UNIT I

Linear Equations

A **linear equation** in the variables $x_1, x_2, \cdots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \text{................(1)}$$

where b and the coefficients $a_1, a_2, \cdots, a_n$ are real or complex numbers, usually known in advance. The subscript n may be any positive integer. The variables are only multiplied by numbers-we don't see x_1 times x_n.

The equations

$$4x_1 - 5x_2 + 2 = x_1 \text{ and } x_2 = 2\left(\sqrt{6} - x_1\right) + x_3$$

are both linear because they can be rearranged algebraically as in equation (1).

The equations

$$4x_1 - 5x_2 = x_1x_2 \text{ and } x_2 = 2\sqrt{x_1} - 6$$

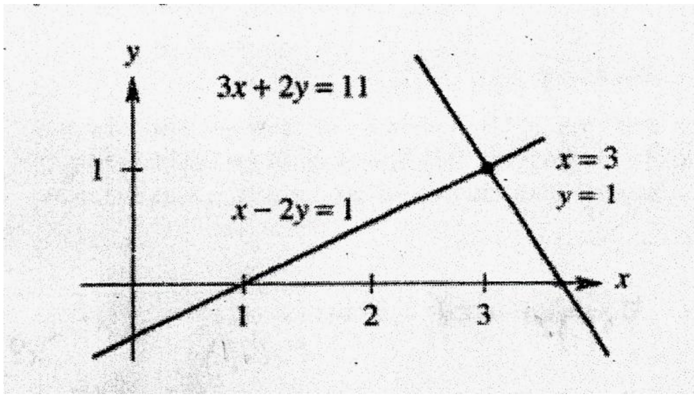
are not linear because of the presence of x_1x_2 in the first equation and $\sqrt{x_1}$ in the second.

A system of linear equations (or a linear system) is a collection of one or more linear equations involving the same variables-say, x and y.

An example is

$$\left.\begin{aligned} x - 2y &= 1 \\ 3x + 2y &= 11 \end{aligned}\right\} \text{.....................(2)}$$

we begin with a *row*. These equations produces straight lines in the xy plane. The first equation $x - 2y = 1$, can be plotted by joining the points $(0, -0.5)$ and (1,0)in the xy plane. Similarly, $3x + 2y = 11$, can be plotted by joining the points (0,5.5) and (3.6,0). These two lines form the *row picture* as shown in the following figure.

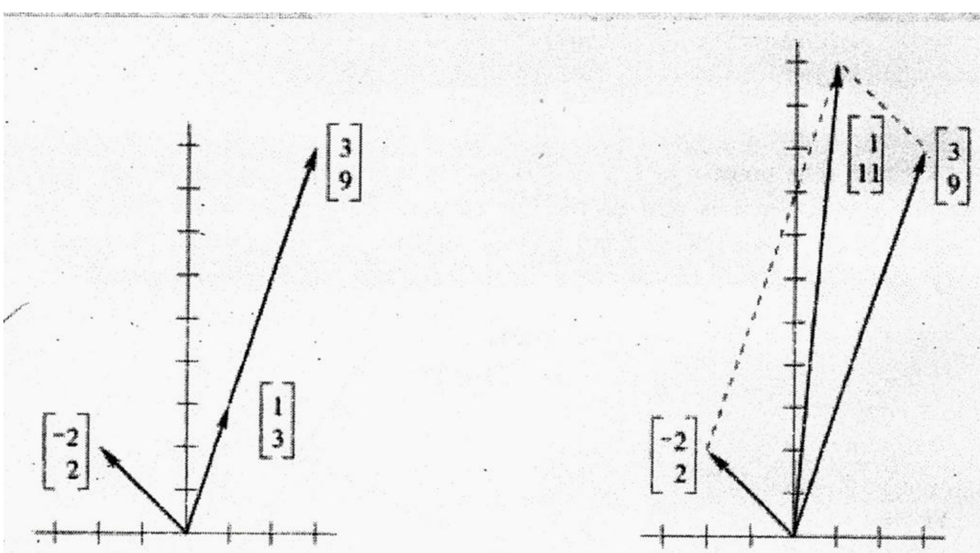


The point (3,1)lies on both the lines and is the point of intersection. The point (3,1) solves both equation at once.

Another way of recognizing the linear system is by *column picture*. In Column picture instead of numbers we need to see *vectors*. If we separate the original system into its column instead of its rows, we get

$$x\begin{bmatrix}1\\3\end{bmatrix}+y\begin{bmatrix}-2\\2\end{bmatrix}=\begin{bmatrix}1\\11\end{bmatrix}=\boldsymbol{b}$$

This has two column vectors on the left side. The problem is to find the combination of those vectors that equals the vector on the right. We are multiplying the first column by x and the second column by y, and adding. With the right choices $x = 3$ and $y = 1$, this produces $3(column1) + 1(column\ 2) = \boldsymbol{b}$. The column picture combines the column vectors on the left side to produce the vector $\boldsymbol{b}$ on the right side.



The first part of the above figure shows two separate columns, and that first column multiplied by 3. This multiplication by a *scalar* (a number) is one of the two basic operations in linear algebra:

$$\text{Scalar multiplication } 3\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}3\\9\end{bmatrix}$$

The other basic operation is vector addition.

$$\text{Vector addition } \begin{bmatrix}3\\9\end{bmatrix} + \begin{bmatrix}-2\\2\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix}$$

The second part of above figure shows a parallelogram. The sum $\begin{bmatrix}1\\11\end{bmatrix}$ is along the diagonal.

We have multiplied the original columns by $x = 3$ and $y = 1$. That combination produces the vector $\boldsymbol{b} = \begin{bmatrix}1\\11\end{bmatrix}$ on the right side of the linear equations.

It is lot easier to see a combination of four vectors in four-dimensional space, than to visualize how four hyperplanes might possibly meet at a point.

The *coefficient matrix* on the left side of the equations is by 2 by 2 matrix A:

$$\textbf{Coefficient matrix } A = \begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}$$

Its rows give the row picture and its columns give the column picture. We write those equations as a matrix problem $A\boldsymbol{x} = \boldsymbol{b}$:

$$\textbf{Matrix equation } \begin{bmatrix}1 & -2\\3 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}1\\11\end{bmatrix}$$

Algebra of Matrices

A ***matrix*** is a rectangular *array* (arrangement) of scalars, usually presented in the following form:

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

The rows of such a matrix A are $m$horizontal list of scalars:

$$(a_{11} \quad a_{12} \quad \cdots \quad a_{1n}), (a_{21} \quad a_{22} \quad \cdots \quad a_{2n}), \cdots, (a_{m1} \quad a_{m2} \quad \cdots \quad a_{mn}),$$

and the columns of A are the n vertical lists of the scalars:

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

The elements a_{ij} are called ij-entry or ij-element, which appears in row i and column j.

We frequently denote such a matrix by simply writing $A = [a_{ij}]$.

A matrix with m rows and n columns is called an ***m* by *n* matrix**, written $m \times n$. The pair of numbers m and n is called the ***size of the matrix***. Two matrices A and B are ***equal***, written $A = B$, if they have the same size and the corresponding elements are equal.

A matrix with only one row is called a ***row matrix*** or ***row vector,*** and a matrix with only one column is called a ***column matrix*** or ***column vector***. A matrix whose elements are all zero is called a ***zero matrix*** and is usually denoted by **0**.

Matrices whose entries are all real numbers are called ***real matrices***.

Ex. Find $x, y, z,$ and t such that

$$\begin{bmatrix} x + y & 2z + t \\ x - y & z - t \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 1 & 5 \end{bmatrix}$$

Ans.: $x = 2, y = 1, z = 4, t = -1$

Matrix Addition and Scalar Multiplication

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two matrices with the same size, say $m \times n$ matrices. The ***sum of A and B***, is the matrix obtained by adding corresponding elements from A and B. That is,

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

The ***product of the matrix A by a scalar k***, written $k.A$ or simply kA, is the matrix obtained by multiplying each element of A by k. That is,

$$kA = \begin{bmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{bmatrix}$$

$A + B$ and kA are also $m \times n$ matrices.

We also define $-A = (-1)A$ and $A - B = A + (-B)$

The matrix $-A$ is called the ***negative of matrix A***, and the matrix $A - B$ is called the ***difference of A and B***. The sum of matrices of different sizes are not defined.

Theorem 1.1: Consider any matrices A, B, C (with the same size) and any scalars k and k'. Then :

i. $(A + B) + C = A + (B + C)$
ii. $A + 0 = 0 + A = A$
iii. $A + (-A) = (-A) + A = 0$
iv. $A + B = B + A$
v. $k(A + B) = kA + kB$
vi. $(k + k')A = kA + k'A$
vii. $(kk')A = k(k'A)$
viii. $1.A = A$

The proof of above theorem reduces to showing that the ij-elements on both the sides of each matrix equation are equal.

Matrix Multiplication

Suppose that $A = [a_{ik}]$ and $B = [b_{kj}]$ are matrices such that the number of columns of A is equal to the number of rows of B; say A is $m \times p$ matrix and B is $p \times n$ matrix. Then the product AB is the $m \times n$ matrix whose ij-element is obtained by multiplying the corresponding elements of i^{th} row of A by the j^{th} column of B and adding. That is,

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i1} & a_{i2} & \cdots & a_{ip} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{p1} & \cdots & b_{pj} & \cdots & b_{pn} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ \cdots & c_{ij} & \cdots \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix}$$

where $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}$

Ex. Find AB where $A = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 & -4 \\ 5 & -2 & 6 \end{bmatrix}$

$$AB = \begin{bmatrix} 1 \times 2 + 3 \times 5 & 1 \times 0 + 3 \times (-2) & 1 \times (-4) + 3 \times 6 \\ 2 \times 2 + (-1) \times 5 & 2 \times 0 + (-1) \times (-2) & 2 \times (-4) + (-1) \times 6 \end{bmatrix}$$

$$= \begin{bmatrix} 17 & -6 & 14 \\ -1 & 2 & -14 \end{bmatrix}$$

Theorem 1.2: Let A, B and C be matrices. Then, whenever the products and sum are defined:

i. $(AB)C = A(BC)$; associative law
ii. $A(B + C) = AB + AC$; left distributive law
iii. $(B + C)A = BA + CA$; right distributive law
iv. $k(AB) = (kA)B = A(kB)$, where k is a scalar

Also, $0A = 0$ and $B0 = 0$, where 0 is the zero matrix.

Transpose of a matrix

The transpose of a matrix A, written A^T, is the matrix obtained by writing the columns of A, in order, as rows. For example,

$$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}^T = \begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix} \text{ and } [1 \quad -3 \quad -5]^T = \begin{bmatrix}1\\-3\\-5\end{bmatrix}$$

In other words, if $A = [a_{ij}]$ is an $m \times n$ matrix, then $A^T = [b_{ij}]$ is the $n \times m$ matrix where $b_{ij} = a_{ji}$.

Theorem 1.3: Let A and B be matrices and let k be a scalar. Then, whenever sum and products are defined:

i. $(A + B)^T = A^T + B^T$
ii. $(kA)^T = kA^T$
iii. $(A^T)^T = A$
iv. $(AB)^T = B^T A^T$

Square Matrices

A ***square matrix*** is a matrix with the same number of rows as columns. An $n \times n$ square matrix is said to be of order n and is sometimes called an n-square matrix.

Diagonal and Trace

Let $A = [a_{ij}]$ be an n-square matrix. The ***diagonal or main diagonal of A*** consist of the elements with the same subscript. That is,

$$a_{11}, a_{22,} a_{33}, \cdots, a_{nn}$$

The ***trace of A***, written $tr(A)$, is the sum of the diagonal elements. Namely,

$$tr(A) = a_{11} + a_{22+} a_{33} + \cdots + a_{nn}$$

Theorem 1.4: Suppose $A = [a_{ij}]$ and $B = [b_{ij}]$ are n-square matrices and k is a scalar. Then:

i. $tr(A + B) = tr(A) + tr(B)$
ii. $tr(A^T) = tr(A)$
iii. $tr(kA) = k\, tr(A)$
iv. $tr(AB) = tr(BA)$

Identity Matrix and Scalar Matrix

The n-square ***identity or unit matrix***, denoted by I_n, or simply I is the n-square matrix with $1's$ on the diagonal and $0's$ elsewhere

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

Also,

$$AI = IA = A$$

More generally, if B is an $m \times n$ matrix $BI = IB = B$

For any scalar $k,$ the matrix kI that contains $k's$ on the diagonal and $0's$ elsewhere is called the ***scalar matrix*** corresponding to the scalar k.

example: $\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}$

Invertible (Nonsingular) Matrices

A square matrix A is said to be invertible or nonsingular if there exists a matrix B such that

$$AB = BA = I$$

where I is the identity matrix. Such a matrix B is **unique.** That is, if $AB_1 = B_1A = I$ and $AB_2 = B_2A = I$, then

$$B_1 = B_1I = B_1(AB_2) = (B_1A)B_2 = IB_2 = B_2$$

We call this matrix B the inverse of $A,$ and denote it by A^{-1}. Observe that the above relation is symmetric; that is, if B is the inverse of $A,$ then A is the inverse of B.

Suppose A and B are invertible. Then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$. More generally, if $A_1, A_2, \cdots, A_k$ are invertible, then their product is invertible and

$$(A_1A_2 \cdots A_k)^{-1} = A_k^{-1} \cdots A_2^{-1}A_1^{-1}$$

Diagonal and Triangular Matrices

A square matrix $D = [d_{ij}]$ is ***diagonal*** if its nondiagonal elements are all zero. Such a matrix is sometimes denoted by

$$D = diag(d_{11}, d_{22}, \cdots, d_{nn})$$

where some or all d_{ii} may be zero. For example: $\begin{bmatrix} 3 & 0 & 0 \\ 0 & -7 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix}$

A square matrix $A = [a_{ij}]$ is ***upper triangular or simply triangular*** if all entries below the (main) diagonal are equal to zero, that is if $a_{ij} = 0$ for $i > j$.

$$\begin{bmatrix} 3 & 5 & 2 \\ 0 & -7 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 5 \\ 0 & -3 \end{bmatrix}$$

Symmetric Matrix

A matrix A is symmetric if $A^T = A$. Equivalently, $A = [a_{ij}]$ is symmetric if symmetric elements (mirror elements with respect to the diagonal) are equal, that is, if each $a_{ij} = a_{ji}$.

A matrix A is ***skew-symmetric*** if $A^T = -A$. Equivalently, if each $a_{ij} = -a_{ji}$. Clearly, the diagonal elements of such a matrix must be zero[1].

Orthogonal Matrices

A real matrix A is ***orthogonal*** if $A^T = A^{-1}$, that is, if $AA^T = A^T A = I$. Thus A must necessarily be square and invertible.

Ex. $A = \begin{bmatrix} \frac{1}{9} & \frac{8}{9} & -\frac{4}{9} \\ \frac{4}{9} & -\frac{4}{9} & -\frac{7}{9} \\ \frac{8}{9} & \frac{1}{9} & \frac{4}{9} \end{bmatrix}$

---

[1] A matrix must be square if $A^T = A$ or $A^T = -A$

Normal Matrix

A real matrix A is ***normal*** if it commutes with its transpose A^T, that is, $AA^T = A^T A$,

All symmetric, skew-symmetric and orthogonal matrices are normal matrices.

Ex. $\begin{bmatrix} 6 & -3 \\ 3 & 6 \end{bmatrix}$

Block Matrix

Using a system of horizontal and vertical (dashed) lines, we can partition a matrix A into submatrices called block (or cells) of A. Clearly a given matrix may be divided into blocks in different ways. For example:

$$\left[\begin{array}{cc:cc:c} 1 & -2 & 0 & 1 & 3 \\ 2 & 3 & 5 & 7 & -2 \\ \hdashline 3 & 1 & 4 & 5 & 9 \\ 4 & 6 & -3 & 1 & 8 \end{array}\right], \quad \left[\begin{array}{cc:ccc} 1 & -2 & 0 & 1 & 3 \\ \hdashline 2 & 3 & 5 & 7 & -2 \\ 3 & 1 & 4 & 5 & 9 \\ \hdashline 4 & 6 & -3 & 1 & 8 \end{array}\right].$$

The convenience of the partition of matrices, say A and B, into blocks is that the result of operations on A and B can be obtained by carrying out the computation with the blocks, just as if they were the actual elements of the matrices.

Suppose that $A = [A_{ij}]$ and $B = [B_{ij}]$ are block matrices with the same numbers of row and column blocks, and suppose that corresponding blocks have the same size. Then adding the corresponding blocks of A and B also adds the corresponding elements of A and B, and multiplying each block of A by a scalar k multiplies each element of A by k. Thus

$$A + B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \cdots & A_{1n} + B_{1n} \\ A_{21} + B_{21} & A_{22} + B_{22} & \cdots & A_{2n} + B_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{m1} + B_{m1} & A_{m2} + B_{m2} & \cdots & A_{mn} + B_{mn} \end{bmatrix}$$

and

$$kA = \begin{bmatrix} kA_{11} & kA_{12} & \cdots & kA_{1n} \\ kA_{21} & kA_{22} & \cdots & kA_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ kA_{m1} & kA_{m2} & \cdots & kA_{mn} \end{bmatrix}$$

The case of matrix multiplication is less obvious, but still true. That is, suppose that $U = [U_{ik}]$ and $V = [V_{kj}]$ are block matrices such that the number of columns of each block $U_{ik}$is equal to the number of rows of each block V_{kj}. Then,

$$UV = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1n} \\ W_{21} & W_{22} & \cdots & W_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ W_{m1} & W_{m2} & \cdots & W_{mn} \end{bmatrix}, \text{ where } W_{ij} = U_{i1}V_{1j} + U_{i2}V_{2j} + \cdots + U_{ip}V_{pj}$$

Kronecker product

The Kronecker product is an operation that transforms two matrices into a larger matrix that contains all the possible products of the entries of the two matrices. It possesses several properties that are often used to solve difficult problems in linear algebra and its applications.

Let A be a $K \times L$ matrix and B an $M \times N$ matrix. Then, the Kronecker product between A and B is the $(KM \times LN)$ block matrix

$$A \otimes B = \begin{bmatrix} A_{11}B & \dots & A_{1L}B \\ \vdots & \ddots & \vdots \\ A_{K1}B & \dots & A_{KL}B \end{bmatrix}$$

where A_{KL} denotes the KL-th entry of A.

Elementary Transformation (or Elementary Operations) of a Matrix

The following three transformations applied on the rows (or columns) of a matrix are called elementary row (or column) transformations.

i. **Interchange of any two rows (or columns):** If i^{th} row (or column) of a matrix is interchanged with the j^{th} row (column). This transformation is denoted by $R_i \leftrightarrow R_j$ $(C_i \leftrightarrow C_j)$.

ii. **Multiplication of (each element of) a row or column by a non-zero number $\boldsymbol{k}$**: If the elements of i^{th} row (or column) are multiplied by a non-zero scalar k. This transformation is denoted by $R_i \to kR_i$ $(or\ C_i \to kC_i)$.

iii. **Addition of $\boldsymbol{k}$ times the elements of a row (or column) to the corresponding elements of another row (or column), $\boldsymbol{k \neq 0}$**: If k times the elements of j^{th} row (or column) are added to the corresponding elements of i^{th} row (or column). It will be denoted by $R_i \to R_i + kR_j$ $(or\ C_i \to C_i + kC_j)$.

Elementary Matrix

A matrix obtained from an identity matrix by a single elementary transformation is called an elementary matrix.

For example: Let $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

then operating $R_1 \leftrightarrow R_3$, we get elementary matrix $E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

Ex. Transform $\begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$, into a unit matrix by using elementary transformation.

Let $A = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$

applying $R_2 \to R_2 - 2R_1$

$$\sim \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

applying $R_1 \to R_1 - 3R_2$

$$\sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Ex. Reduce the matrix A to upper triangular form:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 1 & 2 \end{bmatrix}$$

Inverse of a Matrix by elementary Transformations (Guass Jordon Method)

The elementary row transformations which reduce a square matrix A to the unit matrix , gives the inverse matrix A^{-1}. If we are to find out the inverse of a nonsingular square matrix A, we first write A equivalent to I, a unit matrix of same order $A \sim I$.

Then we apply elementary row transformations on them. The objective is to reduce A to I. As soon as this is achieved, the other matrix gives A^{-1}.

$I \sim A^{-1}$.

Ex. Using elementary row transformations, find the inverse of the matrix.

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

Solution: Let $A \sim I$

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

Applying $R_1 \to R_1 - R_2$

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

Applying $R_2 \to R_2 - 2R_1$ and $R_3 \to R_3 - 3R_1$

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

Applying $R_2 \to \frac{1}{2} R_2$

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

Applying $R_1 \to R_1 + R_2$ and $R_3 \to R_3 + 2R_2$

$$\begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 4 \end{bmatrix} \sim \begin{bmatrix} 0 & 1/2 & 0 \\ -1 & 3/2 & 0 \\ -5 & 6 & 1 \end{bmatrix}$$

Applying $R_3 \to \frac{1}{4} R_3$

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

Applying $R_1 \to R_1 + \frac{1}{2} R_3$ and $R_2 \to R_2 - \frac{1}{2} R_3$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} -5/8 & 5/4 & 1/8 \\ -3/8 & 3/4 & -1/8 \\ -5/4 & 3/2 & 1/4 \end{bmatrix}$$

Therefore, $I \sim A^{-1}$

Thus, $A^{-1} = \begin{bmatrix} -5/8 & 5/4 & 1/8 \\ -3/8 & 3/4 & -1/8 \\ -5/4 & 3/2 & 1/4 \end{bmatrix}$

Rank of a Matrix

Let A be any $m \times n$ matrix. It has square sub-matrices of different orders. The determinants of these square sub-matrices are called minors of A.

A matrix A is said to be of rank r if

i. if it has at least one nonzero minor of order r.
ii. all the minors of order $r + 1$ or higher than r are zero.

It is denoted by $\rho(A) or\ r(A)$[2].

- [2] If A is a null matrix, then $\rho(A) = 0$.

Ex. Find the rank of the following matrices:

1. $\begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix}$
2. $\begin{bmatrix} 5 & 2 & 5 \\ 7 & 3 & 5 \end{bmatrix}$
3. $\begin{bmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \end{bmatrix}$
4. $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$
5. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 4 & 5 \end{bmatrix}$
6. $\begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & -1 \\ -1 & 2 & 7 \end{bmatrix}$
7. $\begin{bmatrix} 1 & -2 \\ -2 & 4 \\ -1 & 2 \end{bmatrix}$

Solution:

1. Let $A = \begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix}$

$$\rho(A) \leq min.(2,2) = 2$$

$$|A| = 10 - 10 = 0$$

Therefore, $\rho(A) < 2$

As there are nonzero elements in A, thus $\rho(A) = 1$.

2. Let $A = \begin{bmatrix} 5 & 2 & 5 \\ 7 & 3 & 5 \end{bmatrix}$

- If A is not a null matrix, then $\rho(A) \geq 1$.
- If A is an $m \times n$ matrix, then $\rho(A) \leq \min(m, n)$.
- If A is a non-singular $n \times n$ matrix, then $\rho(A) = n$
- Rank of I_n is n.
- $\rho(A) = \rho(A^T)$

$$\rho(A) \leq min.(2,3) = 2$$

Therefore, $\rho(A) \leq 2$

As one of the minor of order 2, $\begin{vmatrix} 5 & 2 \\ 7 & 3 \end{vmatrix} = 1 \neq 0$, Therefore $\rho(A) = 2$

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

$$\rho(A) \leq 2$$

Minors of order 2 are as follows

$$M_1 = \begin{vmatrix} 4 & 2 \\ 2 & 1 \end{vmatrix} = 0$$

$$M_2 = \begin{vmatrix} 4 & 6 \\ 2 & 3 \end{vmatrix} = 0$$

$$M_3 = \begin{vmatrix} 2 & 6 \\ 1 & 3 \end{vmatrix} = 0$$

All the minors of order 2 have values equal to zero.

Therefore, $\rho(A) < 2$

Since there are nonzero elements in A. Thus, $\rho(A) = 1$.

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

$$|A| = \begin{vmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{vmatrix} = 5 \neq 0$$

Therefore, $\rho(A) = 3$

5. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 4 & 5 \end{bmatrix}$

$$|A| = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 4 & 5 \end{vmatrix} = 0$$

Therefore, $\rho(A) < 3$

Now since one of the minors of order 2, $\begin{vmatrix} 5 & 6 \\ 4 & 5 \end{vmatrix} = 1 \neq 0$

Hence, $\rho(A) = 2$

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

$$|A| = \begin{vmatrix} 1 & -2 & 3 \\ -2 & 4 & -1 \\ -1 & 2 & 7 \end{vmatrix} = 0$$

Therefore, $\rho(A) < 3$

Now since one of the minors of order 2, $\begin{vmatrix} -2 & 3 \\ 4 & -1 \end{vmatrix} = -10 \neq 0$

Hence, $\rho(A) = 2$

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

$$\rho(A) \leq min.\,(3,2) = 2$$

Minors of order 2 are:

$$M_1 = \begin{vmatrix} 1 & -2 \\ -2 & 4 \end{vmatrix} = 0$$

$$M_2 = \begin{vmatrix} -2 & 4 \\ -1 & 2 \end{vmatrix} = 0$$

$$M_3 = \begin{vmatrix} 1 & -2 \\ -1 & 2 \end{vmatrix} = 0$$

As there are nonzero elements in A, thus $\rho(A) = 1$.

Echelon Form Matrix

A matrix is in echelon form, if it satisfies the following conditions:

i. The first nonzero element in each row, called the leading element, is always strictly to the right of the leading element of the row above it.
ii. Rows with all zero elements, if any are below the rows having a nonzero element.

Reduced row echelon form[3]:

If in addition to above conditions following conditions are also satisfied, then the matrix is said to be in reduced row echelon form:

i. The leading element in each nonzero row should be 1.
ii. Each leading 1 is the only nonzero element in its column.

Examples of Echelon Form Matrices:

$$\begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 5 & 0 & 1 & 3 & 2 \\ 0 & 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 5 \\ 0 & 4 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Examples of Reduced Row Echelon Form Matrices:

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Pivot Position

A position of a leading element in an echelon form of a matrix is called pivot position.

Pivot Column

A column that contains a pivot position is called a pivot column.

Rank of a Matrix by Echelon Form

Once a matrix is transformed into an echelon form by using the **elementary row** operations, rank of the matrix is equal to the number of nonzero rows in its echelon form matrix.

[3] Each matrix is row-equivalent to one and only one reduced row echelon form.

Ex. Find the rank of the following matrix by transforming it into an echelon form

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

Solution:

Applying the elementary transformation $R_1 \leftrightarrow R_2$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 2 & 3 & -1 & -1 \\ 3 & 1 & 3 & -2 \\ 6 & 3 & 0 & -7 \end{bmatrix}$$

$R_2 \to R_2 - 2R_1, R_3 \to R_3 - 3R_1, R_4 \to R_4 - 6R_1$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 0 & 5 & 3 & 7 \\ 0 & 4 & 9 & 10 \\ 0 & 9 & 12 & 17 \end{bmatrix}$$

$R_2 \to R_2 - R_3$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 0 & 1 & -6 & -3 \\ 0 & 4 & 9 & 10 \\ 0 & 9 & 12 & 17 \end{bmatrix}$$

$R_3 \to R_3 - 4R_2$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 0 & 1 & -6 & -3 \\ 0 & 0 & 33 & 22 \\ 0 & 9 & 12 & 17 \end{bmatrix}$$

$R_4 \to R_4 - 9R_2$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 0 & 1 & -6 & -3 \\ 0 & 0 & 33 & 22 \\ 0 & 0 & 66 & 44 \end{bmatrix}$$

$R_4 \to R_4 - 2R_3$

$$\sim \begin{bmatrix} 1 & -1 & -2 & -4 \\ 0 & 1 & -6 & -3 \\ 0 & 0 & 33 & 22 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The above matrix is in echelon form and its nonzero rows is 3. Thus the rank of $A = 3$.

Rank of a matrix by normal form method

In A is an $m \times n$ matrix and by a series of elementary (row or column or both) transformations, it can be put into one of the following forms (called normal or canonical forms):

$$\left[\begin{array}{c:c} I_r & 0 \\ \hdashline 0 & 0 \end{array}\right], \left[\begin{array}{c} I_r \\ \hdashline 0 \end{array}\right], [I_r \,\vdots\, 0], [I_r]$$

where I_r is the unit matrix of order r, then $\rho(A) = r$.

Ex. Reduce the matrix A to its normal form when $A = \begin{bmatrix} 1 & 2 & -1 & 4 \\ 2 & 4 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ -1 & -2 & 6 & -7 \end{bmatrix}$.

Hence find the rank of A.

Solution:

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

$R_2 \to R_2 - 2R_1, R_3 \to R_3 - R_1, R_4 \to R_4 + R_1$

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

$C_2 \to C_2 - 2C_1, C_3 \to C_3 + C_1, C_4 \to C_4 - 4C_1$

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

$C_3 \leftrightarrow C_2$

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

$R_2 \to R_2 - R_3$

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

$R_3 \to R_3 - 4R_2, R_4 \to R_4 - 5R_2$

$$\sim\begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & -4\\0 & 0 & 0 & 16\\0 & 0 & 0 & 17\end{bmatrix}$$

$C_4 \to C_4 + 4C_2$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 16 \\ 0 & 0 & 0 & 17 \end{bmatrix}$$

$C_4 \leftrightarrow C_3$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 17 & 0 \end{bmatrix}$$

$R_3 \to \dfrac{1}{16} R_3$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 17 & 0 \end{bmatrix}$$

$R_4 \to R_4 - 17R_3$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

which is a normal form. Hence $\rho(A) = 3$

Characteristic Equation, Eigen value and Eigen vector

For every square matrix A of order n, we can form a matrix $A - \lambda I$, where λ is a scalar and I is the unit matrix of order n. The determinant of this matrix equated to zero i.e., $|A - \lambda I| = 0$ is called the characteristic equation of A. On expanding the determinant we can write this equation as

$$(-1)^n[\lambda^n + k_1\lambda^{n-1} + k_2\lambda^{n-2} + \cdots + k_n] = 0$$

The roots of this equation are called ***characteristic roots or latent roots or eigen values.***

Now consider $(A - \lambda I)\boldsymbol{X} = \boldsymbol{0}$

$$\text{i.e.} \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

These equations will have a non-trivial solution, only if $\rho(A - \lambda I) < n$ $(= no.of\ unknown)$, which is possible when $(A - \lambda I)$ is singular, i.e., if $|A - \lambda I| = 0$.

This is the characteristic equation of the matrix A and has n roots, which are the eigen values of A. Corresponding to each root, the homogeneous system $(A - \lambda I)\boldsymbol{X} = \boldsymbol{0}$, has a nonzero solution

$$\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

which is called the ***eigen vector or latent vector***.

Ex. Find the eigen values and eigen vectors of the following matrix

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

Solution:

The characteristic equation is

$$|A - \lambda I| = 0$$

$$\Rightarrow \begin{vmatrix} 8 - \lambda & -6 & 2 \\ -6 & 7 - \lambda & -4 \\ 2 & -4 & 3 - \lambda \end{vmatrix} = 0$$

$$\Rightarrow \lambda^3 - 18\lambda^2 - 45\lambda = 0$$

$$\Rightarrow \lambda = 0,3,15$$

Hence the eigen values are 0,3,15

The eigen vector corresponding to $\lambda = 0$ is

$$[A - \lambda I]\boldsymbol{X_1} = 0$$

$$\Rightarrow \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$8x_1 - 6x_2 + 2x_3 = 0$

$-6x_1 + 7x_2 - 4x_3 = 0$

$2x_1 - 4x_2 + 3x_3 = 0$

Solving above equations we get $x_1 = k, x_2 = 2k, x_3 = 2k$

Thus $\boldsymbol{X_1} = k \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

Similarly we can find eigen vectors corresponding to eigen values $\lambda = 0,$ and 15.

Cayley-Hamilton (C-H) theorem

Every square matrix satisfies its own Characteristic equation.

That is, if

$|A - \lambda I| = (-1)^n[\lambda^n + k_1\lambda^{n-1} + k_2\lambda^{n-2} + \cdots + k_n] = 0$

then $(-1)^n[A^n + k_1A^{n-1} + k_2A^{n-2} + \cdots + k_{n-1}A + k_nI] = 0$

Ex. Verify Cayley-Hamilton theorem for the following matrix and hence find its inverse.

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

Solution. The Characteristic equation of A is

$$|A - \lambda I| = 0$$

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

Expanding the above determinant

$$(1-\lambda)\{(3-\lambda)(-4-\lambda) - 12\} - 1\{(-4-\lambda) - 6\} + 3\{-4 + 2(3-\lambda\} = 0$$

$$\lambda^3 - 20\lambda + 8 = 0$$

By C-H theorem we must have

$$A^3 - 20A + 8I = 0$$

LHS = $A^3 - 20A + 8I = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ = RHS

Hence verified

Now, pre multiplying both sides by A^{-1}, we have

$$A^{-1}A^3 - 20A^{-1}A + 8A^{-1}I = A^{-1}0$$

$$\Rightarrow A^2 - 20I + 8A^{-1} = 0$$

$$\Rightarrow A^{-1} = \frac{5}{2}I - \frac{1}{8}A^2$$

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

Ex. Find the characteristic equation of the matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$ and hence find its inverse. Also find the matrix represented by $A^8 - 5A^7 + 7A^6 - 3A^5 + A^4 - 5A^3 + 8A^2 - 2A + I$.

*****************End of Unit I*****************