmatrix}.$$

Performing the row operation $R_3 \to R_3+R_1$ and then the column operation $C_3 \to C_3+C_1$ we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & -\frac{1}{2} & 0\\ 0 & 0 & -12\end{bmatrix}=\begin{bmatrix}1 & 1 & 0\\ -\frac{1}{2} & \frac{1}{2} & 0\\ -3 & -2 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & -\frac{1}{2} & -3\\ 1 & \frac{1}{2} & -2\\ 0 & 0 & 1\end{bmatrix}.$$

$$\therefore\ \mathbf{P}=\begin{bmatrix}1 & -\frac{1}{2} & -3\\ 1 & \frac{1}{2} & -2\\ 0 & 0 & 1\end{bmatrix}.$$

§ 7. Reduction of a real quadratic form.

Theorem 1. *If* **A** *be any n-rowed real symmetric matrix of rank r, then there exists a real non-singular matrix* **P** *such that*

$$\mathbf{P'AP}=diag\ [1, 1, ..., 1, -1, -1, ..., -1, 0, ..., 0]$$

so that 1, *appears p times and,* -1, *appears* $r-p$ *times.*

Proof. A is a real symmetric matrix of rank r. Therefore there exists a non-singular real matrix **Q** such that **Q'AQ** is a diagonal matrix **D** with precisely r non-zero diagonal elements. Let

$$\mathbf{Q'AQ}=\mathbf{D}=\text{diag}\ [\lambda_1, \lambda_2, ..., \lambda_r, 0, ..., 0].$$

Suppose that p of the non-zero diagonal elements are positive. Then $r-p$ are negative.

Since in a diagonal matrix the positions of the diagonal elements occurring in i^{th} and j^{th} rows can be interchanged by applying the congruent operation $R_i \leftrightarrow R_j$, $C_i \leftrightarrow C_j$, therefore without any loss of generality we can take $\lambda_1, ..., \lambda_p$ to be positive and $\lambda_{p+1}, ..., \lambda_r$ to be negative.

Let S be the $n \times n$ (real) diagonal matrix with diagonal elements

$$\frac{1}{\sqrt{\lambda_1}}, ..., \frac{1}{\sqrt{\lambda_p}}, \frac{1}{\sqrt{(-\lambda_{p+1})}}, ..., \frac{1}{\sqrt{(-\lambda_r)}}, 1, ..., 1.$$

Then $\mathbf{S}=\text{diag}\left[\frac{1}{\sqrt{\lambda_1}}, ..., \frac{1}{\sqrt{\lambda_p}}, \frac{1}{\sqrt{(-\lambda_{p+1})}}, ..., \frac{1}{\sqrt{(-\lambda_r)}}, 1, ..., 1\right]$

is a real non-singular diagonal matrix and $\mathbf{S'}=\mathbf{S}$.

If we take $\mathbf{P}=\mathbf{QS}$, then **P** is also real non-singular matrix and we have

$$\mathbf{P'AP}=\mathbf{(QS)'\ A\ (QS)}=\mathbf{S'Q'AQS}=\mathbf{S'DS}=\mathbf{SDS}$$
$$=\text{diag}\ [1, ..., 1, -1, ..., -1, 0, ..., 0]$$

so that 1 and -1 appear p and $r-p$ times respectively.

Corollary. *If* **X'AX** *is a real quadratic form of rank r in n variables, then there exists a real non-singular linear transformation* **X=PY** *which transforms* **X'AX** *to the form*

$$\mathbf{Y'P'APY}=y_1^2+...+y_p^2-y_{p+1}^2-...-y_r^2.$$

Canonical or Normal form of a real quadratic form. Definition. *If* **X'AX** *is a real quadratic form in n variables, then there exists real non-singular linear transformation* **X=PY** *which transforms* **X'AX** *to the form*

$$y_1^2+...+y_p^2-y_{p+1}^2-...-y_r^2.$$

In the new form the given quadratic form has been expressed as a sum and difference of the squares of new variables. This latter expression is called the canonical form or normal form of the given quadratic form.

(Nagarjuna 1980)

If $\phi = X'AX$ is a real quadratic form of rank r, then **A** is a matrix of rank r. If the real non-singular linear transformation $X = PY$ reduces ϕ to normal form, then $P'AP$ is a diagonal matrix having 1 and -1 as its non-zero diagonal elements. Since $P'AP$ is also of rank r, therefore it will have precisely r non-zero diagonal elements. Thus the number of terms in each normal form of a given real quadratic form is the same. Now we shall prove that the number of positive terms in any two normal reductions of a real quadratic form is the same.

Theorem 2. *The number of positive terms in any two normal reductions of a real quadratic form is the same.* **(Banaras 1968)**

Proof. Let $\phi = X'AX$ be a real quadratic form of rank r in n variables. Suppose the real non-singular linear transformations

$$X = PY \text{ and } X = QZ$$

transform ϕ to the normal forms

$$y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2 \qquad \ldots(1)$$

and $$z_1^2 + \ldots + z_q^2 - z_{q+1}^2 - \ldots - z_r^2 \qquad \ldots(2)$$

respectively.

To prove that $p = q$.

Let $p < q$. Obviously $y_1, \ldots, y_n, z_1, \ldots, z_n$ are linear homogeneous functions of $x_1, \ldots, x_n$. Since $q > p$, therefore $q - p > 0$. So $n - (q - p)$ is less than n. Therefore $(n - q) + p$ is less than n.

Now $y_1 = 0, y_2 = 0, \ldots, y_p = 0, z_{q+1} = 0, z_{q+2} = 0, \ldots, z_n = 0$ are $(n - q) + p$ linear homogeneous equations in n unknowns $x_1, \ldots, x_n$. Since the number of equations is less than the number of unknowns n, therefore these equations must possess a non-zero solution. Let $x_1 = a_1, \ldots, x_n = a_n$ be a non-zero solution of these equations and let $X_1 = [a_1, \ldots, a_n]'$. Let $Y = [b_1, \ldots, b_n]' = Y_1$ and $Z = [c_1, \ldots, c_n]'$ when $X = X_1$. Then $b_1 = 0, \ldots, b_p = 0$ and $c_{q+1} = 0, \ldots, c_n = 0$. Putting $Y = [b_1, \ldots, b_n]'$ in (1) and $Z = [c_1, \ldots, c_n]'$ in (2), we get two values of ϕ when $X = X_1$.

These must be equal. Therefore we have

$$-b_{p+1}^2 - \ldots - b_r^2 = c_1^2 + \ldots + c_q^2$$

$$\Rightarrow b_{p+1} = 0, \ldots, b_r = 0$$

$$\Rightarrow Y_1 = O$$

$$\Rightarrow P^{-1} X_1 = O \qquad [\because X_1 = PY_1]$$

$$\Rightarrow X_1 = O,$$

which is a contradiction since X_1 is a non-zero vector.

Thus we cannot have $p < q$. Similarly, we cannot have $q < p$. Hence we must have $p = q$.

Corollary. *The number of negative terms in any two normal reductions of a real quadratic form is the same. Also the excess of the number of positive terms over the number of negative terms in any two normal reductions of a real quadratic form is the same.*

Signature andjindex of a real quadratic form.

Definition. *Let* $y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2$ *be a normal form of a real quadratic form* $X'AX$ *of rank* r. *The number* p *of positive terms in a normal form of* $X'AX$ *is called the index of the quadratic form. The excess of the number of positive terms over the number of negative terms in a normal form of* $X'AX$ *i.e.,* $p - (r - p) = 2p - r$ *is called the signature of the quadratic form and is usually denoted by* s. **(Nagarjuna 1980)**

Thus $s = 2p - r$.

In terms of signature theorem 2 may be stated as follows :

Theorem 3. Sylvester's Law of Inertia. *The signature of a real quadratic form is invariant for all normal reductions.* **(Nagarjuna 1990, Punjab 71)**

For its proof give definition of signature and the proof of theorem 2.

Theorem 4. *Two real quadratic forms in n variables are real equivalent if and only if they have the same rank and index (or signature).* **(Nagarjuna 1990)**

Proof. Suppose $X'AX$ and $Y'BY$ are two real quadratic forms in the same number of variables.

Let us first assume that the two forms are equivalent. Then there exists a real non-singular linear transformation $X = PY$ which transforms $X'AX$ to $Y'BY$ *i.e.* $B = P'AP$.

Now suppose the real non-singular linear transformation $Y = QZ$ transforms $Y'BY$ to normal form $Z'CZ$. Then $C = Q'BQ$. Since P and Q are real non-singular matrices, therefore PQ is also a real non-singular matrix. The linear transformation $X = (PQ)Z$ will transform $X'AX$ to the form

$$(PQZ)' \, A \, (PQZ) = Z'Q'P' \, APQZ = Z'Q'BQZ = Z'CZ.$$

Thus the two given quadratic forms have a common normal form. Hence they have the same rank and same index (or signature).

Conversely, suppose that the two forms have the same rank r and the same signature s. Then they have the same index p where $2p - r = s$. So they can be reduced to the same normal form

$$Z'CZ = z_1^2 + \ldots + z_p^2 - z_{p+1}^2 - \ldots - z_r^2$$

by real non-singular linear transformations, say, $X = PZ$ and $Y = QZ$ respectively. Then $P'AP = C$ and $Q'BQ = C$.

Therefore $Q'BQ = P'AP$. This gives $B = (Q')^{-1} P'APQ^{-1} = (Q^{-1})' P'APQ^{-1} = (PQ^{-1})' A (PQ^{-1})$. Therefore the real non-singular linear transformation $X = (PQ^{-1}) Y$ transforms $X'AX$ to $Y'BY$. Hence the two given quadratic forms are real equivalent.

Reduction of a real quadratic form in the complex field.

Theorem 5. *If* A *be any n-rowed real symmetric matrix of rank r, there exists a non-singular matrix* **P** *whose elements may be any complex numbers such that*

$P'AP = \text{diag } [1, 1, \ldots, 1, 0, \ldots, 0]$ *where* 1, *appears r times.*

Proof. A is a real symmetric matrix of rank r. Therefore there exists a non-singular real matrix Q such that $Q'AQ$ is a diagonal matrix D with precisely r non-zero diagonal elements. Let

$$Q'AQ = D = \text{diag. } [\lambda_1, \ldots, \lambda_r, 0, \ldots, 0].$$

The real numbers $\lambda_1, \ldots, \lambda_r$ may be positive or negative or both.

Let S be the $n \times n$ (complex) diagonal matrix with diagonal elements $\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_r}}, 1, \ldots, 1$. Then $S = \text{diag}\left[\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_r}}, 1, \ldots, 1\right]$ is a complex non-singular diagonal matrix and $S' = S$.

If we take $P = QS$, then P is also a complex non-singular matrix and we have $P'AP = (QS)' A (QS) = S'Q'AQS = S'DS = SDS = \text{diag } [1, 1, \ldots, 1, 0, \ldots, 0]$ so that 1 appears r times. Hence the result.

Corollary 1. *Every real quadratic form* **X'AX** *is complex-equivalent to the form* $z_1^2 + \ldots + z_r^2$ *where r is the rank of* **A**.

Corollary 2. *Two real quadratic forms in n variables are complex equivalent if and only if they have the same rank.*

Orthogonal reduction of a real quadratic form.

Theorem 6. *If* $\phi = X'AX$ *be a real quadratic form of rank r in n variables, then there exists a real orthogonal transformation* $X = PY$ *which transforms* ϕ *to the form*

$$\lambda_1 y_1^2 + \ldots + \lambda_r y_r^2$$

where $\lambda_1, \ldots, \lambda_r$ *are the, r, non-zero eigenvalues of* A, $n-r$ *eigenvalues of* A *being equal to zero.*

Proof. Since A is a real symmetric matrix, therefore there exists a real orthogonal matrix P such that

$$P^{-1}AP = D,$$

where D is a diagonal matrix whose diagonal elements are the eigenvalues of A.

Since A is of rank r, therefore $P^{-1}AP = D$ is also a rank r. So D has precisely r non-zero diagonal elements. Consequently A has exactly r non-zero eigenvalues, the remaining $n-r$ eigenvalues of A being zero. Let $D = \text{diag } [\lambda_1, \ldots, \lambda_r, 0, \ldots, 0]$.

Since $P^{-1} = P'$, therefore $P^{-1}AP = D \Rightarrow P'AP = D \Rightarrow A$ is congruent to D.

Now consider the real orthogonal transformation $X = PY$. We have

$X'AX = (PY)' A (PY) = Y'P'APY = Y'DY = \lambda_1 y_1^2 + \ldots + \lambda_r y_r^2$.

Hence the result.

Theorem. *Every real quadratic form* **X'AX** *in n variables is real equivalent to the form*

$$y_1^2 + \ldots + y_p^2 - y_{p+1}^2 - \ldots - y_r^2,$$

where r is the rank of A *and p is the number of positive eigenvalues of* A.

Proof. A is a real symmetric matrix. Therefore there exists a real orthogonal matrix Q such that

$$Q^{-1}AQ = Q'AQ = D,$$

where D is a diagonal matrix whose diagonal elements are the eigenvalues of A. Since A is of rank r, therefore D is also of rank r. So D has exactly r non-zero diagonal elements. Consequently A has exactly r non-zero eigenvalues, the remaining $n-r$ eigenvalues of A being zero. Let $D = \text{diag } [\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0]$.

Let $\lambda_1, \ldots, \lambda_p$ be positive and $\lambda_{p+1}, \ldots, \lambda_r$ be negative. Let S be the $n \times n$ real diagonal matrix with diagonal elements

$$\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_p}}, \frac{1}{\sqrt{(-\lambda_{p+1})}}, \ldots, \frac{1}{\sqrt{(-\lambda_r)}}, 1, \ldots, 1.$$

Then S is non-singular and $S' = S$. If we take $P = QS$, then P is also a real non-singular matrix and we have

$$P'AP = (QS)' A (QS) = S'Q'AQS = SDS$$
$$= \text{diag } [1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0]$$

so that 1 and -1 appear p and $r-p$ times respectively.

Now the real non-singular linear transformation $X = PY$ reduces $X'AX$ to the form $Y'P'APY$ *i.e.*

$$y_1^2+\ldots+y_p^2-y_{p+1}^2-\ldots-y_r^2.$$

Hence the result.

Corollary. *Two real quadratic forms* $\mathbf{X'AX}$ *and* $\mathbf{Y'BY}$ *in the same number of variables are real equivalent if and only if* **A** *and* **B** *have the same number of positive and negative eigenvalues.*

Important Note. If $\mathbf{X'AX}$ is a real quadratic form, then the number of non-zero eigenvalues of **A** is equal to the rank of $\mathbf{X'AX}$ and the number of positive eigenvalues of **A** is equal to the index of $\mathbf{X'AX}$.

Theorem 8. *Two real quadratic forms* $\mathbf{X'AX}$ *and* $\mathbf{Y'BY}$ *are orthogonally equivalent if and only if* **A** *and* **B** *have the same eigenvalues and these occur with the same multiplicities.*

Proof. If **A** and **B** have eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and **D** is a diagonal matrix with $\lambda_1, \lambda_2, \ldots, \lambda_n$ as diagonal elements, then there exist orthogonal matrices **P** and **Q** such that $\mathbf{P'AP}=\mathbf{D}=\mathbf{Q'BQ}$.

Now $\qquad \mathbf{Q'BQ}=\mathbf{P'AP}$

$\Rightarrow \mathbf{B}=(\mathbf{Q}')^{-1}\,\mathbf{P'APQ}^{-1}=(\mathbf{Q}^{-1})'\,\mathbf{P'APQ}^{-1}=(\mathbf{PQ}^{-1})'\,\mathbf{A}\,(\mathbf{PQ}^{-1}).$

Since $\mathbf{PQ}^{-1}$ is an orthogonal matrix, therefore $\mathbf{Y'BY}$ is orthogonally equivalent to $\mathbf{X'AX}$.

Conversely, if the two forms are orthogonally equivalent, then there exists an orthogonal matrix **P** such that $\mathbf{B}=\mathbf{P'AP}=\mathbf{P}^{-1}\mathbf{AP}$. Therefore **A** and **B** are similar matrices and so have the same eigenvalues with the same multiplicities.

Solved Examples

Ex 1. *Reduce each of the following quadratic forms in three variables to real canonical form and find its rank and signature. Also write in each case the linear transformation which brings about the normal reduction.*

(*i*) $2x_1^2+x_2^2-3x_3^2-8x_2x_3-4x_3x_1+12x_1x_2.$ **[Patna 1969]**

(*ii*) $x^2-2y^2+3z^2-4yz+6zx.$ **[Rajasthan 1966]**

(*iii*) $6x_1^2+3x_2^2+14x_3^2+4x_2x_3+18x_3x_1+4x_1x_2.$ **[Poona 1959]**

(*iv*) $x^2+2y^2+2z^2-2xy-2yz+zx.$ **[Allahabad 1978]**

Solution. (i) The matrix **A** of the given quadratic form is

$$\mathbf{A}=\begin{bmatrix}2 & 6 & -2\\ 6 & 1 & -4\\ -2 & -4 & -3\end{bmatrix}.$$

We write $\mathbf{A}=\mathbf{IAI}$ *i.e.*

$$\begin{bmatrix}2 & 6 & -2\\ 6 & 1 & -4\\ -2 & -4 & -3\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Now we shall reduce **A** to diagonal form by applying congruence operations on it. Performing $R_2 \to R_2-3R_1,\ C_2 \to C_2-3C_1$ and $R_3 \to R_3+R_1,\ C_3 \to C_3+C_1$, we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & -17 & 2\\ 0 & 2 & -5\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -3 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & -3 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

[Note that we apply the row and column operations on **A** in two separate steps. But in order to save labour we should apply them in one step. For this we should not first write the first row. After changing R_2 and R_3 with the corresponding row operations we should simply write 0 in the second and third places of the first row and the first element of the first row should be kept unchanged].

Now performing $R_3 \to R_3+\frac{2}{17}R_2,\ C_3 \to C_3+\frac{2}{17}C_2$, we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & -17 & 0\\ 0 & 0 & -\frac{81}{17}\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -3 & 1 & 0\\ \frac{11}{17} & \frac{2}{17} & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & -3 & \frac{11}{17}\\ 0 & 1 & \frac{2}{17}\\ 0 & 0 & 1\end{bmatrix}.$$

Performing $R_1 \to \frac{1}{\sqrt{2}}R_1,\ C_1 \to \frac{1}{\sqrt{2}}C_1;\ R_2 \to \frac{1}{\sqrt{17}}R_2,\ C_2 \to \frac{1}{\sqrt{17}}C_2;$ and $R_3 \to \sqrt{(17/81)}\,R_3,\ C_3 \to \sqrt{(17/81)}\,C_3$, we get

$$\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}=\begin{bmatrix}a & 0 & 0\\ -3b & b & 0\\ \frac{11}{17}c & \frac{2}{17}c & c\end{bmatrix}\mathbf{A}\begin{bmatrix}a & -3b & \frac{11}{17}c\\ 0 & b & \frac{2}{17}c\\ 0 & 0 & c\end{bmatrix}$$

where $a=1/\sqrt{2},\ b=1/\sqrt{17},\ c=\sqrt{(17/81)}$.

Thus the linear transformation $\mathbf{X}=\mathbf{PY}$

where $\mathbf{P}=\begin{bmatrix}a & -3b & \frac{11}{17}c\\ 0 & b & \frac{2}{17}c\\ 0 & 0 & c\end{bmatrix},\ \mathbf{X}=[x_1\ x_2\ x_3]',\ \mathbf{Y}=[y_1\ y_2\ y_3]',$

transforms the given quadratic form to the normal form

$$y_1^2-y_2^2-y_3^2. \qquad \ldots(1)$$

The rank r of the given quadratic form = the number of non-zero terms in its normal form (1) = 3.

The signature of the given quadratic form = the excess of the number of positive terms over the number of negative terms in its normal form $=1-2=-1$.

The index of the given quadratic form = the number of positive terms in its normal form = 1.

The linear transformation $\mathbf{X}=\mathbf{PY}$ which brings about this normal reduction is given by

$$x=ay_1-3by_2+\frac{11}{7}cy_3,\ x_2=y_2+\frac{2}{17}cy_3,\ x_3=cy_3.$$

(ii) The matrix of the given quadratic form is

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

We write $\mathbf{A}=\mathbf{IAI}$ *i.e.*,

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

Now we shall reduce A to diagonal form by applying congruence operations on it. Performing $R_3\to R_3-3R_1$, $C_3\to C_3-3C_1$, we get

$$\begin{bmatrix}1 & 0 & 0\\ 0 & -2 & -2\\ 0 & -2 & -6\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -3 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & -3\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing $R_3\to R_3-R_2$, $C_3\to C_3-C_2$ we get

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

Performing $R_2\to\frac{1}{\sqrt{2}}R_2$, $C_2\to\frac{1}{\sqrt{2}}C_2$, $R_3\to\frac{1}{\sqrt{4}}R_3$, $C_3\to\frac{1}{\sqrt{4}}C_3$, we get

$$\begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1/\sqrt{2} & 0\\ -\frac{3}{2} & -\frac{1}{2} & \frac{1}{2}\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ 0 & 1/\sqrt{2} & -\frac{1}{2}\\ 0 & 0 & \frac{1}{2}\end{bmatrix}.$$

Thus the linear transformation $\mathbf{X}=\mathbf{PY}$ where

$$\mathbf{A}=\begin{bmatrix}1 & 0 & -\frac{3}{2}\\ 0 & 1/\sqrt{2} & -\frac{1}{2}\\ 0 & 0 & \frac{1}{2}\end{bmatrix},\ \mathbf{X}=[x\ y\ z]',\ \mathbf{Y}=[y_1\ y_2\ y_3]'$$ transforms

the given quadratic form to the normal form

$$y_1^2-y_2^2-y_3^2.$$

The rank of the given quadratic form is 3 and its signature is $1-2=-1$.

(ii) The matrix of the given quadratic form is

$$\mathbf{A}=\begin{bmatrix}6 & 2 & 9\\ 2 & 3 & 2\\ 9 & 2 & 14\end{bmatrix}.$$

We write

$$\begin{bmatrix}6 & 2 & 9\\ 2 & 3 & 2\\ 9 & 2 & 14\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing congruence operations $R_2\to R_2-\frac{1}{3}R_1$, $C_1\to C_2-\frac{1}{3}C_1$; and $R_3\to R_3-\frac{3}{2}R_1$, $C_3\to C_3-\frac{3}{2}C_1$, we get

$$\begin{bmatrix}6 & 0 & 0\\ 0 & \frac{7}{3} & -1\\ 0 & -1 & \frac{1}{2}\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -\frac{1}{3} & 1 & 0\\ -\frac{3}{2} & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & -\frac{1}{3} & -\frac{3}{2}\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing $R_3\to R_3+\frac{3}{7}R_2$, $C_3\to C_3+\frac{3}{7}C_2$, we get

$$\begin{bmatrix}6 & 0 & 0\\ 0 & \frac{7}{3} & 0\\ 0 & 0 & \frac{1}{14}\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -\frac{1}{3} & 1 & 0\\ \frac{23}{14} & \frac{3}{7} & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & -\frac{1}{3} & \frac{23}{14}\\ 0 & 1 & \frac{3}{7}\\ 0 & 0 & 1\end{bmatrix}.$$

Performing $R_1\to 1/\sqrt{6}R_1$, $C_1\to 1/\sqrt{6}C_1$; $R_2\to\sqrt{(3/7)}\,R_2$, $C_2\to\sqrt{(3/7)}\,C_2$; $R_3\to 1/14R_3$, $C_3\to 1/\sqrt{14}C_3$, we get

$$\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}a & 0 & 0\\ -\frac{1}{3}b & b & 0\\ \frac{23}{14}c & \frac{3}{7}c & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}a & -\frac{1}{3}b & \frac{23}{14}c\\ 0 & b & \frac{3}{7}c\\ 0 & 0 & 1\end{bmatrix}$$

where $a=\frac{1}{\sqrt{6}}$, $b=\sqrt{(3/7)}$, $c=\frac{1}{\sqrt{14}}$.

Thus the linear transformation $\mathbf{X}=\mathbf{PY}$ where

$$\mathbf{P}=\begin{bmatrix}a & -\frac{1}{3}b & \frac{23}{14}c\\ 0 & b & \frac{3}{7}c\\ 0 & 0 & 1\end{bmatrix}$$

transforms the given quadratic form to the normal form

$$y_1^2+y_2^2+y_3^2.$$

The rank of the given quadratic form is 3 and its signature is

$$3-0=3.$$

(iv) The matrix of the given quadratic form is

$$\mathbf{A}=\begin{bmatrix}1 & -1 & \frac{1}{2}\\ 1 & 2 & -1\\ \frac{1}{2} & -1 & 2\end{bmatrix}.$$

We write

$$\begin{bmatrix}1 & -1 & \frac{1}{2}\\ -1 & 2 & -1\\ \frac{1}{2} & -1 & 2\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}\mathbf{A}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing congruence operations

$R_2\to R_2+R_1$, $C_2\to C_2+C_1$ and $R_3\to R_3-\frac{1}{2}R_1$, $C_3\to C_3-\frac{1}{2}C_1$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{7}{4} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 1 & -\frac{1}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_3 \to R_3 + \frac{1}{2} R_2$, $C_3 \to C_3 + \frac{1}{2} C_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{3}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_3 \to \sqrt{\frac{2}{3}} R_3$, $C_3 \to \sqrt{\frac{2}{3}} C_3$, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1/\sqrt{6} & \sqrt{(2/3)} \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1/\sqrt{6} \\ 0 & 0 & \sqrt{(2/3)} \end{bmatrix}.$$

Thus the linear transformation $\mathbf{X} = \mathbf{PY}$ where

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1/\sqrt{6} \\ 0 & 0 & \sqrt{(2/3)} \end{bmatrix}$$

transforms the given quadratic form to the normal form

$$y_1^2 + y_2^2 + y_3^2.$$

Rank of the given quadratic form $= 3$.

Signature of the given quadratic form $= 3 - 0 = 3$.

Ex. 2. *Reduce the following quadratic form to canonical form and find its rank and signature :*

$$x^2 + 4y^2 + 9z^2 + t^2 - 12yz + 6zx - 4xy - 2xt - 6zt.$$

Solution. The matrix of the given quadratic form is

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & 3 & -1 \\ -2 & 4 & -6 & 0 \\ 3 & -6 & 9 & -3 \\ -1 & 0 & -3 & 1 \end{bmatrix}.$$

We write

$$\begin{bmatrix} 1 & -2 & 3 & -1 \\ -2 & 4 & -6 & 0 \\ 3 & -6 & 9 & -3 \\ -1 & 0 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Performing the congruent operations

$R_2 \to R_2 + 2R_1$, $C_2 \to C_2 + 2C_1$; $R_3 \to R_3 - 3R_1$, $C_3 \to C_3 - 3C_1$; and $R_4 \to R_4 + R_1$, $C_4 \to C_4 + C_1$, we get

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

Performing $R_2 \to R_2 + R_1$, we get

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

Now performing the corresponding column operation $C_2 \to C_2 + C_4$, we get

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

Performing $R_4 \to R_4 - \frac{1}{2} R_2$, $C_4 \to C_4 - \frac{1}{2} C_2$, we get

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 1 \\ -3 & 0 & 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 3 & -3 & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & \frac{1}{2} \end{bmatrix}.$$

Performing $R_2 \to \frac{1}{2} R_2$, $C_2 \to \frac{1}{2} C_2$, we get

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

The linear transformation $\mathbf{X} = \mathbf{PY}$ where

$\mathbf{X} = [x \quad y \quad z \quad t]'$, $\mathbf{Y} = [y_1 \quad y_2 \quad y_3 \quad y_4]'$ and

$$\mathbf{P} = \begin{bmatrix} 1 & \frac{3}{2} & -3 & -\frac{1}{2} \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

reduces the given quadratic form to the normal form

$$y_1^2 - y_2^2 + y_4^2.$$

Rank of the given quadratic form $=$ number of non-zero terms in its normal form $= 3$.

Signature $= 2 - 1 = 1$.

Ex. 3. *Find an orthogonal matrix* **P** *that will diagonalize the real symmetric matrix*

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}.$$

Interpret the result in terms of quadratic forms.

Solution. The characteristic equation of the given matrix is

$$|\mathbf{A}-\lambda\mathbf{I}|=0$$

i.e.
$$\begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & -1 \\ 1 & -1 & -\lambda \end{vmatrix}=0 \text{ i.e. } (\lambda-1)^2(\lambda+2)=0.$$

$\therefore$ the eigenvalues of A are $1, 1, -2$.

Corresponding to the eigenvalue 1 we can find two mutually orthogonal eigenvectors of A by solving

$$(\mathbf{A}-\mathbf{I})\,\mathbf{X}=\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or
$$-x_1+x_2+x_3=0.$$

Two orthogonal solutions are

$$\mathbf{X}_1=[1, \quad 0, \quad 1]' \text{ and } \mathbf{X}_2=[1, \quad 2, \quad -1]'.$$

An eigenvector corresponding to the eigenvalue -2, is found by solving $2x_1+x_2+x_3=0,\ x_1+2x_2-x_3=0$ to be $\mathbf{X}_3=[-1, 1, 1]'$. The required matrix P is therefore a matrix whose columns are unit vectors which are scalar multiples of $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{X}_3$.

$$\therefore\ \mathbf{P}=\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}.$$

We have $\mathbf{P}'\mathbf{A}\mathbf{P}=\text{diag. }[1, 1, -2]$.

The quadratic form corresponding to the symmetric matrix A is

$$\phi=2x_1x_2+2x_1x_3-2x_2x_3.$$

The orthogonal linear transformation $\mathbf{X}=\mathbf{P}\mathbf{Y}$ will transform it to the diagonal form $\quad y_1^2+y_2^2-2y_3^2$.

The rank of the quadratic form $\phi=$ the number of non-zero eigenvalues of its matrix $\mathbf{A}=3$.

The signature of the quadratic form $\phi=$ the number of positive eigenvalues of A $-$ the number of negative eigenvalues of $\mathbf{A}=2-1=1$. The normal form is $z_1^2+z_2^2-z_3^2$.

Exercises

1. Write the matrix and find the rank of each of the following quadratic forms :
   (i) $x_1^2-2x_1x_2+2x_2^2$.
   (ii) $4x_1^2+x_2^2-8x_3^2+4x_1x_2-4x_1x_3+8x_2x_3$.
2. Find a transformation $\mathbf{X}=\mathbf{P}\mathbf{Y}$ that will transform
   $$x^2+2y^2+3z^2+4xy+4yz$$

   to real canonical form. Also find the rank and signature of the given quadratic form.
3. Reduce each of the following quadratic forms to real canonical form and find its rank and signature :
   (i) $x_1x_2-4x_1x_4-2x_2x_3+12x_3x_4$.
   (ii) $x^2-4y^2+6z^2+2xy-4xz+2w^2-6zw$.
   (iii) $3x^2+3y^2+3z^2-2yz+2zx+2xy$.
   (iv) $2x^2+9y^2+2z^2-2yz+2zx+6xy$.
   (v) $4x_1^2+9x_2^2+2x_3^2+8x_2x_3+6x_3x_1+6x_1x_2$. **(Banaras 1968)**
   (vi) $x_1^2+2x_2^2+3x_3^2+2x_2x_3-2x_3x_1+2x_1x_2$. **(Punjab 1969)**
4. Reduce the quadratic form $7x^2-8y^2-8z^2-2yz-8zx+8xy$ to the canonical form by an orthogonal transformation and hence find the signature of the quadratic form. **(Punjab 1969)**

Answers

1. (i) $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$, rank 2. (ii) $\begin{bmatrix} 4 & 2 & -2 \\ 2 & 1 & 4 \\ -2 & 4 & -8 \end{bmatrix}$, rank 3.
2. Rank 3, signature 1.
3. (i) Rank 4, signature 0. (ii) rank 4, signature 0.
   (iii) rank 3, signature 3. (iv) rank 3, signature 3.
   (v) rank 3, signature 1.

§ 8. Value class of a real quadratic form. Definite, semi-definite and indefinite real quadratic forms.

Definitions. *Let* $\phi=\mathbf{X}'\mathbf{A}\mathbf{X}$ *be a real quadratic form in n variables* $x_1, \ldots, x_n$. *The form* ϕ *is said to be*

(i) **Positive Definite** *if* $\phi \geqslant 0$ *for all real values of the variables* $x_1, \ldots, x_n$ *and* $\phi=0$ *only if* $\mathbf{X}=\mathbf{O}$ *i.e.* $\phi=0 \Rightarrow x_1=x_2=\ldots=x_n=0$. **(Nagarjuna 1977)**

For example the quadratic form $x_1^2-4x_1x_2+5x_2^2$ in two variables is positive definite because it can be written as

$$(x_1-2x_2)^2+x_2^2,$$

which is $\geqslant 0$ for all real values of x_1, x_2 and

$$(x_1-2x_2)^2+x_2^2=0 \Rightarrow x_1-2x_2=0,\ x_2=0$$
$$\Rightarrow x_1=0,\ x_2=0.$$

Similarly the quadratic form $x_1^2+x_2^2+x_3^2$ in three variables is a positive definite form.

(ii) **Negative definite** *if* $\phi \leqslant 0$ *for all real values of the variables* $x_1, \ldots, x_n$ *aud* $\phi=0$ *only if* $x_1=x_2=\ldots=x_n=0$.

For example $-x_1^2-x_2^2-x_3^2$ is a negative definite form in three variables.

(*iii*) **Positive semi-definite** *if $\phi \geqslant 0$ for all real values of the variables $x_1, \ldots, x_n$ and $\phi = 0$ for some non-zero real vector* **X** *i.e. $\phi = 0$ for some real values of the variables $x_1, x_2, \ldots, x_n$ not all zero.*

For example the quadratic form

$$x_1^2 + x_2^2 + 2x_3^2 - 2x_1x_3 - 2x_2x_3$$

is positive semi-definite because it can be written in the form

$$(x_1 - x_3)^2 + (x_2 - x_3)^2,$$

which is $\geqslant 0$ for all real values of x_1, x_2, x_3 but is zero for non-zero values also, for example, $x_1 = x_2 = x_3 = 1$.

Similarly the quadratic form $x_1^2 + x_2^2 + 0x_3^2$ in three variables x_1, x_2, x_3 is positive semi-definite. It is non-negative for all real values of x_1, x_2, x_3 and it is zero for values $x_1 = 0$, $x_2 = 0$, $x_3 = 2$ which are not all zero.

(*iv*) **Negative semi-definite** *if $\phi \leqslant 0$ for all real values of the variables $x_1, \ldots, x_n$ and $\phi = 0$ for some values of the variables $x_1, \ldots, x_n$ not all zero.*

For example the quadratic form $-x_1^2 - x_2^2 - 0x_3^2$ in three variables x_1, x_2, x_3 is negative semi-definite.

(*v*) **Indefinite** *if ϕ takes positive as well as negative values for real values of the variables $x_1, \ldots, x_n$.*

For example the quadratic form $x_1^2 - x_2^2 + x_3^2$ in three variables is indefinite. It takes positive value 1 when $x_1 = 1$, $x_2 = 1$, $x_3 = 1$ and it takes negative value -1 when $x_1 = 0$, $x_2 = 1$, $x_3 = 0$.

Note 1. The above five classes of real quadratic forms are mutually exclusive and are called **value classes of real quadratic forms.** *Every real quadratic form must belong to one and only one value class.*

Note 2. A form which is positive definite or negative definite is called *definite* and a form which is positive semi-definite or negative semi-definite is called *semi-definite*.

Non-negative definite quadratic form.

Definition. *A real quadratic form $\phi = $* **X'AX** *in n variables $x_1, \ldots, x_n$, is said to be non-negative definite if it takes only non-negative values for all real values of $x_1, \ldots, x_n$.*

Thus ϕ is non-negative definite if $\phi \geqslant 0$ for all real values of $x_1, \ldots, x_n$. A non-negative definite quadratic form may be positive definite or positive semi-definite. It is positive definite if it takes the value 0 only when $x_1 = x_2 = \ldots = x_n = 0$.

Classification of real-symmetric matrices. Definite, semi-definite and indefinite real symmetric matrices.

Definition. *A real symmetric matrix* **A** *is said to be definite, semi-definite or indefinite if the corresponding quadratic form* **X'AX** *is definite, semi-definite or indefinite respectively.*

Positive definite real symmetric matrix.

Definition. *A real symmetric matrix* **A** *is said to be positive definite if the corresponding form* **X'AX** *is positive definite.*

Non-negative definite real symmetric matrix.

Definition. *A real symmetric matrix* **A** *is said to be non negative definite if the associated quadratic form* **X'AX** *is non-negative definite.*

Theorem 1. *All real equivalent real quadratic forms have the same value class.*

Proof. Let $\phi = \mathbf{X'AX}$ and $\psi = \mathbf{Y'BY}$ be two real equivalent real quadratic forms. Then there exists a real non-singular matrix **P** such that $\mathbf{P'AP} = \mathbf{B}$ and $(\mathbf{P}^{-1})'\,\mathbf{BP}^{-1} = \mathbf{A}$. The real non-singular linear transformation $\mathbf{X} = \mathbf{PY}$ transforms the quadratic form ϕ into the quadratic form ψ and the inverse transformation $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$ transforms the quadratic form ψ into the quadratic form ϕ. The two quadratic forms have the same ranges of values. The vectors **X** and **Y** for which ϕ and ψ have the same value are connected by the relations $\mathbf{X} = \mathbf{PY}$ and $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$. Thus the vector **Y** for which ψ has the same value as ϕ has for the vector **X** is given by $\mathbf{Y} = \mathbf{P}^{-1}\mathbf{X}$. Similarly the vector **X** for which ϕ has the same value as ψ has for the vector **Y** is given by $\mathbf{X} = \mathbf{PY}$.

Now we shall discuss the five cases separately.

Case I. *ϕ is positive definite if and only if ψ is positive definite.*

Suppose ϕ is positive definite.

Then $\phi \geqslant 0$ and $\phi = 0 \Rightarrow \mathbf{X} = \mathbf{O}$.

Since ϕ and ψ have same ranges of values, therefore

$$\phi \geqslant 0 \Rightarrow \psi \geqslant 0.$$

Also $\psi = 0 \Rightarrow \mathbf{Y'BY} = 0$

$\Rightarrow (\mathbf{PY})'\,\mathbf{A}\,(\mathbf{PY}) = 0$

[$\because$ ϕ has the same value for the vector **PY** as ψ has for the vector **Y**]

$\Rightarrow \mathbf{PY} = \mathbf{O}$ [$\because$ ϕ is positive definite means $\mathbf{X'AX} = 0 \Rightarrow \mathbf{X} = \mathbf{O}$]

$\Rightarrow \mathbf{P}^{-1}(\mathbf{PY}) = \mathbf{P}^{-1}\mathbf{O}$

$\Rightarrow \mathbf{Y} = \mathbf{O}$.

Thus ψ is also positive definite.

Conversely suppose that ψ is positive definite.

Then $\psi \geqslant 0$ and $\psi=0 \Rightarrow \mathbf{Y}=\mathbf{O}$.

Since ϕ and ψ have the same ranges of values, therefore

$$\psi \geqslant 0 \Rightarrow \phi \geqslant 0.$$

Also $\phi=0 \Rightarrow \mathbf{X'AX}=0$

$\Rightarrow (\mathbf{P}^{-1}\mathbf{X})'\ \mathbf{B}\ (\mathbf{P}^{-1}\ \mathbf{X})=0$ [$\because$ ψ has the same value for the vector $\mathbf{P}^{-1}\ \mathbf{X}$ as ϕ has for the vector $\mathbf{X}$]

$\Rightarrow \mathbf{P}^{-1}\mathbf{X}=\mathbf{O}$ [$\because$ ψ is positive definite]

$\Rightarrow \mathbf{X}=\mathbf{O}$.

Thus ϕ is also positive definite.

Case II. ϕ *is negative definite if and only if* ψ *is negative definite.*

The proof is the same as in case I.

The only difference is that we are to replace the expressions $\phi \geqslant 0, \psi \geqslant 0$ by the expressions $\phi \leqslant 0, \psi \leqslant 0$.

Case III. ϕ *is positive semi-definite if and only if* ψ *is positive semi-definite.*

Since ϕ and ψ have the same ranges of values, therefore $\phi \geqslant 0$ if and only if $\psi \geqslant 0$.

Further since $\mathbf{P}$ is non-singular, therefore

$$\mathbf{X} \neq \mathbf{O} \Rightarrow \mathbf{Y}=\mathbf{P}^{-1}\ \mathbf{X} \neq \mathbf{O}$$

and

$$\mathbf{Y} \neq \mathbf{O} \Rightarrow \mathbf{X}=\mathbf{PY} \neq \mathbf{O}.$$

Also the vectors $\mathbf{X}$ and $\mathbf{Y}$ for which ϕ and ψ have the same values are connected by the relations $\mathbf{X}=\mathbf{PY}$ and $\mathbf{Y}=\mathbf{P}^{-1}\ \mathbf{X}$. Therefore $\phi=0$ for some non-zero vector $\mathbf{X}$ if and only if $\psi=0$ for some non-zero vector $\mathbf{Y}$.

Hence ϕ is positive semi-definite if and only if ψ is positive semi-definite.

Case IV. ϕ *is negative semi-definite if and only if* ψ *is negative semi-definite.*

For proof replace the expression $\phi \geqslant 0, \psi \geqslant 0$ in case III by the expressions

$$\phi \leqslant 0, \psi \leqslant 0.$$

Case V. ϕ *is indefinite if and only if* ψ *is indefinite.* Since ϕ and ψ have the same ranges of values, therefore the result follows immediately.

Thus the proof of the theorem is complete.

Criterion for the value of a real quadratic form in terms of its rank and signature.

Theorem 2. *Suppose r is the rank and s is the signature of a real quadratic form $\phi=\mathbf{X'AX}$ in n variables. Then ϕ is (i) positive definite if and only if $s=r=n$, (ii) negative definite if and only if $-s=r=n$, (iii) positive semi-definite if and only if $s=r<n$, (iv) negative semi-definite if and only if $-s=r<n$; and (v) indefinite if and only if $|s| \neq r$.*

Proof. Let $\psi=y_1^2+\ldots+y_p^2-y_{p+1}^2-\ldots-y_r^2$ ...(1)

be the real canonical form of the real quadratic form ϕ of rank r and signature s. Then $s=2p-r$. Since ϕ and ψ are real equivalent real quadratic forms, therefore they have the same value class.

(i) Suppose $s=r=n$. Then $p=n$ and the real canonical form of ϕ becomes $y_1^2+\ldots+y_n^2$. But this is a positive definite quadratic form. So ϕ is also positive definite.

Conversely suppose that ϕ is positive definite. Then ψ is also a positive definite form in n variables. So we must have

$$\psi=y_1^2+\ldots+y_n^2.$$

Hence $r=n, p=n,$

$$2p-r=s=n.$$

(ii) Suppose $-s=r=n$. Then $s=2p-r$ gives $p=0$. The real canonical form of ϕ becomes $-y_1^2-\ldots-y_n^2$ which is negative definite and so ϕ is also negative definite.

Conversely if ϕ is negative definite, then ψ is also negative definite and so we must have $\psi=-y_1^2-\ldots-y_n^2$. Hence $r=n, p=0$, $2p-r=s=-n$ i.e. $-s=n$.

(iii) Suppose $s=r<n$. Then $s=2p-r$ gives $p=r$ and the real canonical form of ϕ becomes $y_1^2+\ldots+y_r^2$ where $r<n$. But this is a positive semi-definite form in n variables. So ϕ is also positive semi-definite.

Conversely if ϕ is positive semi-definite, then ψ is also a positive semi-definite form in n variables. So we must have $\psi=y_1^2+\ldots+y_r^2$ where $r<n$. Therefore $p=r<n$ and $s=2p-r=r$. Thus $s=r<n$.

(iv) Suppose $-s=r<n$. Then $s=2p-r$ gives $p=0$ and the real canonical form of ϕ becomes $-y_1^2-\ldots-y_r^2$ where $r<n$. This is a negative semi-definite form in n variables. So ϕ is also negative semi-definite.

Conversely if ϕ is negative semi-definite, then ψ is also a negative semi-definite form in n variables. So we must have

$$\psi = -y_1^2 - \ldots - y_r^2 \text{ where } r < n.$$

Therefore $p=0$ and $s=2p-r=-r$. Thus $-s=r<n$.

(v) Suppose $|s| \neq r$. Then $|2p-r| \neq r$. Therefore $p \neq 0$ and $p \neq r$ and so $0<p<r$. Then in this case the canonical form of ϕ has positive as well as negative terms and so it is an indefinite form. Consequently ϕ is also indefinite.

Conversely if ϕ is indefinite, then ψ is also indefinite. So there must be positive as well as negative terms in ψ. Therefore

$$|s| \neq r.$$

Criterion for the value class of a real quadratic form in terms of the eigenvalues of its matrix.

Suppose $\phi = X'AX$ is a real quadratic form in n variables. Then A is a real symmetric matrix of order n. Suppose r is the number of non-zero eigenvalues of A. Then r=rank of the quadratic form ϕ. Further if s=the number of positive eigenvalues of A – the number of negative eigenvalues of A, then s=signature of ψ. Hence with the help of theorem 2, we arrive at the following conclusion.

Theorem 3. *A real quadratic form $\phi = X'AX$ in n variables is*

(*i*) *positive definite if and only if all the eigenvalues of* A *are positive.* **(Nagarjuna 1977)**

(*ii*) *negative definite if and only if all the eigenvalues of* A *are negative.*

(*iii*) *positive semi-definite if and only if all the eigenvalues of* A *are $\geqslant 0$ and at least one eigenvalue is* 0.

(*iv*) *negative semi-definite if and only if all the eigenvalues of* A *are $\leqslant 0$ and at least one eigenvalue of* A *is* 0.

(*v*) *indefinite if and only if* A *has positive as well as negative eigenvalues.*

On account of its importance we shall give an independent proof of case (i).

Theorem 4. *A real symmetric matrix is positive definite if and only if all its eigenvalues are positive.* **(Nagarjuna 1977)**

Proof. Let A be a real symmetric matrix of order n. Then there exists an orthogonal matrix P such that

$$P^{-1}AP = P'AP = D = \text{diag}\,[\lambda_1, \lambda_2, \ldots, \lambda_n]$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the eigenvalues of A.

Let $X'AX$ be the real quadratic form corresponding to the matrix A. Let us transform this quadratic form by the real non-singular linear transformation $X = PY$ where $Y = [y_1, y_2, \ldots, y_n]'$. Then

$$X'AX = (PY)'A(PY) = Y'P'APY = Y'DY.$$

Therefore $$X'AX = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2. \quad \ldots(1)$$

Now suppose that $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive. Then the right hand side of (1) ensures that $X'AX \geqslant 0$ for all real vectors X. Also

$$\begin{aligned} & X'AX = 0 \\ \Rightarrow\ & \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 = 0 \\ \Rightarrow\ & y_1 = y_2 = \ldots = y_n = 0 \quad [\because \lambda_1, \ldots, \lambda_n \text{ are all positive}] \\ \Rightarrow\ & Y = O \\ \Rightarrow\ & P^{-1}X = O \quad [\because X = PY \Rightarrow Y = P^{-1}X] \\ \Rightarrow\ & P(P^{-1}X) = PO \\ \Rightarrow\ & X = O. \end{aligned}$$

Thus if $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive, then $X'AX$ is positive definite and so the matrix A is positive definite.

Conversely suppose that A is a positive definite matrix. Then the quadratic form $X'AX$ is positive definite. So

$$\begin{aligned} & X'AX \geqslant 0 \text{ for all real vectors X} \\ \Rightarrow\ & \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 \geqslant 0 \text{ for all real vectors Y} \\ \Rightarrow\ & \lambda_1, \ldots, \lambda_n \text{ are all} \geqslant 0. \end{aligned}$$

$$\begin{aligned} & \text{Also } X'AX = 0 \text{ only if } X = O \\ \Rightarrow\ & \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 = 0 \text{ only if } PY = O \\ \Rightarrow\ & \lambda_1 y_1^2 + \ldots + \lambda_n y_n^2 = 0 \text{ only if } Y = O \\ & [\because P \text{ is non-singular means } PY = O \text{ only if } Y = O] \\ \Rightarrow\ & \lambda_1, \ldots, \lambda_n \text{ are all not equal to zero.} \end{aligned}$$

Therefore if A is positive definite, then $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all >0.

This completes the proof of the theorem.

Corollary. ***A positive definite real symmetric matrix is non-singular.***

Proof. Suppose A is a positive definite real symmetric matrix. Then the eigenvalues of A are all positive. Also there exists an orthogonal matrix P such that

$$P^{-1}AP = D,$$

where D is a diagonal matrix having the eigenvalues of A as its diagonal elements. So all diagonal elements of D are positive and thus D is non-singular.

Now $$A = PDP^{-1} \Rightarrow A \text{ is non-singular.}$$

Theorem 5. *A real symmetric matrix* **A** *is positive definite if and only if there exists a non-singular matrix* **Q** *such that*

$$A=Q'Q.$$

Proof. Suppose **A** is positive definite. Then all the eigenvalues of **A** are positive and we can find an orthogonal matrix **P** such that

$$P^{-1}AP=P'AP=D=\text{diag. }[\lambda_1, \ldots, \lambda_n]$$

where each $\lambda_i>0$. Let $D_1=\text{diag. }[\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}]$. Then $D_1^2=D$ and $D_1'=D_1$. We have

$$A=PDP^{-1}=PD_1^2P^{-1}=PD_1D_1P'$$
$$=(PD_1)(PD_1)'=Q'Q \text{ where } Q=(PD_1)'.$$

Clearly, **Q** is non-singular since **P** and D_1 are non-singular.

Conversely suppose that $A=Q'Q$ where **Q** is non-singular. We have for all real vectors **X**,

$$X'AX=X'Q'QX=(QX)'(QX)$$
$$=Y'Y, \text{ where } Y=QX \text{ is a real } n\text{-vector}$$
$$\geqslant 0.$$

Also $X'AX=0 \Rightarrow Y'Y=0$

$$\Rightarrow Y=O$$
$$\Rightarrow QX=O \quad [\because \ Y=QX]$$
$$\Rightarrow Q^{-1}(QX)=Q^{-1}O$$
$$\Rightarrow X=O.$$

$\therefore$ **X'AX** is a positive definite real quadratic form and so the symmetric matrix **A** is positive definite.

Theorem 6. *Every real non-singular matrix* **A** *can be written as a product* $A=PS$, *where* **S** *is a positive definite symmetric matrix and* **P** *is orthogonal.*

Proof. Since **A** is non-singular, therefore by theorem 5, **A'A** is a positive definite real symmetric matrix. Let **Q** be an orthogonal matrix such that

$$Q^{-1}(A'A)Q=Q'(A'A)Q=D=\text{diag. }[\lambda_1, \ldots, \lambda_n],$$

where $\lambda_1, \ldots, \lambda_n$ are the positive real eigenvalues of **A'A**. Let $D_1=\text{diag. }[\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}]$. Then $D_1^2=D$ and $D_1'=D_1$.

Now let $S=QD_1Q'$. Clearly $S'=S$ and so **S** is symmetric. Moreover **S** is positive definite because it is similar to D_1 which has positive eigenvalues.

Also $S^2=QD_1Q'QD_1Q'=QD_1Q^{-1}QD_1Q'$

$$=QD_1^2Q^{-1}=QDQ^{-1}=A'A.$$

Now let $P=AS^{-1}$. Then **P** is orthogonal because

$$P'P=(AS^{-1})'AS^{-1}=(S^{-1})'A'AS^{-1}$$

$$=(S^{-1})'S^2S^{-1} \quad [\because \ A'A=S^2]$$
$$=(S^{-1})'S\,SS^{-1}=(S^{-1})'S$$
$$=(S')^{-1}S=S^{-1}S \quad [\because \ S \text{ is symmetric}]$$
$$=I.$$

Thus $S=QD_1Q'$ is a positive definite real symmetric matrix and $P=AS^{-1}$ is an orthogonal matrix and we have

$$PS=AS^{-1}S=A.$$

Hence the result.

Note. The decomposition $A=PS$ obtained in theorem 6 is called the *polar factorization of* **A**.

§ 9. Criterion for positive-definiteness of a quadratic form in terms of leading principal minors of its matrix.

Leading principal minors of a matrix. Definition. *Let*

$$A=\left[a_{ij}\right]_{n\times n}$$

be a square matrix of order n. Then

$$A_1=a_{11},\ A_2=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ A_3=\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ldots, A_n=\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

are called the leading principal minors of **A**.

Before stating the main theorem we shall prove the following Lemma.

Lemma. *If* **A** *is the matrix of a positive definite form, then* $|A|>0$.

Proof. If **X'AX** is a positive definite real quadratic form, then there exists a real non-singular matrix **P** such that

$$P'AP=I.$$

$\therefore \quad |P'AP|=|I|=1$

or $\quad |P'|\,|A|\,|P|=1$

or $\quad |A|=1/|P|^2 \quad [\because \ |P|=|P'|\neq 0]$

Therefore $|A|$ is positive.

Now we shall state and prove the main theorem.

Theorem. *A necessary and sufficient condition for a real quadratic form* **X'AX** *to be positive definite is, that the leading principal minors of* **A** *are all positive.* **(Nagarjuna 1978; I.A.S. 84)**

Proof. The condition is necessary. Suppose **X'AX** is a positive definite quadratic form in n variables. Let k be any natural number such that $k\leqslant n$. Putting $x_{k+1}=0, \ldots, x_n=0$ in the positive definite form **X'AX**, we get a positive definite form in k variables $x_1, \ldots, x_k$. The determinant of the matrix of this new quadratic form is the leading principal minor of order k of A and is positive by virtue of the lemma we have just proved. Thus every leading principal minor of the matrix of a positive definite quadratic form is positive.

The condition is sufficient. Now it is given that the leading principal minors of A are all positive and we are to prove that the form X'AX is positive definite. Here we shall use the principle of mathematical induction.

The result is true for quadratic forms in one variable since $a_{11}x^2$ is positive definite when a_{11} is positive.

Assume as our induction hypothesis that the theorem is true for quadratic forms in m variables. Then we shall prove that it is also true for quadratic forms in $(m+1)$ variables.

Now let S be any real symmetric matrix of order $(m+1)$ and let the leading principal minors of S be all positive. We partition S as follows :

$$S=\begin{bmatrix} B & B_1 \\ B_1' & \lambda \end{bmatrix},$$

where B is a real symmetric matrix of order m and B_1 is an $m\times 1$ column matrix.

By hypothesis the leading principal minors of S are all positive. Therefore $|S|$ and the leading principal minors of B are all positive. Thus B is a real symmetric matrix of order m having all its leading principal minors positive. So by induction hypothesis the quadratic form corresponding to B is positive definite. Therefore there exists a non-singular matrix P of order m such that

$$P'BP=I_m.$$

Since $|B|>0$, therefore B is non-singular. Let $C=-B^{-1}B_1$. Then C is an $m\times 1$ column matrix. Also

$$C'=-(B^{-1}B_1)'=-B'_1\,(B^{-1})'$$
$$=-B_1'\,(B')^{-1}=-B'_1\,B^{-1},$$

since $B'=B$, B being symmetric. We have

$$\begin{bmatrix} P' & O \\ C' & 1 \end{bmatrix} S \begin{bmatrix} P & C \\ O & 1 \end{bmatrix}=\begin{bmatrix} P' & O \\ C' & 1 \end{bmatrix}\begin{bmatrix} B & B_1 \\ B'_1 & \lambda \end{bmatrix}\begin{bmatrix} P & C \\ O & 1 \end{bmatrix}$$
$$=\begin{bmatrix} P'B & P'B_1 \\ C'B+B'_1 & C'B_1+\lambda \end{bmatrix}\begin{bmatrix} P & C \\ O & 1 \end{bmatrix}$$
$$=\begin{bmatrix} P'BP & P'BC+P'B_1 \\ C'BP+B'_1P & C'BC+B'_1C+C'B_1+\lambda \end{bmatrix}$$
$$=\begin{bmatrix} I_m & O \\ O & B'_1C+\lambda \end{bmatrix}$$
$$[\because\ P'BP=I_m,\ C=-B^{-1}B_1,\ C'=-B'_1B^{-1}]$$

Thus $\begin{bmatrix} P' & O \\ C' & 1 \end{bmatrix} S \begin{bmatrix} P & C \\ O & 1 \end{bmatrix}=\begin{bmatrix} I_m & O \\ O & B'_1C+\lambda \end{bmatrix}.$

Taking determinants of both sides, we get

$$|P'|\,.\,|S|\,.\,|P|=|I_m|\,.\,|B_1'C+\lambda|=B_1'C+\lambda$$

because $B'_1C+\lambda$ is an 1×1 matrix.

$$\therefore\ |P|^2\,.\,|S|=B'_1C+\lambda\quad [\because\ |P|=|P'|].$$

Since $|S|>0$ and $|P|\neq 0$, therefore $B'_1C+\lambda$ is positive. Let $B'_1C+\lambda=\alpha^2$, where α is real. Then

$$\begin{bmatrix} P' & O \\ C' & 1 \end{bmatrix} S \begin{bmatrix} P & C \\ O & 1 \end{bmatrix}=\begin{bmatrix} I_m & O \\ O & \alpha^2 \end{bmatrix}.$$

Pre-multiplying and post-multiplying both sides with

$$\begin{bmatrix} I_m & O \\ O & \alpha^{-1} \end{bmatrix},\ \text{we get}$$

$$\begin{bmatrix} I_m & O \\ O & \alpha^{-1} \end{bmatrix}\begin{bmatrix} P' & O \\ C' & 1 \end{bmatrix} S \begin{bmatrix} P & C \\ O & 1 \end{bmatrix}\begin{bmatrix} I_m & O \\ O & \alpha^{-1} \end{bmatrix}=I_{m+1}.$$

Now let $Q=\begin{bmatrix} P & C \\ O & 1 \end{bmatrix}\begin{bmatrix} I_m & O \\ O & \alpha^{-1} \end{bmatrix}.$

Then Q is non-singular as it is the product of two non-singular matrices. Also

$$Q'=\begin{bmatrix} I_m & O \\ O & \alpha^{-1} \end{bmatrix}\begin{bmatrix} P' & O \\ C' & 1 \end{bmatrix}.$$

Therefore, we have

$$Q'SQ=I_{m+1}.$$

Thus the real symmetric matrix S of order $m+1$ is congruent to I_{m+1}. So the quadratic form corresponding to S is positive definite.

The proof is now complete by induction.

Solved Examples

Ex. 1. *Prove that the quadratic form*

$$6x_1^2+3x_2^2+14x_3^2+4x_2x_3+18x_3x_1+4x_1x_2$$

in three variables is positive definite.

Solution. The matrix A of the given quadratic form is

$$A=\begin{bmatrix} 6 & 2 & 9 \\ 2 & 3 & 2 \\ 9 & 2 & 14 \end{bmatrix}.$$

Let r be the rank and s be the signature of the given quadratic form. Then proceeding as in Ex. 1, part (iii) page 358, we find that $r=3$, $s=3$. Thus $r=s=n$. Therefore the given quadratic form is positive definite.

Ex. 2. *Prove that the quadratic form*

$$6x_1^2+3x_2^2+3x_3^2-4x_1x_2-2x_2x_3+4x_3x_1$$

in three variables is positive definite.

Solution. The matrix A of the given quadratic form is

$$A=\begin{bmatrix}6 & -2 & 2\\ -2 & 3 & -1\\ 2 & -1 & 3\end{bmatrix}.$$

The leading principal minors of A are

$$A_1=6,\ A_2=\begin{vmatrix}6 & -2\\ -2 & 3\end{vmatrix}=18-4=14,$$

$$A_3=\begin{vmatrix}6 & -2 & 2\\ -2 & 3 & -1\\ 2 & -1 & 3\end{vmatrix}=\begin{vmatrix}0 & 1 & -7\\ 0 & 2 & 2\\ 2 & -1 & 3\end{vmatrix}, \text{ by } R_2+R_3,\ R_1-3R_3$$

$$=2(2+14)=\text{positive}.$$

Since the leading principal minors of A are all positive, therefore the given quadratic form is positive definite.

Ex. 3. *Prove that the quadratic form*

$$2x_1^2+x_2^2-3x_3^2-8x_2x_3-4x_3x_1+12x_1x_2$$

in three variables is indefinite. **(Poona 1959)**

Solution. The matrix A of the given quadratic form is

$$A=\begin{bmatrix}2 & 6 & -2\\ 6 & 1 & -4\\ -2 & -4 & -3\end{bmatrix}.$$

Proceeding as in Ex. 1 part (i) page 356, we find that the real canonical form of the given quadratic form is $y_1^2-y_2^2-y_3^2$ which is indefinite form. Hence the given quadratic form is indefinite.

Ex. 4. *Prove that the quadratic form*

$$6x^2+49y^2+51z^2-82yz+20zx-4xy$$

in three variables is positive definite. **(Agra 1967, Rajasthan 64)**

Solution. The matrix A of the given quadratic form is

$$A=\begin{bmatrix}6 & -2 & 10\\ -2 & 49 & -41\\ 10 & -41 & 51\end{bmatrix}.$$

The leading principal minors of A are

$$A_1=6,\ A_2=\begin{vmatrix}6 & -2\\ -2 & 49\end{vmatrix}=6\times 49-4=\text{positive},$$

$$A_3=\begin{vmatrix}6 & -2 & 10\\ -2 & 49 & -41\\ 10 & -41 & 51\end{vmatrix}=\begin{vmatrix}0 & 145 & -113\\ -2 & 49 & -41\\ 0 & 204 & -154\end{vmatrix}\ \begin{matrix}R_1+3R_2,\\ R_3+5R_2\end{matrix}$$

$$=2\,(-145\times 154+204\times 113)=\text{positive}.$$

Since the leading principal minors of A are positive, therefore the given quadratic form is positive definite.

Ex. 5. *Write the matrix A of the quadratic form*

$$6x^2+65y^2+11z^2+4zx.$$

Find the eigenvalues of A and hence determine the value class of the given quadratic form.

Solution. The matrix A of the given quadratic form is

$$A=\begin{bmatrix}6 & 0 & 2\\ 0 & 35 & 0\\ 2 & 0 & 11\end{bmatrix}.$$

The characteristic equation of A is

$$\begin{vmatrix}6-\lambda & 0 & 2\\ 0 & 35-\lambda & 0\\ 2 & 0 & 11-\lambda\end{vmatrix}=0$$

or $\quad (6-\lambda)\,(35-\lambda)\,(11-\lambda)+2\times 2\,(35-\lambda)=0$

or $\quad (35-\lambda)\,[(6-\lambda)\,(11-\lambda)+4]=0$

or $\quad (35-\lambda)\,[\lambda^2-17\lambda+70]=0$

or $\quad (35-\lambda)\,(\lambda-7)\,(\lambda-10)=0.$

$\therefore$ the eigenvalues of A are 35, 10, 7.

Since the eigenvalues of A are all positive, therefore the quadratic form is positive definite.

Ex. 6. *Show that the quadratic form*

$$5x_1^2+26x_2^2+10x_3^2+4x_2x_3+14x_3x_1+6x_1x_2$$

in three variables is positive semi-definite and find a non-zero set of values of x_1, x_2, x_3 which makes the form zero.

Solution. The matrix of the given form is

$$A=\begin{bmatrix}5 & 3 & 7\\ 3 & 26 & 2\\ 7 & 2 & 10\end{bmatrix}.$$

We write

$$\begin{bmatrix}5 & 3 & 7\\ 3 & 26 & 2\\ 7 & 2 & 10\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}A\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Now we shall reduce A to diagonal form by applying congruence operations on it. Performing $R_2 \to R_2 - 3/5\, R_1$, $C_2 \to C_2 - 3/5 C_1$; $R_3 \to R_3 - 7/5\, R_1$, $C_3 \to C_3 - 7/5\, C_1$, we get

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 121/5 & -11/5 \\ 0 & -11/4 & 1/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3/5 & 1 & 0 \\ -7/5 & 0 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & -\frac{3}{5} & -17/5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Performing $R_3 \to R_3 + \frac{1}{11} R_2$, $C_3 \to C_3 + \frac{1}{11} C_2$, we get

$$\begin{bmatrix} 5 & 0 & 0 \\ 0 & 121/5 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3/5 & 1 & 0 \\ -16/11 & 1/11 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & -\frac{3}{5} & -16/11 \\ 0 & 1 & 1/11 \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore the linear transformation

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & -3/5 & -16/11 \\ 0 & 1 & 1/11 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

i.e. $\quad x_1 = y_1 - \dfrac{3}{5} y_2 - \dfrac{16}{11} y_3,\quad x_2 = y_2 + \dfrac{1}{11} y_3,\quad x_3 = y_3$

transforms the given form to the diagonal form

$$5y_1^2 + \frac{121}{5} y_2^2 + 0y_3^2. \qquad \ldots(1)$$

But the quadratic form (1) in 3 variables is positive semi-definite and equivalent quadratic forms have the same value class. Therefore the given quadratic form is positive semi-definite.

The set of values $y_1 = 0$, $y_2 = 0$, $y_3 = 1$ makes (1) zero. Corresponding to this set of values, we have $x_3 = 1$, $x_2 = \frac{1}{11}$, $x_1 = -\frac{16}{11}$. This is a non-zero set of values of x_1, x_2, x_3 which makes the given quadratic form zero.

Ex. 7. *Show that the quadratic form*

$$6x^2 + 17y^2 + 3z^2 - 20xy - 14yz + 8zx$$

in three variables is positive semi-definite and find a non-zero set of values of x, y, z which makes the form zero.

Solution. The matrix A of the given form is

$$\mathbf{A} = \begin{bmatrix} 6 & -10 & 4 \\ -10 & 17 & -7 \\ 4 & -7 & 3 \end{bmatrix}.$$

We write

$$\begin{bmatrix} 6 & -10 & 4 \\ -10 & 17 & -7 \\ 4 & -7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

To avoid fractions, we first perform the row operations $R_2 \to 3R_2$, $R_3 \to 3R_3$ and obtain

$$\begin{bmatrix} 6 & -10 & 4 \\ -30 & 51 & -21 \\ 12 & -21 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now we perform the corresponding column operations $C_2 \to 3C_2$, $C_3 \to 3C_3$ and obtain

$$\begin{bmatrix} 6 & -30 & 12 \\ -30 & 153 & -63 \\ 12 & -63 & 27 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Performing the congruent operations

$R_2 \to R_2 + 5R_1$, $C_2 \to C_2 + 5C_1$; and $R_3 \to R_3 - 2R_1$, $C_3 \to C_3 - 2C_1$, we get

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & -3 \\ 0 & -3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 3 & 0 \\ -2 & 0 & 3 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 5 & -2 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

Performing $R_3 \to R_3 + R_2$, $C_3 \to C_3 + C_2$, we get

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 5 & 3 & 0 \\ 3 & 3 & 3 \end{bmatrix} \mathbf{A} \begin{bmatrix} 1 & 5 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{bmatrix}.$$

Therefore the linear transformation

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 5 & 3 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

i.e. $\quad x = y_1 + 5y_2 + 3y_3,\quad y = 3y_2 + 3y_3,\quad z = 3y_3$

transforms the given form to the diagonal form

$$6y_1^2 + 3y_2^2 + 0y_3^2. \qquad \ldots(1)$$

But the quadratic form (1) in three variables is positive semi-definite. Therefore the given quadratic form is also positive semi-definite.

The set of values $y_1 = 0$, $y_2 = 0$, $y_3 = 1$ makes (1) zero. Corresponding values of x, y, z are $z = 3$, $y = 3$, $x = 3$. Thus $x = y = z = 3$ is a non-zero set of values of x, y, z which makes the given form

Ex. 8. *Show that the form*

$$x_1^2+2x_2^2+3x_3^2+2x_2x_3-2x_3x_1+2x_1x_2$$

in three variables is indefinite and find two sets of values of x_1, x_2, x_3 for which the form assumes positive and negative values. **(Punjab 1967)**

Solution. The matrix of the given quadratic form is

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

We write

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

Now we shall reduce A to diagonal form by applying congruent operations on it. Performing

$R_2\to R_2-R_1$, $C_2\to C_2-C_1$; and $R_3\to R_3+R_1$, $C_3\to C_3+C_1$, we get

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

Performing $R_3\to R_3-2R_2$, $C_3\to C_3-2C_2$, we get

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

$\therefore$ the linear transformation

$$\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}1 & -1 & 3\\ 0 & 1 & -2\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}$$

i.e.

$$\left.\begin{aligned}x_1&=y_1-y_2+3y_3\\ x_2&=y_2-2y_3\\ x_3&=y_3\end{aligned}\right\}\quad\ldots(A)$$

transforms the given form to the diagonal form $y_1^2+y_2^2-2y_3^2$. ...(1)

The form (1) is indefinite and so the given quadratic form is also indefinite.

Obviously $y_1=0$, $y_2=0$, $y_3=1$ makes the form (1) negative and $y_1=0$, $y_2=1$, $y_3=0$ makes the form (1) positive. Substituting these values in the relations (A), we see that the sets of values $x_1=3$, $x_2=-2$, $x_3=1$; $x_1=-1$, $x_2=1$, $x_3=0$ respectively make the given form negative and positive.

Ex. 9. *Classify the following forms in three variables as definite, semi-definite and indefinite*

(i) $2x^2+2y^2+3z^2-4yz-4zx+2xy$. **(Nagarjuna 1990)**

(ii) $26x^2+20y^2+10z^2-4yz-16zx-36xy$

(iii) $x_1^2+4x_2^2+x_3^2-4x_2x_3+2x_3x_1-4x_1x_2$.

Solution. (i) The matrix of the form is

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

We write

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

We shall reduce A to diagonal form by congruent operations. To avoid fractions we first apply the row operation $R_2\to 2R_2$ and obtain

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

Now applying the corresponding column operation $C_2\to 2C_2$, we obtain

$$\begin{bmatrix}2 & 2 & -2\\ 2 & 8 & -4\\ -2 & -4 & 3\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}A\begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing $R_2\to R_2-R_1$, $C_2\to C_2-C_1$; and $R_3\to R_3+R_1$, $C_3\to C_3+C_1$, we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & 6 & -2\\ 0 & -2 & 1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -1 & 2 & 0\\ 1 & 0 & 1\end{bmatrix}A\begin{bmatrix}1 & -1 & 1\\ 0 & 2 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Performing $R_3\to 3R_3$, $C_3\to 3C_3$, we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & 6 & -6\\ 0 & -6 & 9\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -1 & 2 & 0\\ 3 & 0 & 3\end{bmatrix}A\begin{bmatrix}1 & -1 & 3\\ 0 & 2 & 0\\ 0 & 0 & 3\end{bmatrix}.$$

Performing $R_3\to R_3+R_2$, $C_3\to C_3+C_2$, we get

$$\begin{bmatrix}2 & 0 & 0\\ 0 & 6 & 0\\ 0 & 0 & 3\end{bmatrix}=\begin{bmatrix}1 & 0 & 0\\ -1 & 2 & 0\\ 2 & 2 & 3\end{bmatrix}A\begin{bmatrix}1 & -1 & 2\\ 0 & 2 & 2\\ 0 & 0 & 3\end{bmatrix}.$$

$\therefore$ the given quadratic form transforms to the diagonal form $2y_1^2+6y_2^2+3y_3^2$.

$\therefore$ r=rank of the given quadratic form=3

s=signature of the given quadratic form $=2r-p$

$$=(2\times 3)-3=3.$$

$\therefore$ $r=s=3$.

Hence the given form is positive definite.

(ii) The matrix of the given form is

$$\mathbf{A}=\begin{bmatrix} 26 & -18 & -8 \\ -18 & 20 & -2 \\ -8 & -2 & 10 \end{bmatrix}.$$

The characteristic equation of **A** is

$$\begin{vmatrix} 26-\lambda & -18 & -8 \\ -18 & 20-\lambda & -2 \\ -8 & -2 & 10-\lambda \end{vmatrix}=0$$

or $$\begin{vmatrix} -\lambda & -18 & -8 \\ -\lambda & 20-\lambda & -2 \\ -\lambda & -2 & 10-\lambda \end{vmatrix}=0, \text{ by } C_1+C_2+C_3$$

or $$-\lambda\begin{vmatrix} 1 & -18 & -8 \\ 1 & 20-\lambda & -2 \\ 1 & -2 & 10-\lambda \end{vmatrix}=0$$

or $$-\lambda\begin{vmatrix} 1 & -18 & -8 \\ 0 & 2-\lambda & 6 \\ 0 & 16 & 18-\lambda \end{vmatrix}=0, \ R_2-R_1, R_3-R_1$$

or $\quad -\lambda[(2-\lambda)(18-\lambda)-96]=0$

or $\quad -\lambda[\lambda^2-20\lambda-60]=0.$

$\therefore$ the eigenvalues of **A** are

$$\lambda=0, \lambda=\frac{20\pm\sqrt{(400+240)}}{2}$$

i.e. $\lambda=0$, λ=a positive number, λ=a negative number.

Since **A** has positive as well as negative eigenvalues, therefore the given form is indefinite.

(iii) The matrix of the given form is

$$\mathbf{A}=\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}.$$

The characteristic equation of **A** is

$$\begin{vmatrix} 1-\lambda & -2 & 1 \\ -2 & 4-\lambda & -2 \\ 1 & -2 & 1-\lambda \end{vmatrix}=0$$

or $$\begin{vmatrix} -\lambda & -2 & 1 \\ -\lambda & 4-\lambda & -2 \\ -\lambda & -2 & 1-\lambda \end{vmatrix}=0, \text{ by } C_1+C_2+C_3$$

or $$-\lambda\begin{vmatrix} 1 & -2 & 1 \\ 1 & 4-\lambda & -2 \\ 1 & -2 & 1-\lambda \end{vmatrix}=0$$

or $$-\lambda\begin{vmatrix} 1 & -2 & 1 \\ 0 & 6-\lambda & -3 \\ 0 & 0 & -\lambda \end{vmatrix}=0, \ R_2-R_1, R_3-R_1$$

or $\quad -\lambda[-\lambda(6-\lambda)]=0$

or $\quad \lambda^2(6-\lambda)=0.$

$\therefore$ the eigenvalues of **A** are 0, 0, 6.

Since the eigenvalues of **A** are all non-negative and **A** has at least one zero eigenvalue, therefore the given quadratic form is positive semi-definite.

Ex. 10. *Show that every real non-singular matrix* **A** *can be expressed as*

$$\mathbf{A}=\mathbf{QDR},$$

where **Q** *and* **R** *are orthogonal and* **D** *is real diagonal.*

Solution. Since **A** is a real non-singular matrix, therefore $\mathbf{A'A}$ is a positive definite real symmetric matrix. Let **P** be an orthogonal matrix such that

$$\mathbf{P}^{-1}(\mathbf{A'A})\mathbf{P}=\mathbf{P'}(\mathbf{A'A})\mathbf{P}=\text{diag.}[\lambda_1,\ldots,\lambda_n],$$

where $\lambda_1,\ldots,\lambda_n$ are the positive real eigenvalues of the positive defiite matrix $\mathbf{A'A}$.

Let $\mathbf{D}=\text{diag}[\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n}]$. Then **D** is a real diagonal matrix and $\mathbf{D'}=\mathbf{D}$.

We have

$\mathbf{D'D}=\mathbf{D}^2=\text{diag}[\lambda_1,\ldots,\lambda_n]$

$\Rightarrow \mathbf{D'D}=\mathbf{P'A'AP}$

$\Rightarrow (\mathbf{P'})^{-1}\mathbf{D'DP}^{-1}=(\mathbf{P'})^{-1}\mathbf{P'A'APP}^{-1}$

$\Rightarrow (\mathbf{P'})^{-1}\mathbf{D'DP}^{-1}=\mathbf{A'A}$

$\Rightarrow (\mathbf{A'})^{-1}(\mathbf{P'})^{-1}\mathbf{D'D}\,\mathbf{P}^{-1}\mathbf{A}^{-1}=\mathbf{I}$

$\Rightarrow (\mathbf{A}^{-1})'\,\mathbf{PD'DP'A}^{-1}=\mathbf{I}$ $\quad$ [$\because$ **P** is orthogonal $\Rightarrow \mathbf{P'}=\mathbf{P}^{-1}$]

$\Rightarrow (\mathbf{DP'A}^{-1})'(\mathbf{DP'A}^{-1})=\mathbf{I}$

$\Rightarrow \mathbf{DP'A}^{-1}$ is orthogonal.

Let $\mathbf{S}=\mathbf{DP'A}^{-1}$.

Then **S** is an orthogonal matrix.

Now let $Q=S^{-1}$; then Q is an orthogonal matrix. Also let $R=P'$. Then R is an orthogonal matrix.

We have

$QDR=(DP'A^{-1})^{-1}DP'=A(P')^{-1}D^{-1}DP'=A(P')^{-1}P'=A.$

Hence the result.

Ex. 11. *If* A *is a positive definite real symmetric matrix, show that there exists a positive definite real symmetric matrix* B *such that*

$$B^2=A.$$

Solution. Since A is a positive definite real symmetric matrix, therefore the eigenvalues $\lambda_1,\ldots,\lambda_n$ of A are all real and positive. Also there exists an orthogonal matrix P such that

$$P^{-1}AP=D=\text{diag}[\lambda_1,\ldots,\lambda_n].$$

Let $D_1=\text{diag}[\sqrt{\lambda_1},\ldots,\sqrt{\lambda_n}]$. Then $D_1^2=D$, $D_1'=D_1$, and the eigenvalues of D_1 are all positive.

Now suppose that $B=PD_1P^{-1}=PD_1P'$.

We have $B'=(PD_1P')'=PD_1'P'=PD_1P'=B$.

$\therefore$ B is a real symmetric matrix.

Also $B=PD_1P^{-1}\Rightarrow B$ is similar to D_1. So B and D_1 have the same eigenvalues. Therefore the eigenvalues of B are all positive. So B is a positive definite real symmetric matrix.

Finally, we have

$B^2=(PD_1P^{-1})^2=PD_1P^{-1}PD_1P^{-1}=PD_1^2P^{-1}=PDP^{-1}=A.$

Hence the result.

Exercises

1. Show that the quadratic form $2x^2-4xy+3xz+6y^2+6yz+8z^2$ in three variables is positive definite. **(Punjab 1971)**
2. Determine the value class of the form $6x^2+12y^2+8yz+4zx$ in three variables.
3. Show that the quadratic form $y^2+2z^2-2yz+2zx-2xy$ in three variables is indefinite.
4. Determine the value class of the form $-y^2+2yz-2xy$ in three variables.
5. Show that a real symmetric matrix is non-negative definite if and only if all its eigenvalues are non-negative.
6. Prove that every positive semi-definite symmetric matrix A has a positive semi-definite square root B such that $B^2=A$.

§ 10. Hermitian Forms.

Definition. *If* $A=[a_{ij}]_{n\times n}$ *is a Hermitian matrix of order n, then the expression*

$$X^\theta AX=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}\bar{x}_i x_j,$$

is called a Hermitian form in n variables $x_1,\ldots,x_n$. *The Hermitian matrix* A *is called the matrix of this Hermitian form.*

Theorem 1. *A Hermitian form* $X^\theta AX$ *assumes only real values for all complex n-vectors* X.

Proof. Suppose $X^\theta AX$ is a Hermitian form. Then A is a Hermitian matrix and $A^\theta=A$.

Since $X^\theta AX$ is a 1×1 matrix, therefore it is symmetric and so

$$(X^\theta AX)'=X^\theta AX.$$

Now $\overline{(X^\theta AX)}=\overline{(X^\theta AX)'}=(X^\theta AX)^\theta=X^\theta A^\theta X=X^\theta AX.$

Thus $X^\theta AX$ and its conjugate are equal.

$\therefore$ $X^\theta AX$ is a real 1×1 matrix.

Hence $X^\theta AX$ has only real values.

Theorem 2. *The determinant and every leading principal minor of a Hermitian matrix* A *are real.*

Proof We have $\overline{|A|}=|\bar{A}|=|(\bar{A})'|=|A^\theta|$

$=|A|$ [$\because$ $A^\theta=A$, A being Hermitian]

Thus $|A|$ and its conjugate are equal.

$\therefore$ $|A|$ is real.

Since every leading principal sub-matrix of a Hermitian matrix is Hermitian, we conclude that every leading principal minor of a Hermitian matrix is real.

Non-negative definite and Positive definite Hermitian forms and matrices. Definition. *A Hermitian form* $X^\theta AX$ *is said to be non-negative definite if* $X^\theta AX\geqslant0$ *for all complex n-vectors* X. *It is positive definite if* $X^\theta AX\geqslant0$ *for all complex n-vectors* X *and in addition* $X^\theta AX=0$ *implies* $X=O$. *A Hermitian matrix* A *is called positive (non-negative) definite if the Hermitian form* $X^\theta AX$ *is positive (non-negative) definite.*

§ 11. Unitary Reduction of a Hermitian form.

Theorem *Every Hermitian form* $X^\theta AX$ *is unitarily equivalent (under a transformation* $X=PY$, P *unitary) to the form*

$$\lambda_1\bar{y}_1y_1+\ldots+\lambda_n\bar{y}_ny_n$$

where $\lambda_1,\ldots,\lambda_n$ *are eigenvalues of* A.

Proof. Since A is a Hermitian matrix, therefore there exists a unitary matrix P such that

$$P^{-1}AP = P^{\theta}AP = D = \text{diag}\,[\lambda_1, \dots, \lambda_n]$$

where $\lambda_1, \dots, \lambda_n$ are eigenvalues of A.

Consider the unitary linear transformation $X = PY$. We have

$$X^{\theta}AX = (PY)^{\theta}\,A\,(PY) = Y^{\theta}P^{\theta}APY = Y^{\theta}DY$$
$$= \lambda_1\,\bar{y}_1\,y_1 + \dots + \lambda_n\,\bar{y}_n\,y_n.$$

Hence the result.

Exercises

1. Show that a Hermitian matrix H is positive definite if and only if its eigenvalues are all positive and is non-negative definite if and only if its eigenvalues are non-negative.
2. Show that a Hermitian matrix H is positive definite if and only if there exists a non-singular matrix Q such that
$$H = Q^{\theta}Q.$$
3. If A is positive definite or a positive semi-definite Hermitian matrix show that there exists a Hermitian matrix B such that
$$B^2 = A.$$