7. Theory of Matrices

7.0. Introduction

In chapter 5 we have introduced $m \times n$ matrices and we have represented linear transformations by these matrices. In this chapter we shall develop the general theory of matrices. Throughout this chapter we deal with matrices whose entries are from the field F of real or complex numbers.

7.1. Algebra of Matrices

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

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

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

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

Addition of matrices. We have already defined the addition of two $m \times n$ matrices $A = (a_{ij})$ and

$B = (b_{ij})$ by $A + B = (a_{ij} + b_{ij})$.

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

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

$A + B = \begin{bmatrix} 1 & 6 \\ 5 & 5 \\ 8 & 5 \end{bmatrix}$

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

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

$\alpha A = (\alpha a_{ij})$.

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

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

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

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

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

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

Examples

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

$B = \begin{bmatrix} 1 & 1 \\ 1 & 5 \\ 3 & 2 \\ 1 & 0 \end{bmatrix}$. A is a 3×4 matrix and B is a 4×2 matrix. Hence the product AB is a 3×2

matrix and

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

$$= \begin{bmatrix} 10 & 5 \\ 6 & 12 \\ 2 & 1 \end{bmatrix}$$

Note that in this example the product BA is not defined. Even if the product BA is defined, AB need not be equal to BA.

2. Let $A = \begin{bmatrix} 2 & 4 & 0 \\ 9 & 3 & 1 \\ 4 & 7 & 2 \end{bmatrix}$ and

$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then $AI = IA = A$.

(Verify)

3. Consider the square matrix of order n given by

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

Let A be any $m \times n$ matrix. Then $I_n A = A$.

Also if A is an $m \times n$ matrix, $AI_n = A$.

If A is any $n \times n$ matrix, $AI_n = I_n A = A$.

I_n is called the **identity matrix** of order n.

We shall denote the identity matrix of any order by the symbol I.

Solved problems

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

Solution.

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

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

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

Now,

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

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

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

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

$$= \mathbf{0}.$$

Hence $A(A - I)(A + 2I) = \mathbf{0}$.

Problem 2. Prove that $\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^n = \begin{bmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{bmatrix}$

Solution. We prove this result by induction on n.

When $n = 1$ result is obviously true.

Let us assume that the result is true for $n = k$.

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

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

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

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

The result is true for $n = k + 1$

Hence the result is true for all positive integers n.

Exercises

1. Write down six pairs of matrices A and B such that the product AB is defined and in each case compute the product AB.

2. (a) Show that if A is an $m \times n$ matrix, then AB and BA are both defined iff B is an $n \times m$ matrix.

   (b) Write down six pairs of matrices A and B such that both AB and BA are defined and compute the products AB and BA.

3. If A and B are two matrices such that AB and $A + B$ are both defined, show that A, B are square matrices of the same order.

4. Let $A = \begin{bmatrix} 1 & -2 & 4 \\ -3 & 0 & 2 \\ 7 & 4 & 3 \end{bmatrix}$ and

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

   Compute A, B^2, AB and BA.

5. If $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$ show that

   $A^2 - 4A - 5I = \mathbf{0}$.

6. If $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{bmatrix}$ prove that

   $A^3 - 6A^2 + 7A + 2I = \mathbf{0}$.

7. Prove that if $A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix}$, then

   $A^k = \begin{bmatrix} 1 + 2k & -4k \\ k & 1 - 2k \end{bmatrix}$ for any positive integer k.

8. Decide which of the following statements are true and which are false.

   (a) For any two matrices A and B, $A + B$ is defined.

   (b) AB is defined $\Rightarrow$ BA is defined.

   (c) For any matrix A, A^2 is defined.

   (d) For any square matrix A, A^2 is defined.

   (e) Matrix addition is commutative.

   (f) Matrix addition is associative.

   (g) Matrix multiplication is commutative.

   (h) If A and B are 3×3 matrices then $(A + B)^2 = A^2 + 2AB + B^2$.

   (i) If A and B are 3×3 matrices then $(A + B)(A - B) = A^2 - B^2$.

   (j) (h) and (i) are true if $AB = BA$.

Answers.

8. (a) F (b) F (c) F (d) T (e) T
   (f) T (g) F (h) F (i) F (j) T

Theorem 7.1. Let A be an $m \times n$ matrix, B an $n \times p$ matrix and C a $p \times q$ matrix. Then $A(BC) = (AB)C$.

Proof. Let $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$. Let us find the rs^{th} entry in $A(BC)$.

The r^{th} row in A is $a_{r1}, a_{r2}, \ldots\ldots, a_{rm}$. The s^{th} column in BC consists of the elements $\sum b_{1j}c_{js}, \ldots\ldots, \sum b_{nj}c_{js}$. Hence the rs^{th} entry in $A(BC)$ is $a_{r1}\sum b_{1j}c_{js} + \cdots + a_{rn}\sum b_{nj}c_{js}$

$$= \sum_{i=1}^{n} a_{ri} \sum_{j=1}^{p} b_{ij}c_{js} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ri}b_{ij}c_{js}.$$

Let us now find the rs^{th} entry in $(AB)C$.

The r^{th} row in AB is

$\sum a_{ri}b_{i1}, \sum a_{ri}b_{i2}, \ldots\ldots, \sum a_{ri}b_{ip}$.

The s^{th} column in C is $c_{1s}, c_{2s}, \ldots\ldots, c_{ps}$.

Hence the rs^{th} entry in $(AB)C$ is

$$\left(\sum a_{ri}b_{i1}\right)c_{1s} + \left(\sum a_{ri}b_{i2}\right)c_{2s} + \cdots + \left(\sum a_{ri}b_{ip}\right)c_{ps} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ri}b_{ij}c_{js}$$

Thus $A(BC) = (AB)C$.

Theorem 7.2. Let U, V, W be vector spaces of dimensions m, n and p respectively over a field F with respective bases $\{u_1, u_2, \dots u_m\}$, $\{v_1, v_2, \dots, v_n\}$, and $\{w_1, w_2, \dots\dots, w_p\}$. Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations and $M(T_1)$ and $M(T_2)$ their corresponding matrices with respect to these bases.

Then $M(T_2 \circ T_1) = M(T_1)M(T_2)$.

Proof. $M(T_1)$ and is an $m \times n$ matrix and $M(T_2)$ is an $n \times p$ matrix. Hence the product $M(T_1)M(T_2)$ is defined and is an $m \times p$ matrix.

Let $M(T_1) = (a_{ij})$ and $M(T_2) = (b_{ij})$.

Then, $T_1(u_i) = \sum_{j=1}^{n} a_{ij} v_j$ and $T_2(v_j) = \sum_{k=1}^{p} b_{jk} w_k$.

$$\therefore \quad (T_2 \circ T_1)(u_i) = T_2\left(\sum_{j=1}^{n} a_{ij} v_j\right).$$

$$= \sum_{j=1}^{n} a_{ij} T_2(v_j)$$

$$= \sum_{j=1}^{n} a_{ij} \sum_{k=1}^{p} b_{jk}\, w_k$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{p} (a_{ij} b_{jk})(w_k)$$

Thus $M(T_2 \circ T_1) = M(T_1)M(T_2)$.

Note 1. Thus multiplication of two matrices is equivalent to the composition of their corresponding linear transformations in the reverse order. Since composition of linear transformation is associative we get matrix multiplication is associative.

Note 2. Let $M_n(F)$ denote the set of all square matrices of order n over the field F. Then matrix multiplication is an associative binary operation on $M_n(F)$. If $A, B, C \in M_n(F)$ the two distributive laws.

$A(B+C) = AB+AC$ and $(A+B)C = AC+BC$ can be verified.

Since $M_n(F)$ is already an abelian group under matrix addition we see that $M_n(F)$ is a ring.

Exercises

1. Using $A = \begin{bmatrix} 1 & -1 & 1 \\ 5 & 0 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ test the associative law $A(BC) = (AB)C$ for matrix multiplication.

2. Compute $(2 \;\; 1 \;\; -1) \begin{bmatrix} 4 & -1 & 2 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix}$

3. Find for what values of x will $(x \;\; 4 \;\; 1) \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ 4 \\ 1 \end{bmatrix} = 0$

4. Given that $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix} A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$ find the matrix A.

Answers.

2. (3) $3 \cdot x = -2 \pm i\sqrt{6}$

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

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

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

For example, if $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 0 & 1 \\ 0 & 3 & 1 & 5 \end{bmatrix}$ then

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

Clearly if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix.

Theorem 7.3. Let A and B be two $m \times n$ matrices. Then

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

Proof.

(i) The $(i, j)^{\text{th}}$ entry of $(A^T)^T$

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

$\therefore \quad (A^T)^T = A$

(ii) The $(i, j)^{\text{th}}$ entry of $(A+B)^T$

$$= (j, i)^{\text{th}} \text{ entry of } A+B$$
$$= (j, i)^{\text{th}} \text{ entry of } A + (j, i)^{\text{th}} \text{ entry of } B$$
$$= (i, j)^{\text{th}} \text{ entry of } A^T + (i, j)^{\text{th}} \text{ entry of } B^T$$
$$= (i, j)^{\text{th}} \text{ entry of } (A^T + B^T).$$

$\therefore \quad (A+B)^T = A^T + B^T$.

Theorem 7.4. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then $(AB)^T = B^T A^T$.

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

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

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

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

The $(i, j)^{\text{th}}$ entry of $AB = c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

$\therefore$ The $(i, j)^{\text{th}}$ entry of $(AB)^T = c_{ji} = \sum_{k=1}^{n} a_{jk} b_{ki}$

Now the i^{th} row of B^T is the i^{th} column of B and it consists of the elements $b_{1i}, b_{2i}, \ldots\ldots, b_{ni}$. Also the j^{th} column of A^T is the j^{th} row of A and it consists of the elements $a_{j1}, a_{j2}, \ldots\ldots, a_{jn}$. Hence the $(i, j)^{\text{th}}$ entry of $B^T A^T = b_{1i} a_{j1} + b_{2i} a_{j2} + \ldots\ldots + b_{ni} a_{jn}$.

$$= \sum_{k=1}^{n} b_{ki} a_{jk}$$
$$= (i, j)^{\text{th}} \text{ entry of } (AB)^T.$$

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

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

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

For example

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

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

Theorem 7.5. Let A and B be matrices with entries from C. Then

(i) $\overline{(\overline{A})} = A$.
(ii) $\overline{A+B} = \overline{A} + \overline{B}$
(iii) $\overline{kA} = \overline{k}\,\overline{A}$, where $k \in C$.
(iv) $A = \overline{A} \Leftrightarrow$ all entries of A are real.
(v) $\overline{AB} = \overline{A}\,\overline{B}$ provided AB is defined.
(vi) $(\overline{A})^T = \overline{A^T}$

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

Exercises

1. Let $A = \begin{bmatrix} 3 & 4 & 6 \\ -1 & 7 & 2 \\ 4 & 3 & 0 \end{bmatrix}$ and

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

Find $A^T, B^T, (A+B)^T, (AB)^T$ and $B^T A^T$.

2. Let $A = \begin{bmatrix} 2i & 3+4i & 0 \\ 1+i & 1-i & i \\ 3 & 2i & 4 \end{bmatrix}$ and

$B = \begin{bmatrix} 0 & 2 & 6 \\ -1 & 4 & 6 \\ 2 & 0 & 2 \end{bmatrix}$

Find $\overline{A}$, $\overline{A+B}$, $\overline{AB}$, $\overline{A}\,\overline{B}$, $\overline{A}^T$, $\overline{A^T}$, $\overline{B}^T$, $\overline{A}^T B$ and $A\overline{B}^T$

7.2. Types of Matrices

Definition. An $1 \times n$ matrix is calleds a ***row matrix***.

Thus a row matrix consists of 1 row and n columns.

It is of the form $(a_{11}, a_{12}, a_{13}, \ldots\ldots, a_{1n})$.

Definition. An $m \times 1$ matrix is called a ***column matrix***. Thus a column matrix consists of m rows and 1 column and it is of the form $\begin{bmatrix} a_{11} \\ a_{21} \\ \cdots \\ a_{m1} \end{bmatrix}$

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

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

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

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

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

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

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

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

For example, $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$ is lower triangular

$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 2 & 3 & 0 \\ 2 & 3 & 2 & 4 \end{bmatrix}$ is upper triangular.

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

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

Example.

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

Theorem 7.6. A square matrix A is symmetric iff $A = A^T$.

Proof. Let A be a symmetrix matrix.

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

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

Hence $A = A^T$.

Conversely let $A = A^T$.

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

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

Hence A is symmetric.

Theorem 7.7. Let A be any square matrix.

Then $A + A^T$ is symmetric.

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

$$= A^T + A$$

$$= A + A^T.$$

Hence $A + A^T$ is symmetric.

Theorem 7.8. Let A and B be symmetric matrices of order n. Then

(i) $A + B$ is symmetric.

(ii) AB is symmetric iff $AB = BA$.

(iii) $AB + BA$ is symmetric.

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

Proof.

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

$$= A + B \quad \text{(since } A \text{ and } B \text{ are symmetric)}$$

$\therefore$ $A + B$ is symmetric.

(ii) AB is symmetric

$\Leftrightarrow (AB)^T = AB$

$\Leftrightarrow B^T A^T = AB$ (by Theorem 7.4)

$\Leftrightarrow BA = AB.$

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

$$= B^T A^T + A^T B^T$$

$$= BA + AB \quad \text{(since } A \text{ and } B \text{ are symmetric)}$$

$$= AB + BA.$$

$\therefore$ $AB + BA$ is symmetric.

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

$\therefore$ kA is symmetric.

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

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

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

Theorem 7.9. A square matrix A is skew symmetric matrix iff $A = -A^T$.

Proof is similar to that of Theorem 7.6

Theorem 7.10. Let A be any square matrix. Then $A - A^T$ is skew symmetric.

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

$$= A^T - A$$

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

Hence $A - A^T$ is skew symmetric.

Theorem 7.11. Any square matrix A can be expressed uniquely as the sum of a symmetric matrix and a skew symmetric matrix.

Proof. Let A be any square matrix.

Then $A + A^T$ is a symmetric matrix (by Theorem 7.7)

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

Also, $\frac{1}{2}(A - A^T)$ is a skew symmetric matrix (by Theorem 7.10)

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

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

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

$$A = S + R \qquad \dots (1)$$

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

$$= S^T + R^T$$

$= S - R$ (since S is symmetric and R is skew symmetric) (2)

$\therefore\ A^T = S - R$ ....

From (1) and (2) we get

$$S = \tfrac{1}{2}(A + A^T) \text{ and } R = \tfrac{1}{2}(A - A^T).$$

Theorem 7.12. Let A and B be skew symmetric matrices of order n. Then

(i) $A + B$ is skew symmetric.

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

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

Proof. Let A, B be skew symmetric.

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

$= -A - B$ (by Theorem 7.9)

$= -(A + B)$.

$\therefore\ A + B$ is skew symmetric.

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

(iii) Let m be any positive integer.

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

$= A^T A^T \ldots A^T$ (m times)

$= (-A)(-A) \ldots (-A)$ (m times)

(since $A^T = -A$)

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

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

$\therefore\ A^m$ is symmetric when m is even and skew symmetric when m is odd.

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

Example. $\begin{bmatrix} 1 & -1+2i & 3+4i \\ -1-2i & -2 & 3 \\ 3-4i & 3 & 2 \end{bmatrix}$ is a Hermitian matrix.

$\begin{bmatrix} 0 & -a+ib \\ a+ib & 0 \end{bmatrix}$, $\begin{bmatrix} ib & c+id \\ -c+id & ib \end{bmatrix}$ are skew Hermitian matrices.

Note.

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

Theorem 7.13. Let A be a square matrix.

(i) A is Hermitian iif $A = \overline{A}^T$.

(ii) A is skew Hermitian iff $A = -\overline{A}^T$.

Proof. The result is an immediate consequence of the definition.

Theorem 7.14. Let A and B be square matrices of the same order. Then

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

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

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

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

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

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

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

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

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

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

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

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

$\therefore$ $A+B$ is Hermitian.

(iii) $\overline{(iA)}^T = (-i\overline{A})^T$

$= i\overline{A}^T$

$= iA$ (since A is Hermitian)

$\therefore$ iA is skew Hermitian.

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

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

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

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

$= BA + AB$

$= AB + BA$

$\therefore$ $AB+BA$ is Hermitian.

Theorem 7.15. Let A be any square matrix.

Then

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

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

Proof.

(i) Let $A + \overline{A}^T = B$.
Then $\overline{B} = \overline{A} + A^T$

$\therefore$ $\overline{B}^T = (\overline{A} + A^T)^T$.

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

$= B$.

Hence $A + \overline{A}^T$ is Hermitian.

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

Theorem 7.16. Any square matrix A can be uniquely expressed as the sum of a Hermitian matrix and a skew Hermitian matrix.

Proof. The proof is similar to that of Theorem 7.11.

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

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

Theorem 7.17. Let A and B be orthogonal matrices of the same order. Then

(i) A^T is orthogonal.

(ii) AB is orthogonal.

Proof. **(i)** $A^T(A^T)^T = A^TA = I$ (since A is orthogonal).

Similarly we can prove $(A^T)^TA^T = I$.

$\therefore$ A^T is orthogonal.

(ii)

$(AB)(AB)^T = (AB)(B^TA^T)$

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

$= AIA^T$ (since B is orthogonal)

$= AA^T$

$= I$.

Similarly $(AB)^T(AB) = I$.
Hence AB is orthogonal.

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

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

Note. Any unitary matrix over **R** is an orthogonal matrix.

Theorem 7.18. If A and B are unitary matrices of the same order, then AB is also an unitary matrix.

Proof. Similar to the proof of **(ii)** of Theorem 7.17.

Exercises

1. Give examples of each of the following types of matrices; upper triangular matrix, lower triangular matrix, diagonal matrix, scalar matrix, symmetric matrix, Hermitian matrix, skew Hermitian matrix, orthogonal matrix and unitary matrix.

2. Give examples of matrices over the field of complex numbers which are

   (a) symmetric but not Hermitian.

   (b) skew symmetric but not skew Hermitian.

3. Show that the product of two upper (lower) triangular matrices of the same order is again an upper (lower) triangular matrix.

4. Show that the product of two diagonal matrices of the same order is again a diagonal matrix.

5. Show that any two diagonal matrices of the same order commute.

6. For any square matrix A show that AA^T and A^TA are symmetric.

7. Show that if A is symmetric then A^T is symmetric.

8. Show that if A is skew symmetric then A^2 is symmetric and A^3 is skew symmetric.

9. Show that if A and B are symmetric matrices of the same order then $AB - BA$ is skew symmetric.

10. Show that if A and B are skew symmetric matrices then AB is symmetric iff $AB = BA$.

11. Show that any Hermitian matrix A can be written as $A = B + iC$ where B is a real symmetric matrix and C is a real skew symmetric matrix. State and prove a similar result for a skew Hermitian matrix.

12. Show that every square matrix A can be uniquely expressed as $A = B + iC$ where B and C are Hermitian.

13. A square matrix A is called an ***idempotent matrix*** if $A^2 = A$.

    Show that $\begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix}$ and $\begin{bmatrix} -1 & 3 & 5 \\ 1 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}$ are idempotent matrices.

14. Show that if $AB = A$ and $BA = B$ then A and B are idempotent matrices.

15. Show that if A is an idempotent matrix, then $B = I - A$ is also an idempotent matrix and $AB = BA = \mathbf{0}$.

16. A square matrix A is said to be ***nilpotent*** if $A^n = \mathbf{0}$ for some positive integer n. Show that $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{bmatrix}$ are nilpotent.

17. A square matrix A is said to be ***involutory*** if $A^2 = I$.

    Show that $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ are involutory.

18. Show that a square matrix A is involutory iff $(I + A)(I - A) = \mathbf{0}$

19. Show that $\frac{1}{2}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$ is an orthogonal matrix.

20. If (l_i, m_i, n_i) where $i = 1, 2, 3$ are the direction cosines of three mutually perpendicular lines referred to an orthogonal cartesian coordinate system, then show that $\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix}$ is an orthogonal matrix.

21. Show that $\frac{1}{2}\begin{bmatrix} l+i & -1+i \\ 1+i & 1-i \end{bmatrix}$ is an unitary matrix.

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

    Let A and B be square matrices of order n.

    (a) A, B are symmetric $\Rightarrow AB$ is symmetric.

(b) A, B are skew symmetric $\Rightarrow AB$ is skew symmetric.

(c) A, B are upper triangular matrices $\Rightarrow$ AB is upper triangular.

(d) A, B are lower triangular matrices $\Rightarrow$ AB is lower triangular.

(e) A, B are diagonal matrices $\Rightarrow AB$ is a diagonal matrix.

(f) A, B are scalar matrices $\Rightarrow AB$ is a scalar matrix.

(g) Conjugate of a symmetric matrix is symmetric.

(h) Conjugate of a skew symmetric matrix is skew symmetric.

(i) Conjugate of a Hermitian matrix is Hermitian.

(j) Conjugate of a skew Hermitian matrix is skew Hermitian.

(k) Any real symmetric matrix is Hermitian.

Answers.

22. (a) F (b) F (c) T (d) T
(e) T (f) T (g) T (h) T
(i) T (j) T (k) T

7.3. The Inverse of a Matrix

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

Determinants. We can associate with any $n \times n$ matrix $A = (a_{ij})$ over a field F an element of F given by the

$$\text{determinant} \begin{vmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ a_{n1} & a_{n2} & \cdots\cdots & a_{nn} \end{vmatrix}$$

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

For example,

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

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

$$|A| = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = 1.$$

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

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

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

Hence if A and B are two $n \times n$ matrices

$|AB| = |A||B|$.

Theorem 7.19. The product of any two non-singular matrices is non-singular.

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

$$\therefore \quad |AB| = |A||B| \neq 0.$$

Hence AB is non-singular.

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

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

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

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

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

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

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

$$M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}.$$

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

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

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

Thus the $(i, j)^{\text{th}}$ entry of $adj\ A$ is A_{ji}.

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

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

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

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

Similarly other co-factors can be calculated and we get

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

We notice that

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

$$= \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix}$$

$= (adj\ A)A.$ (verify)

Exercises

1. Write down six square matrices A and calculate $adj\ A$, $A(adj\ A)$ and $(adj\ A)A$.
2. Prove that $adj\ A^T = (adj\ A)^T$.
3. If A is symmetric prove that $adj\ A$ is symmetric.

Theorem 7.20. Let A be any square matrix of order n. Then $(adj\ A)A = A(adj\ A) = |A|I$ where I is the identity matrix of order n.

Proof. The $(i, j)^{\text{th}}$ element of$(A(adj\ A))$

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

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

$$\therefore \quad A(adj\ A) = \begin{bmatrix} |A| & 0 & \cdots\cdots & 0 \\ 0 & |A| & \cdots\cdots & 0 \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ 0 & 0 & \cdots\cdots & |A| \end{bmatrix}$$

$$= |A|I.$$

Similarly, $(adj\ A)A = |A|I$.

Hence $(adj\ A)A = A(adj\ A) = |A|I$.

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

$$\text{Then} \quad AB = A\left[\frac{1}{|A|}(adj\ A)\right]$$

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

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

$$= I.$$

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

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

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

Theorem 7.21. A square matrix A of order n is non-singular iff A is invertible.

Proof. Suppose A is invertible.

Then there exists a matrix B such that

$AB = BA = I$.

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

$\therefore \; |A|\,|B| = 1$.

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

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

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

Then $AB = BA = I$. (refer the note above)

$\therefore$ A is invertible and B is the inverse of A.

Solved problems

Problem 1. Compute the inverse of the matrix

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

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

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

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

Now, we find $$adj\, A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

where A_{ij}, $(i, j = 1, 2, 3)$ are cofactors of a_{ij}.

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

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

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

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

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

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

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

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

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

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

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

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

Problem 2. If $\omega = e^{2\pi i/3}$ find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}$$

Solution. We note that $\omega^3 = 1$.

$\therefore \quad |A| = 3(\omega^2 - \omega)$. (verify)

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

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

(verify)

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

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

Problem 3. Show that a square matrix A is orthogonal iff $A^{-1} = A^T$.

Solution. Suppose A is orthogonal. Then $A\,A^T = I$.

$\therefore \quad |A\,A^T| = |I| = 1.$

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

$\therefore \quad |A||A| = 1.$

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

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

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

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

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

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

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

Then $A\,A^T = A\,A^{-1} = I$. Similarly $A^T A = I$.

Hence A is orthogonal.

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

Solution. Suppose A is involutory. Then $A^2 = I$. Hence $|A^2| = 1$.

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

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

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

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

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

$\therefore \quad IA = A^{-1}.$

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

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

Then $A^2 = AA = AA^{-1} = I$.

$\therefore \quad A$ is involutory.

Exercises

1. Compute the inverse of each of the following matrices.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 2 & 2 & -3 \\ -3 & 2 & 2 \\ 2 & -3 & 2 \end{bmatrix}$

(d) $\begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. Show that the set of all non-singular matrices of order n over a field F is a group under matrix multiplication.
3. If A and B are non-singular matrices of order n prove that $(AB)^{-1} = B^{-1}A^{-1}$.
4. If A is a non-singular symmetric matrix prove that A^{-1} is also a symmetric matrix.
5. If A is a non-singular matrix, prove that $(A^T)^{-1} = (A^{-1})^T$.
6. If A is orthogonal, prove that A^{-1} is orthogonal.
7. Determine which of the following statements are true and which are false. Let A, B and C be square matrices of order n. Then
   (a) A, B are non-singular $\Rightarrow AB$ is non-singular.
   (b) A, B are non-singular $\Rightarrow A + B$ is non-singular.
   (c) A, B are singular $\Rightarrow AB$ is singular.
   (d) A is singular, B is non-singular $\Rightarrow$ AB is singular.
   (e) A is non-singular, B singular AB is singular.

(f) AB is non-singular $\Rightarrow$ BA is non-singular.

(g) $AB = AC \Rightarrow B = C$.

(h) $AB = AC$ and A non-singular $\Rightarrow$ $B = C$.

(i) $AB = 0 \Rightarrow A$ and B are singular.

(j) $A(B + C) = AB + AC$.

Answers.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ (b) $-\frac{1}{31}\begin{bmatrix} -9 & 4 & 11 \\ -8 & 7 & -4 \\ -2 & -6 & -1 \end{bmatrix}$

(c) $\frac{1}{5}\begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$ (d) $\begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$

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

(g) F (h) T (i) F (j) T.

In the following theorem we bring out the connection between non- singular linear transformations and non-singular matrices.

Theorem 7.22. Let V and W be vector spaces of dimension n over a field F with bases $v_1, v_2, \ldots\ldots, v_n$ and $w_1, w_2, \ldots\ldots\ldots, w_n$ respectively. Then a linear transformation $T : V \to W$ is non-singular iff the associated matrix is non- singular.

Proof. Let $T : V \to W$ be a non-singular linear transformation.

Then T is 1-1 and onto.

Hence $T^{-1} : W \to V$ is also a linear transformation.

Let A and B be the matrices representing the linear transformations T and T^{-1} with respect to the chosen bases.

By theorem 7.2, multiplication of the matrices A and B is equivalent to the composition of the corresponding linear transformation T and T^{-1}.

Also $T \circ T^{-1}$ and $T^{-1} \circ T$ are identity transformations.

Hence $AB = BA = I$. Thus A has an inverse B.

Hence A is non-singular.

Conversely, let A be a non-singular matrix. Then A^{-1} exists.

Let $S : W \to V$ be the linear transformation determined by the matrix A^{-1}.

It is easily verified that $T \circ S = S \circ T = I$

Hence T has an inverse linear transformation S.

Hence T is a non-singular linear transformation.

7.4. Elementary Transformations

Definition. Let A be an $m \times n$ matrix over a field F. An elementary row-operation on A is of any one of the following three types.

1. *The interchange of any two rows.*
2. *Multiplication of a row by a non-zero element c in F.*
3. *Addition of any multiple of one row with any other row.*

Similarly we define an elementary column operation on A as any one of the following three types.

1. *The interchange of any two columns.*
2. *Multiplication of a column by a non-zero element c in F.*
3. *Addition of any multiple of one column with any other column.*

Example. Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & -1 \end{bmatrix}$, $A_1 = \begin{bmatrix} 3 & -1 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}$,

$A_2 = \begin{bmatrix} 2 & 2 \\ 4 & 1 \\ 6 & -1 \end{bmatrix}$ $A_3 = \begin{bmatrix} 1 & 2 \\ 5 & 7 \\ 3 & -1 \end{bmatrix}$. A_1 is obtained from A by interchanging the first and third rows.

A_2 is obtained from A by multiplying the first column of A by 2.

A_3 is obtained from A by adding to the second row the multiple by 3 of the first row.

Notation. We shall employ the following notations for elementary transformations.

(i) Interchange of i^{th} and j^{th} rows will be denoted by $R_i \leftrightarrow R_j$.

(ii) Multiplication of i^{th} row by a non-zero element $c \in F$ will be denoted by $R_i \to cR_i$.

(iii) Addition of k times the j^{th} row to the i^{th} row will be denoted by $R_i \to R_i + kR_j$.

The corresponding column operations will be denoted by writing C in the place of R.

Definition. An $m \times n$ matrix B is said to be ***row equivalent (column equivalent)*** to an $m \times n$ matrix A if B can be obtained from A by a finite succession of elementary row operations (column operations).

A and B are said to be ***equivalent*** if B can be obtained from A by a finite succession of elementary row *or* column operations.

If A and B are equivalent. We write $A \sim B$.

Exercise. Prove that row equivalence, column equivalence and equivalence are equivalence relations in the set of all $m \times n$ matrices.

Definition. A matrix obtained form the identity matrix by applying a single elementary row or column operation is called an ***elementary matrix***.

For example, $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ $\begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$ are elementary matrices obtained from the identity matrix $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ by applying the elementary operations $R_1 \leftrightarrow R_2$, $R_1 \to 4R_1$, $R_3 \to R_3 + 2R_2$ respectively.

Exercise. Give examples of elementary matrices of order 4.

Theorem 7.23. Any elementary matrix is non-singular.

Proof. The determinant of the identity matrix of any order is 1. Hence the determinant of an elementary matrix obtained by interchanging any two rows is -1. The determinant of an elementary matrix obtained by multiplying any row by $k \neq 0$ is k. The determinant of an elementary matrix obtained by adding a multiple of one row with another row is 1. Hence any elementary matrix is non-singular.

Theorem 7.24. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then every elementary row (column) operation of the product AB can be obtained by subjecting the matrix A (matrix B) to the same elementary row (column) operation.

Proof. Let $R_1, R_2, \ldots\ldots, R_m$ denote the rows of the matrix A and $C_1, C_2, \ldots\ldots, C_p$ denote the columns of B. By the definition of matrix multiplication

$$AB = \begin{bmatrix} R_1C_1 & R_1C_2 & \cdots\cdots & R_1C_p \\ R_2C_1 & R_2C_2 & \cdots\cdots & R_2C_p \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ R_mC_1 & R_mC_2 & \cdots\cdots & R_mC_p \end{bmatrix}$$

It is obvious from the above representation of AB that if we apply any elementary row operation on A the matrix AB is also subjected to the same elementary row operation. Also if we apply any elementary column operation on B the matrix AB is also subjected to the same elementary column operation.

Theorem 7.25. Each elementary row operation on an $m \times n$ matrix A is equivalent to *pre-multiplying* the matrix A by the corresponding elementary $m \times m$ matrix.

Proof. Since A is an $m \times n$ matrix we can write

$A = IA$ where I is the identity matrix of order m. By theorem 7.24 an elementary row operation IA is equivalent to the same row operation on I. But an elementary row operation on I gives an elementary matrix. Hence by pre-multiplying A by the corresponding elementary matrix we get the required row operation on A.

Note. Similarly each elementary column operation of an $m \times n$ matrix A is equivalent to post-multiplying the matrix A by the corresponding elementary $n \times n$ matrix.

Corollary 1. If two $m \times n$ matrices A and B are row equivalent then $A = PB$ where P is a non-singular $m \times m$ matrix.

Proof. Since A is row equivalent to B, A can be obtained from B by applying successive elementary row operations. Hence $A = E_1 E_2 \ldots E_n B$ where each E_i is an elementary matrix. Since each E_i is non-singular, $A = PB$ where $P = E_1 E_2 \ldots E_n$ and P is non-singular.

Corollary 2. If two matrices A and B are column equivalent then $A = BQ$ where Q is a non-singular matrix.

Corollary 3. If two $m \times n$ matrices A and B are equivalent then $A = PBQ$ where P is a non-singular $m \times m$ matrix and Q is a non- singular $n \times n$ matrix.

Corollary 4. The inverse of an elementary matrix is again an elementary matrix.

Proof. Let E be an elementary matrix obtained from I by applying some elementary operations. If we apply the reverse operation on E, then E is carried back to I. Let E^* be the elementary matrix corresponding to the reverse operation.

Then $E^*E = EE^* = I$. Hence $E^* = E^{-1}$.

Hence E^{-1} is also an elementary matrix.

Canonical form of a matrix. We now use elementary row and column operations to reduce any matrix to a simple form, called the *canonical form of a matrix.*

Theorem 7.26. By successive applications of elementary row and column operations, any non-zero $m \times n$ matrix A can be reduced to a diagonal matrix D in which the diagonal entries are either 0 or 1 and all the 1's proceeding all the zeros on the diagonal. In other words, any non-zero $m \times n$ matrix is equivalent to a matrix of the form $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ where I_r is the $r \times r$ identity matrix and O is the zero matrix.

Proof. We shall prove the theorem by induction on the number of rows of A. Suppose A has just one row. Let $A = (a_{11} a_{12} \ldots \ldots a_{1n})$.

Since $A \neq \mathbf{0}$, by interchanging columns, if necessary, we can bring a non-zero entry c to the position a_{11}.

Multiplying A by c^{-1} we get 1 as the first entry.

Other entries in A can be made zero by adding suitable multiples of 1. Thus the result is true when $m = 1$.

Now, suppose that the result is true for any non-zero matrix with $m - 1$ rows.

Let A be a non-zero $m \times n$ matrix. By permuting rows and columns we can bring some non-zero entry c to the position a_{11}.

Multiplying the first row by c^{-1} we get 1 as the first entry.

All other entries in the first column can be made zero by adding suitable multiples of the first row to each other row.

Similarly all the other entries in the first row can be made zero.

This reduces A to a matrix of the form

$B = \begin{pmatrix} I_1 & O \\ O & C \end{pmatrix}$ where C is an $(m-1) \times (n-1)$ matrix.

Now by induction hypothesis C can be reduced to the desired form by elementary row and column operations.

Hence A is equivalent to a matrix of the required form.

Corollary 1. If A is an $m \times n$ matrix there exist non-singular square matrices P and Q of orders m and n respectively such that $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$

The result follows from corollary 3 of theorem 7.25.

Corollary 2. Any non-singular square matrix A of order n is equivalent to the identity matrix.

Proof. By corollary 1, $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.

Since P, A, Q are all non-singular $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$ is non-singular. This is possible iff $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} = I_n$.

Corollary 3. Any non-singular matrix A can be expressed as a product of elementary matrices.

Proof. By corollary 2, $PAQ = I_n$. Hence $A = P^{-1}Q^{-1}$. Further by corollary 4 of theorem 7.25, P^{-1} and Q^{-1} are products of elementary matrices.

Hence A is a product of elementary matrices.

Note. The inverse of a non-singular matrix A can be computed by using elementary transformations. Let A be a non-singular matrix of order n. Then $AA^{-1} = A^{-1}A = I$. Now, the non-singular matrix A^{-1} can be expressed as the product of elementary matrices.

Let $A^{-1} = E_1E_2 \ldots\ldots E_n$.

Then $I = A^{-1}A = E_1E_2 \ldots\ldots E_nA$.

Thus every non-singular matrix A can be reduced to I by pre-multiplying A by elementary matrices.

Hence A can be reduced to the identity matrix by applying successive elementary row operations.

Now, $A = IA$. Reduce the matrix A in the left hand side to I by applying successive elementary row operations and apply the same elementary row operations to the factor I in the right hand side.

Then we get $I = BA$ so that $B = A^{-1}$.

Solved problems

Problem 1. Reduce the matrix $A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{bmatrix}$ to the canonical form.

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

$$\sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{array}{l} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - 2R_1 \end{array}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{array}{l} C_2 \to C_2 - 2C_1 \\ C_3 \to C_3 + C_1 \end{array}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad C_3 \to C_3 + 3C_2$$

$$\sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad R_2 \to -R_2$$

Problem 2. Find the inverse of the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{bmatrix}$

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

$$\Rightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 1 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} A,$$

$$\begin{array}{l} R_2 \to R_2 - 3R_1 \\ R_3 \to R_3 + 2R_1 \end{array}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 0 & 14 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 5 & -1 & 1 \end{bmatrix} A,$$

$$R_3 \to R_3 - R_2$$

$$\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{bmatrix}$$

$$\begin{array}{l} R_1 \to R_1 - \frac{1}{7}R_3 \\ R_2 \to R_2 + \frac{1}{2}R_3 \\ R_3 \to \frac{1}{14}R_3 \end{array}$$

$$\Rightarrow A^{-1} = \begin{bmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{bmatrix}$$

Exercises

1. Write down six matrices (not necessarily square matrices) and reduce them to the canonical form.
2. Find the inverse of the following matrices by using elementary operations.

(a) $\begin{bmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Answers.

2. (a) $\begin{bmatrix} -9 & 8 & -5 \\ -8 & 7 & -4 \\ -2 & 2 & -1 \end{bmatrix}$ (b) $\frac{1}{2}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$

Definition. Let A and B be two square matrices of order n. B is said to be ***similar*** to A if there exists a $n \times n$ non-singular matrix P such that $B = P^{-1}AP$.

Solved problems

Problem 1. Similarity of matrices is an equivalence relation in the set of all $n \times n$ matrices.

Proof. Let S be the set of all $n \times n$ matrices.

Let $A \in S$.

Since $A = I^{-1}AI$ and I is non-singular, A is similar to A.

Hence similarity of matrices is reflexive.

Now, let $A, B \in S$ and let A be similar to B.

$\therefore$ $A = P^{-1}BP$ where $P \in S$ is a non-singular matrix.

Now, $$P^{-1}BP = A \Rightarrow PP^{-1}B\,PP^{-1} = PA\,P^{-1}$$
$$\Rightarrow B = PA\,P^{-1}$$
$$\Rightarrow B = (P^{-1})^{-1}A(P^{-1}).$$

Since P is non-singular $P^{-1} \in S$ is also non-singular.

$\therefore$ B is similar to A.

Hence similarity of matrices is symmetric.

Now, let $A, B, C \in S$.

Let A be similar to B to B be similar to C. Hence there exist non-singular matrices $P, Q \in S$ such that

$$A = P^{-1}BP \text{ and } B = Q^{-1}CQ.$$

Now, $$\begin{aligned} A &= P^{-1}BP \\ &= P^{-1}(Q^{-1}CQ)P \\ &= (P^{-1}Q^{-1})CQP \\ &= (QP)^{-1}C(QP). \end{aligned}$$

Since $P, Q \in S$ are non-singular, $QP \in S$ is also non-singular.

Hence A is similar to C.

$\therefore$ Similarity of matrices is transitive.

Hence similarity of matrices is an equivalence relation.

Problem 2. If A and B are similar matrices show that their determinants are same.

Solution. Let A and B be two similar matrices.

$\therefore$ There exists a non-singular matrix P such that $B = P^{-1}AP$.

Now, $$\begin{aligned} |B| &= |P^{-1}AP| \\ &= |P^{-1}||A||P| \\ &= |A| \quad (\text{since } |P^{-1}| = \frac{1}{|P|}) \end{aligned}$$

Hence the result.

7.5. Rank of a Matrix

We now proceed to introduce the concept of the rank of a matrix.

Definition. Let $A = (a_{ij})$ be an $m \times n$ matrix. The rows $R_i = (a_{i1}, a_{i2}, \ldots\ldots\ldots, a_{in})$ of A can be thought of as elements of F^n. The subspace of F^n generated by the m rows of A is called the ***row space*** of A.

Similarly, the subspace of F^m generated by the n columns of A is called the ***column space*** of A.

The dimension of the row space (column space) of A is called the ***row rank (column rank)*** of A.

Theorem 7.27. Any two row equivalent matrices have the same row space and have the same row rank.

Proof. Let A be an $m \times n$ matrix.

It is enough if we prove that the row space of A is not altered by any elementary row operation.

Obviously the row space of A is not altered by an elementary row operation of the type $R_i \leftrightarrow R_j$.

Now, consider the elementary row operation

$R_i \to cR_i$ where $c \in F - \{0\}$.

Since $L(\{R_1, R_2, \ldots\ldots, R_i, \ldots\ldots\ldots, R_n\}) = L(\{R_1, R_2, \ldots\ldots\ldots cR_i, \ldots\ldots, R_n\})$ the row space of A is not altered by this type of elementary row operation.

Similarly we can easily prove that the row space of A is not altered by an elementary row operation of the type $R_i \to R_i + cR_i$.

Hence row equivalent matrices have the same row space and hence the same row rank.

Similarly we can prove the following theorem.

Theorem 7.28. Any two column equivalent matrices have the same column rank.

Theorem 7.29. The row rank and the column rank of any matrix are equal.

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

Let $R_1, R_2, \ldots\ldots\ldots, R_m$ denote the rows of A.

Hence $R_i = (a_{i1}, a_{i2}, \ldots\ldots\ldots, a_{in})$.

Suppose the row rank of A is r.

Then the dimension of the row space is r.

Let $v_1 = (b_{11}, \ldots, b_{1n}), v_2 = (b_{21}, \ldots, b_{2n}), \ldots \ldots, v_r = (b_{r1}, \ldots\ldots\ldots, b_{rn})$ be a basis for the row space of A.

Then each row is a linear combination of the vectors $v_1, v_2, \ldots\ldots\ldots, v_r$.

$$\begin{aligned}
\text{Let, } R_1 &= k_{11}v_1 + k_{12}v_2 + \cdots\cdots + k_{1r}v_r \\
R_2 &= k_{21}v_1 + k_{22}v_2 + \cdots\cdots + k_{2r}v_r \\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
R_m &= k_{m1}v_1 + k_{m2}v_2 + \cdots\cdots + k_{mr}v_r
\end{aligned}$$

where $k_{ij} \in F$.

Equating the i^{th} component of each of the above equations, we get

$$\begin{aligned}
a_{1i} &= k_{11}b_{1i} + k_{12}b_{2i} + \cdots\cdots + k_{1r}b_{ri} \\
a_{2i} &= k_{21}b_{1i} + k_{22}b_{2i} + \cdots\cdots + k_{2r}b_{ri} \\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
&\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \\
a_{mi} &= k_{m1}b_{1i} + k_{m2}b_{2i} + \cdots\cdots + k_{mr}b_{ri}
\end{aligned}$$

Hence

$$\begin{bmatrix} a_{1i} \\ \cdot \\ \cdot \\ \cdot \\ a_{mi} \end{bmatrix} = b_{1i}\begin{bmatrix} k_{11} \\ \cdot \\ \cdot \\ \cdot \\ k_{m1} \end{bmatrix} + b_{2i}\begin{bmatrix} k_{12} \\ \cdot \\ \cdot \\ \cdot \\ k_{m2} \end{bmatrix} + \cdots + b_{ri}\begin{bmatrix} k_{1r} \\ \cdot \\ \cdot \\ \cdot \\ k_{mr} \end{bmatrix}$$

Thus each column of A is a linear combination of r vectors.

Hence the dimension of the column space $\le r$.

$\therefore$ Column rank of $A \le r =$ row rank of A.

Similarly, row rank of $A \le$ column rank of A.

Hence the row rank and the column rank of A are equal.

Definition. The ***rank*** of a matrix A is the common value of its row and column rank.

Note 1. Since the row rank and the column rank of a matrix are unaltered by elementary row and column operations, *equivalent matrices have the same rank.* In particular if a matrix A is reduced to its canonical form $\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$, then *rank of* $A = r$.

Thus to find the rank of a matrix A, we reduce A to the canonical form and find the number of non-zero entries in the diagonal.

Note that in the canonical form of the matrix A, there exists an $r \times r$ sub-matrix, namely, I_r, whose determinant is not zero.

Further every $(r+1)\times(r+1)$ sub-matrix contains a row of zeros and hence its determinant is zero.

Also under any elementary row or column operation the value of a determinant is either unaltered or multiplied by a non-zero constant.

Hence the matrix A is also such that

(i) there exists an $r \times r$ sub-matrix whose determinant is nonzero.

(ii) the determinant of every $(r+1)\times(r+1)$ sub-matrix is zero.

Hence one can also define the rank of a matrix A to be r if A satisfies (i) and (ii).

Note 2. Any non-singular matrix of order n is equivalent to the identity matrix and hence its rank is n.

Note 3. The rank of a matrix is not altered on multiplication by non-singular matrices, since premultiplication by a non-singular matrix is equivalent to applying elementary row operaptions and post-multiplication by a non- singular matrix is equivalent to applying elementary column operations.

Solved problems

Problem 1. Find the rank of the matrix

$$A=\begin{bmatrix}4&2&1&3\\6&3&4&7\\2&1&0&7\end{bmatrix}.$$

Solution.

$$A=\begin{bmatrix}4&2&1&3\\6&3&4&7\\2&1&0&7\end{bmatrix}$$

$$\sim\begin{bmatrix}1&2&4&3\\4&3&6&7\\0&1&2&7\end{bmatrix} \quad C_1 \leftrightarrow C_3$$

$$\sim\begin{bmatrix}1&0&0&0\\4&-5&-10&-5\\0&1&2&7\end{bmatrix} \quad \begin{matrix}C_1 \to C_2-2C_1\\ C_3 \to C_3-4C_1\\ C_4 \to C_4-3C_1\end{matrix}$$

$$\sim\begin{bmatrix}1&0&0&0\\0&-5&-10&-5\\0&1&2&7\end{bmatrix} \quad R_2 \to R_2-4R_1$$

$$\sim\begin{bmatrix}1&0&0&0\\0&-5&0&0\\0&1&0&6\end{bmatrix} \quad \begin{matrix}C_3 \to C_3-2C_2\\ C_4 \to C_4-C_2\end{matrix}$$

$$\sim\begin{bmatrix}1&0&0&0\\0&-5&0&0\\0&0&0&6\end{bmatrix} \quad R_3 \to R_3+\frac{1}{5}R_2$$

$$\sim\begin{bmatrix}1&0&0&0\\0&-5&0&0\\0&0&6&0\end{bmatrix} \quad C_2 \leftrightarrow C_3$$

$$\sim\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\end{bmatrix} \quad \begin{matrix}R_2 \to -\frac{1}{5}R_2\\ R_3 \to \frac{1}{6}R_3\end{matrix}$$

$\therefore$ Rank of $A=3$.

Problem 2. Find the rank of the matrix

$$A=\begin{bmatrix}1&1&1&1\\4&1&0&2\\0&3&4&2\end{bmatrix}$$ by examining the determinant minors.

Solution.

$$\begin{vmatrix}1&1&1\\4&1&0\\0&3&4\end{vmatrix}=0=\begin{vmatrix}1&1&1\\1&0&2\\3&4&2\end{vmatrix}$$

$$\begin{vmatrix}1&1&1\\4&1&2\\0&3&2\end{vmatrix}=0=\begin{vmatrix}1&1&1\\4&0&2\\0&4&2\end{vmatrix}$$

$\therefore$ Every 3×3 submatrix of A has determinant zero.

Also, $\begin{vmatrix}1&1\\4&1\end{vmatrix}=-3\neq 0.$

$\therefore$ Rank of $A=2$.

Exercises

1. Determine the rank of any six matrices of your choice.

2. Find the rank of the following matrices,

(a) $\begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & -3 \\ 6 & -1 & 1 \end{bmatrix}$ (b) $\begin{bmatrix} 0 & 1 & 2 & 1 \\ 2 & -3 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & -1 & 0 & 2 & 1 \\ 3 & 1 & 1 & -1 & 2 \\ 4 & 0 & 1 & 0 & 3 \\ 9 & -1 & 2 & 3 & 7 \end{bmatrix}$

3. Find the column rank of the matrices

(a) $\begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{bmatrix}$ (b) $\begin{bmatrix} 3 & 1 & -5 & -1 \\ 1 & -2 & 1 & -5 \\ 1 & 5 & -7 & 2 \end{bmatrix}$

(Hint: Row rank = rank of the matrix = column rank)

4. Find the row rank of the matrix

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

Answers. 2.(a) 2 (b) 3 (c) 3 3.(a) 2 (b) 3 4. 2

7.6. Simultaneous Linear Equations

In this section we shall apply the theory of matrices developed in the preceeding sections to study the existence of solutions of simultaneous linear equations.

Matrix form of a set of linear equations.

Consider a system of m linear equations in n unknowns $x_1, x_2, \ldots\ldots\ldots, x_n$ given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots\cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots\cdots + a_{2n}x_n &= b_2 \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad & \ldots \quad \ldots \quad \ldots \\ \ldots \quad \ldots \quad \ldots \quad \ldots \quad & \ldots \quad \ldots \quad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots\cdots + a_{mn}x_n &= b_m \end{aligned}$$

Using the concept of matrix multiplication and equality of matrices this system can be written as $AX = B$ where,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1n} \\ a_{21} & a_{22} & \cdots\cdots & a_{2n} \\ \cdot & \cdot & \cdots\cdots & \cdot \\ \cdot & \cdot & \cdots\cdots & \cdot \\ a_{m1} & a_{m2} & \cdots\cdots & a_{mn} \end{bmatrix},$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_m \end{bmatrix}$$

The $m \times n$ matrix A is called the ***coefficient matrix***.

Definition. A set of values of $x_1, x_2, \ldots\ldots\ldots, x_n$ which satisfy the above system of equations is called a **solution** of the system. The system of equations is said to be ***consistent*** if it has at least one solution. Otherwise the system is said to be ***inconsistent***.

The $m \times (n+1)$ matrix given by

$$\begin{bmatrix} a_{11} & \cdots\cdots & a_{1n} & b_1 \\ a_{21} & \cdots\cdots & a_{2n} & b_2 \\ \cdot & \cdots\cdots & \cdot & \cdot \\ \cdot & \cdots\cdots & \cdot & \cdot \\ a_{m1} & \cdots\cdots & a_{mn} & b_m \end{bmatrix}$$

is called the ***augmented matrix*** of the system and is denoted by (A, B).

Thus the augmented matrix (A, B) is obtained by annexing to A the column matrix B, which becomes the $(n+1)^{\text{th}}$ column in (A, B).

Note. Since every column in A appears in (A, B) the column space of the matrix A is a subspace of the column space of the matrix (A, B).

Hence the rank of $A \leq$ rank of (A, B).

Theorem 7.30. The system of linear equations $AX = B$ is consistent iff rank of A = rank of (A, B).

Proof. Let the system be consistent.

Let $u_1, u_2, \ldots\ldots, u_n$ be a solution of the system.

Then $B = u_1C_1 + u_2C_2 + \cdots\cdots + u_nC_n$ where $C_1, C_2, \ldots\ldots, C_n$ denote the columns of A.

$$\begin{matrix} a_{12} & \cdots\cdots & a_{1n} \\ a_{22} & \cdots\cdots & a_{2n} \\ \cdot & \cdots\cdots & \cdot \\ \cdot & \cdots\cdots & \cdot \\ a_{m2} & \cdots\cdots & a_{mn} \end{matrix}\Bigg],$$

$$, B = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_m \end{bmatrix}$$

is called the ***coefficient matrix***.

f values of $x_1, x_2, \ldots\ldots, x_n$
ve system of equations is called
em. The system of equations is
if it has at least one solution.
is said to be ***inconsistent***.
atrix given by

$$\begin{matrix} \ldots & a_{1n} & b_1 \\ \ldots & a_{2n} & b_2 \\ \ldots & \cdot & \cdot \\ \ldots & \cdot & \cdot \\ \ldots & a_{mn} & b_m \end{matrix}\Bigg]$$

matrix of the system and is

atrix (A, B) is obtained by
n matrix B, which becomes
$A, B)$.

nn in A appears in (A, B)
atrix A is a subspace of the
(A, B).
ank of (A, B).

of linear equations
iff rank of A = rank of

onsistent.
a solution of the system.
$+ \ldots\ldots + u_n C_n$ where
e columns of A.

Hence the column space of the augmented matrix (A, B), namely $\langle C_1, C_2, \ldots\ldots, C_n, B\rangle$ is the same as the column space $\langle C_1, C_2, \ldots\ldots, C_n\rangle$ of A.

Hence the rank of A = rank of (A, B).

Conversely let rank of A = rank of (A, B).

Then the column rank of A = column rank of (A, B).

$$dim\langle C_1, C_2, \cdots, C_n\rangle = dim\langle C_1, C_2, \ldots, C_n, B\rangle.$$

But $\langle C_1, C_2, \ldots\ldots, C_n\rangle$ is a subspace of $\langle C_1, C_2, \ldots\ldots, C_n, B\rangle$.

$\therefore$ B is a linear combination of $C_1, C_2, \ldots\ldots, C_n$.

If $B = u_1C_1 + \ldots + u_nC_n$ then $u_1, u_2, \ldots\ldots, u_n$ is a solution of the system.

Hence the theorem.

Remark. The solution of a given system of simultaneous equations is not altered by interchanging any two equations or by multiplying any equation by a non-zero constant or by adding a multiple of one equation to another. Hence we can reduce the given system of equations to an equivalent system by applying elementary row operations to the augmented matrix. This reduced form will enable us to test for the consistency and to find the solution if it exists. This is illustrated in the following problems.

Solved problems

Problem 1. Show that the equations

$$x + y + z = 6$$
$$x + 2y + 3z = 14$$
$$x + 4y + 7z = 30$$

are consistent and solve them.

Solution. The given system of equations can be put in the matrix form

$$AX = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 7 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \\ 30 \end{bmatrix} = B.$$

The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 14 \\ 1 & 4 & 7 & 30 \end{bmatrix}$$
$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 3 & 6 & 24 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$
$$\sim \begin{bmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{bmatrix} R_3 \to R_3 - 3R_2$$

Hence rank of A = rank of $(A, B) = 2$.

Hence the given system is consistent.

Also the given system of equations reduces to

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 0 \end{bmatrix}$$

$$\therefore \quad x + y + z = 6$$
$$y + 2z = 8.$$

Putting $z = c$ we obtain the general solution of the system as $x = c - 2, y = 8 - 2c, z = c$.

Problem 2. Verify whether the following system of equations is consistent. If it is consistent, find the solution.

$$x - 4y - 3z = -16$$
$$4x - y + 6z = 16$$
$$2x + 7y + 12z = 48$$
$$5x - 5y + 3z = 0.$$

Solution. The matrix form of the system is given by

$$\begin{bmatrix} 1 & -4 & -3 \\ 4 & -1 & 6 \\ 2 & 7 & 12 \\ 5 & -5 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 16 \\ 48 \\ 0 \end{bmatrix}$$

$\therefore$ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & -4 & -3 & -16 \\ 4 & -1 & 6 & 16 \\ 2 & 7 & 12 & 48 \\ 5 & -5 & 3 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 2R_1 \\ R_4 \to R_4 - 5R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} R_3 \to R_3 - R_2 \\ R_4 \to R_4 - R_2 \end{matrix}$$

$\therefore$ Rank of A = Rank of $(A, B) = 2$ and hence the system is consistent. Also the system of equations reduces to

$$\begin{bmatrix} 1 & -4 & -3 \\ 0 & 15 & 18 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -16 \\ 80 \\ 0 \\ 0 \end{bmatrix}$$

$\therefore$ $x - 4y - 3z = -16$ and $15y + 18z = 80$.

Putting $z = c$ we obtain the general solution of the systems as $x = -(9c/5) + (16/3)$,

$y = -(6c/5) + (16/3)$;

$z = c$.

Problem 3. For What values of η the equations

$$x + y + z = 1$$
$$x + 2y + 4z = \eta$$
$$x + 4y + 10z = \eta^2 \text{ are consistent?}$$

Solution. The matrix form of the system is given by

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 10 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \eta \\ \eta^2 \end{bmatrix}$$

$\therefore$ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \eta \\ 1 & 4 & 10 & \eta^2 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta - 1 \\ 0 & 3 & 9 & \eta^2 - 1 \end{bmatrix} \begin{matrix} R_2 \to R_2 - R_1 \\ R_3 \to R_3 - R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta - 1 \\ 0 & 0 & 0 & \eta^2 - 3\eta + 2 \end{bmatrix} R_3 \to R_3 - 3R_2$$

$\therefore$ The given system is consistent iff $\eta^2 - 3\eta + 2 = 0$

$\therefore$ $\eta = 2$ or 1.

Problem 4. Show that the system of equations

$$x + 2y + z = 11$$
$$4x + 6y + 5z = 8$$
$$2x + 2y + 3z = 19 \text{ is inconsistent.}$$

Solution. The matrix form of the system is given by

$$\begin{bmatrix} 1 & 2 & 1 \\ 4 & 6 & 5 \\ 2 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 11 \\ 8 \\ 19 \end{bmatrix}$$

$\therefore$ The augmented matrix is given by

$$(A, B) = \begin{bmatrix} 1 & 2 & 1 & 11 \\ 4 & 6 & 5 & 8 \\ 2 & 2 & 3 & 19 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & -2 & 1 & -3 \end{bmatrix} \begin{matrix} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 2R_1 \end{matrix}$$

$$\sim \begin{bmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & 0 & 0 & 33 \end{bmatrix} R_3 \to R_3 - R_2$$

$\therefore$ Rank of $A = 2$ and rank of $(A, B) = 3$.

$\therefore$ The given system is inconsistent.

Exercises

1. Solve or prove the inconsistency of the following systems of equations.

(a)
$$2x - y + 3z = 8$$
$$x - 2y - z = -4$$
$$3x + y - 4z = 0$$

(b)
$$x + 2y - 5z = 0$$
$$3x + 4y + 6z = 0$$
$$x + y + z = 0$$

(c)
$$\begin{aligned} x+2y-5z&=-9\\ 3x-y+2z&=5\\ 2x+3y-z&=3\\ 4x-5y+z&=-3 \end{aligned}$$

(d)
$$\begin{aligned} x+y+z&=1\\ x+2y+3z&=1\\ x+3y+5z&=7\\ x+4y+7z&=10 \end{aligned}$$

(e)
$$\begin{aligned} x-2y-z-t&=-1\\ 3x-2z+3t&=-4\\ 5x-4y+t&=-3 \end{aligned}$$

(f)
$$\begin{aligned} x+y+z&=7\\ x+2y+3z&=8\\ y+2z&=6 \end{aligned}$$

2. For what values of λ and μ the system of equations

$$\begin{aligned} x+y+z&=6\\ x+2y+3z&=10\\ x+2y+\lambda z&=\mu \end{aligned}$$

is (a) inconsistent (b) consistent (c) consistent and the solution is unique.

3. Show that the set of all solutions of the system of homogeneous equations $AX = 0$ forms a vector space.

4. Show that a system of n equations in n unknowns given by $AX = Y$ has a unique solution if the $n \times n$ matrix A is non-singular.

Answers.

1. (a) consistent; $x = y = z = 2$

(b) consistent; $x = y = z = 0$;

(c) consistent; $x = \frac{1}{2}, y = \frac{3}{2}, z = \frac{5}{2}$;

(d) consistent; $x = c-2, y = 3-2c; z = c$

(e) inconsistent;

(f) inconsistent;

(g) inconsistent.

2. If $\lambda = 3$ and $\mu \neq 10$, inconsistent.
If $\lambda = 3$ and $\mu = 10$, consistent.
If $\lambda \neq 3$, consistent and the solution is unique.

7.7. Characteristic Equation And Cayley Hamilton Theorem

Definition. An expression of the form $A_0 + A_1x + A_2x^2 + \cdots + A_nx^n$ where $A_0, A_1, \ldots, A_n$ are square matrices of the same order and $A_n \neq \mathbf{0}$ is called a **matrix polynomial** of degree n.

For example, $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}x + \begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix}x^2$ is a matrix polynomial of degree 2 and it is simply the matrix $\begin{pmatrix} 1+x+2x^2 & 2+x \\ 2x+3x^2 & 3+x+x^2 \end{pmatrix}$.

Definition. Let A be any square matrix of order n and let I be the identity matrix of order n. Then the matrix polynomial given by $A - xI$ is called the **characteristic matrix** of A.

The determinant $|A - xI|$ which is an ordinary polynomial in x of degree n is called the **characteristic polynomial** of A.

The equation $|A - xI| = 0$ is called the **characteristic equation** of A.

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

Then the characteristic matrix of A is $A - xI$ given by

$$A - xI = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix}$$

$\therefore$ The characterisitc polynomial of A is

$$|A - xI| = \begin{vmatrix} 1-x & 2 \\ 3 & 4-x \end{vmatrix}$$

$$= (1-x)(4-x) - 6$$
$$= x^2 - 5x - 2$$

$\therefore$ The characteristic equation of A is $|A - xI| = 0$

$\therefore$ $x^2 - 5x - 2 = 0$ is the characteristic equation of A.

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

The characteristic matrix of A is $A - xI$ given by

$$A - xI = \begin{pmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{pmatrix}$$

The characteristic polynomial of A is

$$|A - xI| = \begin{vmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{vmatrix}$$
$$= (1-x)\,[(1-x)(-x) - 4] - 2(1-x)]$$
$$= -x(1-x)^2 - 4(1-x) - 2 + 2x$$
$$= -x^3 + 2x^2 - x - 4 + 4x - 2 + 2x$$
$$= -x^3 + 2x^2 + 5x - 6$$

$\therefore$ The characteristic equation of A is

$$-x^3 + 2x^2 + 5x - 6 = 0$$
$$\text{(i.e)} \quad x^3 - 2x^2 - 5x + 6 = 0$$

Theorem 7.31. (Cayley Hamilton Theorem).

Any square matrix A satisfies its characteristic equation.

(i.e.) if $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ is the characteristic polynomial of degree n of A then

$$a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = \mathbf{0}.$$

Proof. Let A be a square matrix of order n.

Let $|A - xI| = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \qquad (1)$

be the characteristic polynomial of A.

Now, $adj(A - xI)$ is a matrix polynomial of degree $n-1$ since each entry of the matrix $adj(A - xI)$ is a cofactor of $A - xI$ and hence is a polynomial of degree $\leq n-1$.

$\therefore$ Let $adj\ (A - xI) = B_0 + B_1x + B_2x^2 + \cdots + B_{n-1}x^{n-1}. \qquad (2)$

Now, $(A - xI)adj\ (A - xI) = |A - xI|I.$

$(\because \quad (adj A)A = A(adj A) = |A|I)$

$\therefore \quad (A - xI)(B_0 + B_1x + \cdots + B_{n-1}x^{n-1})$
$= (a_0 + a_1x + \cdots + a_nx^n)I$ using (1) and (2).

$\therefore$ Equating the coefficients of the corresponding powers of x we get

$$AB_0 = a_0I$$
$$AB_1 - B_0 = a_1I$$
$$AB_2 - B_1 = a_2I$$
$$\cdots \qquad \cdots\cdots$$
$$\cdots \qquad \cdots\cdots$$
$$AB_{n-1} - B_{n-2} = a_{n-1}I$$
$$-B_{n-1} = a_nI$$

Pre-multiplying the above equations by $I, A, A^2, \ldots, A^n$ respectively and adding we get

$$a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = 0.$$

Note. The inverse of a non-singular matrix can be calculated by using the Cayley Hamilton theorem as follows.

Let $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ be the characteristic polynomial of A.

Then by theorem 1.1 we have

$$a_0I + a_1A + a_2A^2 + \cdots + a_nA^n = 0. \qquad (3)$$

Since $|A - xI| = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ we get $a_0 = |A|$ (by putting $x = 0$).

$\therefore \quad a_0 \neq 0$ ($\because A$ is a non singular matrix.)

$$\therefore \quad I = -\frac{1}{a_0}\left[a_1A + a_2A^2 + \cdots + a_nA^n\right] \text{(by 3)}$$

$$A^{-1} = -\frac{1}{a_0}\left[a_1 I + a_2 A + \cdots + a_n A^{n-1}\right].$$

Solved Problems

Problem 1. Find the characteristic equation of the matrix

$$A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$

Solution. The characteristic equation of A is given by $|A-\lambda I| = 0$.

(i.e.) $\begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} = 0.$

$$(8-\lambda)[(7-\lambda)(3-\lambda)-16]+6[-6(3-\lambda)+8] + 2[24-2(7-\lambda)] = 0$$

(i.e.) $(8-\lambda)(\lambda^2-10\lambda+5)+6(6\lambda-10) + 2(2\lambda+10) = 0$

(i.e.) $(8\lambda^2-80\lambda+40-\lambda^3+10\lambda^2-5\lambda) + (36\lambda-60)+(4\lambda+20) = 0$

(i.e.) $\lambda^3-18\lambda^2+45\lambda = 0$, which represents the characteristic equation of A.

Problem 2. Show that the non-singular matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ satisfies the equation $A^2-2A-5I = \mathbf{0}$. Hence evaluate A^{-1}.

Solution. The characteristic polynomial of A is

$$|A-xI| = \begin{vmatrix} 1-x & 2 \\ 3 & 1-x \end{vmatrix} = x^2-2x-5.$$

By Cayley-Hamilton theorem $A^2-2A-5I = \mathbf{0}$.

$$\therefore \quad I = \frac{1}{5}(A^2-2A).$$

$$\therefore \quad A^{-1} = \frac{1}{5}(A-2I)$$

$$= \frac{1}{5}\left[\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$$

$$= \frac{1}{5}\begin{pmatrix} -1 & 2 \\ 3 & -1 \end{pmatrix}$$

Problem 3. Show that the matrix

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

$A(A-I)(A+2I) = \mathbf{0}$.

Solution. The characteristic polynomial of A is

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

$= -\lambda^3-\lambda^2+2\lambda$ (verify).

$\therefore$ By Cayley-Hamilton theorem $-A^3-A^2+2A = \mathbf{0}$.

(i.e.) $A^3+A^2-2A = \mathbf{0}$. Hence $A(A^2+A-2I) = \mathbf{0}$.

$\therefore \quad A(A+2I)(A-I) = \mathbf{0}$.

Problem 4. Using Cayley-Hamilton theorem find the inverse of the matrix

$$\begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$$

Solution. Let $A = \begin{bmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{bmatrix}$

The characteristic polynomial of $A = |A-xI|$

$$= \begin{vmatrix} 7-x & 2 & -2 \\ -6 & -1-x & 2 \\ 6 & 2 & -1-x \end{vmatrix}$$

$$= (7-x)[(1+x)^2-4]-2[6(1+x)-12] - 2[-12+6(1+x)]$$

$$= \begin{pmatrix} 17 & 0 & -102 \\ 102 & 68 & 204 \\ 0 & 0 & 68 \end{pmatrix} - \begin{pmatrix} 36 & 0 & -72 \\ 72 & 72 & 144 \\ 0 & 0 & 72 \end{pmatrix} + \begin{pmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{pmatrix}$$

$$\therefore \ A^4 = \begin{pmatrix} 1 & 0 & -30 \\ 30 & 16 & 60 \\ 0 & 0 & 16 \end{pmatrix}$$

Exercises

1. Obtain the characteristic polynomial for the following matrices.

   (i) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

2. Find the characteristic equation of the following matrices.

   (i) $\begin{pmatrix} -b & -c \\ 1 & 0 \end{pmatrix}$ (ii) $\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$

   (iii) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -8 \\ 2 & -4 & 3 \end{bmatrix}$ (iv) $\begin{bmatrix} -b & -c & -d \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

3. Verify Cayley-Hamilton theorem for the matrix $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$ and hence find A^{-1}.

4. If $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{bmatrix}$ prove that $A^3 - 2A^2 - 5A + 6I = \mathbf{0}$.

5. Verify Cayley-Hamilton theorem for A and hence find A^{-1}.

   (a) $A = \begin{bmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{bmatrix}$

   (b) $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 3 \end{bmatrix}$

   (c) $A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$

   (d) $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$

6. If $A = \begin{pmatrix} 2 & 4 \\ 1 & 1 \end{pmatrix}$ find A^3 and A^{-3}.

7. Verify that the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & -1 \end{pmatrix}$ satisfies its own characteristic equation and hence find A^{-1} and A^4.

7.8. Eigen Values And Eigen Vectors

Definition. Let A be an $n \times n$ matrix. A number λ is called an **eigen value** of A if there exists a non-zero vector $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ such that $AX = \lambda X$ and X is called an **eigen vector** corresponding to the eigen value λ.

Remark 1. If X is an eigen vector corresponding to the eigen value λ of A, then αX where α is any non-zero number, is also an eigen vector corresponding to λ.

Remark 2. Let X be an eigen vector corresponding to the eigen value λ of A. Then $AX = \lambda X$ so that $(A - \lambda I)X = \mathbf{0}$. Thus X is a non-trivial solution of the system of homogeneous linear equations $(A - \lambda I)X = \mathbf{0}$. Hence $|A - \lambda I| = 0$, which is the characteristic polynomial of A.

$$\text{Let } |A - \lambda I| = a_0\lambda^n + a_1\lambda^{n-1} + \cdots + a_n.$$

The roots of this polynomial give the eigen values of A. Hence eigen values are also called **characteristic roots**.