BASICS OF MATRIX ALGEBRA

- A ***matrix*** is an ordered rectangular array of numbers or functions.
- The numbers or functions are called the *elements* or *entries* of the matrix.
    - Matrices are denoted by capital letters.
    - Ex: $A = \begin{bmatrix} 3 & -1 & 2 \\ 8 & 10 & 1 \end{bmatrix}$; $B = \begin{bmatrix} -9 & 5 \\ 5 & -2 \\ 6 & 7 \end{bmatrix}$; $C = \begin{bmatrix} 2 & 0 \\ -1 & 9 \end{bmatrix}$
- A matrix having m rows and n columns is called a matrix of *order* $m \times n$.
  We observe in above matrices, A is of order 2×3, matrix B is of order 3×2 and C is a 2×2 matrix.
- In general, an $m \times n$ matrix can be written as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}_{m\times n}$$

    - Here a_{ij} is the element or entry of the matrix in i^{th} row and j^{th} column.
    - We also write the matrix as $\left[a_{ij}\right]_{m\times n}$ or simply $\left[a_{ij}\right]$.
    - An $m \times n$ matrix is called a *rectangular matrix*.
- **Types of Matrices:**
    - A $1 \times n$ matrix having only one row is called a *<u>Row Matrix</u>*.
    - An $m \times 1$ matrix having only one column is called a *<u>Column Matrix</u>*.
    - An $m \times n$ matrix whose all elements are '0' is called *<u>Zero</u>* or *<u>Null matrix</u>* and is denoted by $O_{m\times n}$ or simply O.
    - An $n \times n$ matrix, where the number of rows and columns are same is called a *Square matrix*.
        - In this case, we say that the square matrix A is of order n and denote it as A_n .
        - The diagonal containing the elements a_{11}, a_{22}, …, a_{nn} is called the main diagonal or principal diagonal of the square matrix and the elements are called the diagonal elements.

- Sum of all the diagonal elements in a square matrix A of order n is called the *trace of A* and is denoted by tr(A). i.e. $tr(A)=\sum_{i=1}^{n} a_{ii}$

- A square matrix whose all non-diagonal elements are '0' is said to be a *Diagonal matrix*.
- A diagonal matrix whose all diagonal elements are equal is called a *Scalar matrix*.
- A diagonal matrix whose all diagonal element are equal to '1' is called a *Unit* or *Identity matrix*. An identity matrix of order n is denoted by I_n or I.
- The matrix obtained by interchanging rows and columns of a given matrix A of order $m \times n$ is called the *Transpose of a matrix*. It is denoted by A' or A^T and it has order $n \times m$.
- A square matrix A is said to be a *symmetric matrix* if $A' = A$.
- A square matrix A is said to be a *Skew-symmetric matrix* if $A' = -A$.
- A square matrix whose all elements below the principal diagonal are '0' is called an *Upper triangular matrix*.
- A square matrix whose all elements above the principal diagonal are '0' is called a *Lower triangular matrix*.
- A matrix obtained from a given matrix by omitting some rows and/or columns of the matrix is called its *Sub-matrix*.
- A matrix is said to be an *Orthogonal matrix* if $A'A = I$.

❖ **Operations on matrices:**

- *Equality of matrices*: Two matrices $A=\left[a_{ij}\right]$ and $B=\left[b_{ij}\right]$ of same order are said to be equal if $a_{ij}=b_{ij}$ for each i and j.
- Addition/Subtraction of two matrices: Addition/subtraction of two matrices $A=\left[a_{ij}\right]$ and $B=\left[b_{ij}\right]$ is possible if they are of same order and is defined as $A \pm B=\left[a_{ij}\right] \pm \left[b_{ij}\right]=\left[a_{ij} \pm b_{ij}\right]$.
- For any $m \times n$ matrix A and a scalar 'k' the scalar multiplication of A is denoted by kA and is defined as $kA = k\left[a_{ij}\right]=\left[ka_{ij}\right]$.
- Multiplication of matrices $A=\left[a_{ij}\right]_{m\times p}$ and $B=\left[b_{ij}\right]_{r\times n}$ is defined if $p = r$, i.e. number of columns of matrix A is equal to number of rows of matrix B then the resultant matrix is of order $m \times n$ and is defined as $C=\left[c_{ij}\right]_{m\times n}$ where $c_{ij}=\sum_{l=1}^{n} a_{il}b_{lj} = a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}$.

- To every square matrix $A=\left[a_{ij}\right]$ of order *n*, we can associate a number (real or complex) called <u>*determinant*</u> of the square matrix A. It is denoted by $|A|$ or det(*A*).
  - If $A=[a]$ is a matrix of order 1, then $|A|=a$.
  - If $A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a matrix of order 2, then $|A|=a_{11}a_{22}-a_{12}a_{21}$.
  - If $A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ is a matrix of order 3, then

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

- A square matrix is said to be <u>*singular matrix*</u> if its determinant is '0'.
- A square matrix is said to be <u>*non-singular matrix*</u> if its determinant is not zero.
- If *A* is a square matrix with $|A|\neq 0$, then there exists a matrix *B* such that *AB* = *BA* = *I*, where *I* is of the same order as *A*. Here *B* is called the <u>*inverse*</u> *of A* and is denoted by A^{-1}.
- The <u>*minor of an element*</u> in a square matrix is the determinant of a square matrix obtained by deleting a row and a column corresponding to the element of the matrix.
- **Elementary row/column operations:**
  - *Multiplication of a row/column by a scalar*
    - Multiplication of i^{th} row by *k* is denoted by kR_i.
    - Multiplication of i^{th} column by *k* is denoted by kC_i.
  - *Interchanging any two rows/columns*
    - The interchange of i^{th} and j^{th} row is denoted by R_{ij}.
    - The interchange of i^{th} and j^{th} column is denoted by C_{ij}.
  - *Multiplying a row/column by scalar and adding it to another row/column*
    - Multiplying the j^{th} row by a scalar *k* and adding it to i^{th} row is denoted by $R_i + kR_j$.
    - Multiplying the j^{th} column by a scalar *k* and adding it to the i^{th} column is denoted by $C_i + kC_j$.
- **Equivalent matrices:** If a matrix B is obtained from matrix A by performing a series of elementary transformation(s) on A, then B is said to be <u>*equivalent*</u> to A. We write it as A ~ B.
- In a rectangular matrix, **Leading entry** in a row is the left most non-zero entry in a non-zero row.

- **Non zero row/column** in a matrix is a row/column that contains at least one non zero entry.
- **Row-Echelon form**: A rectangular matrix is said to be in *echelon form* or *row echelon form(REF)* if it has the following properties:
    - All non-zero rows are above any rows of all zeros.
    - Each leading entry is 1.
    - Each leading entry of a row is in a column to the right of the leading entry of the row above it.
    - All entries in a column below a leading entry are zeros.

  We understand the definition by the given matrix: $\begin{bmatrix} 1 & \# & \# & \# \\ 0 & 0 & 1 & \# \\ 0 & 0 & 0 & 1 \end{bmatrix}$,

  where ‘ # ’ denotes any number which may be non-zero or zero.
- Pivot (element) is the first non-zero element in the first non-zero column and the row/column containing the pivot (element) is called the pivot row/column.
- **Working rules to reduce a matrix in row echelon form:**
    1. Only elementary row operations are applied to reduce in row echelon form.
    2. Identify the pivot element of the matrix.
    3. Move the pivot to the first row of the matrix (if necessary) by interchanging two rows. Thus, the first row becomes the pivot row.
    4. Make the pivot element as 1 (if required) by multiplying the entire row by the reciprocal of pivot.
    5. Make all elements **below** the pivot in the column as 0.
    6. Repeat steps (1) to (4) ignoring the pivot row.
    7. Continue the process till there are no more leading entries.
    8. The so obtained matrix is reduced to row echelon form.

- **Rank of a matrix:** A matrix is said to be of rank ‘r’ if:
    - There is at least one minor of order ‘r’ which is not equal to zero and
    - Every minor of order $(r + 1)$ is equal to zero.

    In other words, “The order of the largest non-zero minor of a matrix is called the rank of a matrix.”

    Rank of a matrix A is denoted by $\rho(A)$.

    Thus, we can deduce the following from the definition of the rank of the matrix:

    1. $\rho(A) \geq r$, if there exists a non-zero minor of order r.

2. $\rho(A) \le r$, if all minors of order ‘ $r + 1$’ are zero.
3. $\rho(A) \le \min\{m, n\}$ for any $m \times n$ matrix A.
4. Rank of every non-zero matrix is ≥ 1.
5. Rank of a zero matrix is 0.

Note: Equivalent matrices have same rank. i.e. Rank of a matrix remains unaltered by elementary transformations.

- We can find the rank of a matrix in different ways:
  - Using definition, i.e. by evaluating the minors of the matrix.
    - Begin with the highest order minor(s) of say order ‘r’ for the given matrix. If at least one of them is non-zero, then $\rho(A) = r$.
    - If all minors of order ‘r’ are zero, then begin with minor(s) of order ‘$r - 1$’ and so on till you get a non-zero minor.
    - The rank of the matrix is the order of the so obtained first non-zero minor.
  - By reducing into row echelon form (REF).
    - Reduce the given matrix into REF.
    - The number of non-zero rows obtained in REF is the rank of the matrix.
- **Gauss-Jordan method for finding the inverse of a matrix**: The method of finding the inverse of a given non-singular matrix by elementary row transformations is called *Gauss-Jordan method*.

  Working rules:
  1. Write the given square matrix A and the identity matrix I of the same order side by side as $[\mathrm{A} : \mathrm{I}]$.
  2. Apply series of elementary row transformations simultaneously on both A and I such that matrix A reduces to I. In this process, the identity matrix I reduces to A^{-1} giving the matrix of the form $[\mathrm{I} : \mathrm{A}^{-1}]$.
- **System of Simultaneous Linear Equations (SLE)**

  Consider a set of ‘m’ linear equations in ‘n’ unknowns $x_1, x_2, \ldots, x_n$.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

  These can be written in matrix form as:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \text{ or AX = B}$$

Where the matrix $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} \end{bmatrix}$ is called the co-efficient matrix and

$\begin{bmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & \cdots & a_{mn} & b_m \end{bmatrix}$ is called the *augmented matrix* denoted by [A : B]

- An SLE, AX = B is called a *homogeneous linear equation* if B = O. i.e. a homogeneous linear equation is of the form AX = O.
- An SLE, AX = B is called a *non-homogeneous linear equation* if B ≠ O.
- For a system of linear equations, there are three possibilities:
  - It has a unique solution
  - It has infinitely many solutions
  - It has no solution
- If an SLE has one or more solution, the SLE is said to be *consistent* system of equations.
- If an SLE does not have any solution, the SLE is said to be *inconsistent* system of equations.
- In an SLE, the solution X = O, is called a *trivial* solution.
- **Theorem:** A system equations AX = B is consistent iff the coefficient matrix and the augmented matrix have the same rank.
- ***Condition for the consistency of system of equations AX = B.***
  - For a non-homogeneous linear equation AX = B, i.e. where B ≠ O, Consider a system of 'm' linear equations in 'n' variables (unknowns) with $m \geq n$. Find ρ(A) and ρ[A : B].
    - If ρ[A : B] = ρ(A) = n, then the system is consistent with unique solution.
    - If ρ[A : B] = ρ(A) < n, then the system is consistent with infinite solutions.
    - If ρ[A : B] ≠ ρ(A), then the system is inconsistent and hence it has no solution.

We note that if $m < n$, then the system always has infinite or no solution.

  - For a homogeneous linear equation AX = O,
    - X = O, i.e. trivial solution is always a solution. Hence, a homogeneous linear equation is always consistent.

- If $\rho(A) = n$, the system have a trivial solution only.
    - If AX = O has 'n' linear equations in 'n' unknowns with $\rho(A) = n$, then $|A| \neq 0$ in this case.
- If $\rho(A) < n$, then the system has infinitely many solutions.

❖ **Eigen values and Eigen vectors** play an important role in the study of ordinary differential equations. Applications like analyzing population growth model and in calculating power of matrices, these are used. Vibration of beams, probability (Markov process), Economics (Leontif model), genetics, quantum mechanics, population dynamics and geometry are few of the applications of Eigen values and Eigen vectors. In a mechanical system normal modes of vibration are represented by Eigen vectors.

❖ **Definition**: Let A be a square matrix of order n. If there exists a non-zero column vector X such that $AX = \lambda X$ for some scalar λ, then X is called the Eigen vector of A. The corresponding scalar λ is called the Eigen value of A.

Now, $AX = \lambda X \Rightarrow AX - \lambda X = O \Rightarrow (A - \lambda I)X = O$

❖ For a given square matrix A of order n, we have the following:

1. $A - \lambda I$ is called the characteristic matrix where λ is a scalar and I is a unit matrix of order n.
2. $\det(A - \lambda I)$ or $|A - \lambda I|$ is called the characteristic polynomial of matrix A.
3. $|A - \lambda I| = 0$ is called the characteristic equation of matrix A.
4. Solutions or roots of $|A - \lambda I| = 0$ are called characteristic roots or Eigen values or characteristic values or latent roots or proper values.
5. The set of Eigen values of a matrix is called its spectrum.

❖ For a given matrix A of order n, its characteristic equation is of degree n.

❖ **Method to find the characteristic equation of a matrix of order 3.**

For a matrix A of order 3, its characteristic equation is of degree 3. Assume that the characteristic equation is given by: $\lambda^3 - S_1\lambda^2 + S_2\lambda - S_3 = 0$. The co-efficients are given by:

1. S_1 = sum of diagonal elements of A = trace(A) = $a_{11} + a_{22} + a_{33}$.
2. S_2 = sum of minors of diagonal elements = $M_{11} + M_{22} + M_{33}$.
   [Minor M_{ij} of an element a_{ij} is obtained by taking the determinant of the matrix obtained by deleting the row and the column in which a_{ij} lies.]
3. $S_3 = \det(A) = |A|$.

❖ Method to find the characteristic equation of a matrix of order 2.

For a matrix A of order 2, its characteristic equation is of degree 2. Assume that the characteristic equation is given by: $\lambda^2 - S_1\lambda + S_3 = 0$. The co-efficients are given by:

1. S_1 = sum of diagonal elements of A = trace(A) = $a_{11} + a_{22} + a_{33}$.
2. S_3 = det(A) = |A|.

❖ Properties of Eigen Values

1. Eigen values of an upper triangular matrix or lower triangular matrix are the same as the diagonal elements on its principal diagonal.
2. For any square matrix A and its transpose A^T have same Eigen values.
3. Square matrix A is non-singular if an only if all of its Eigen values are non-zero.
4. Assuming that A is a non-singular matrix, the Eigen values of A^{-1} are the reciprocals of the Eigen values of A.
5. For a non-zero scalar k, the Eigen values of kA are the same as the values obtained by multiplying each Eigen value of A by k.
   i.e. if λ is the Eigen value of A, then $k\lambda$ is the Eigen value of kA.
6. If λ is the Eigen value of A, then λ^m is the Eigen value of A^m.
7. Eigen values of a symmetrical matrix are always real.
8. Eigen values of a skew-symmetric matrix are either purely imaginary or 0.

❖ Properties of Eigen Vectors

1. An Eigen vector corresponding to an Eigen value is not unique.
2. Eigen vectors for A and A^T are different unless A is a symmetric matrix.
3. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are distinct Eigen values then the corresponding Eigen vectors $X_1, X_2, \ldots, X_n$ are distinct and form a linearly independent set.
   [Vectors $v_1, v_2, \ldots, v_n$ are said to be linearly independent vectors if $a_1 \cdot v_1 + a_2 \cdot v_2 + \ldots + a_n \cdot v_n = \theta \Rightarrow a_1 = a_2 = \ldots = a_n = 0$. Otherwise the vectors are said to be linearly dependent.]
4. Eigen vectors of A, A^{-1}, A^n are same.
5. If two or more Eigen values are equal for a square matrix A of order n, then it may or may not be possible to get linearly independent Eigen vectors corresponding to the repeated Eigen value.
6. Eigen vectors of a symmetric matrix corresponding to different Eigen values are orthogonal.
   [Two vectors X_1 and X_2 are said to be orthogonal if $X_1^T \cdot X_2 = 0$.

7. Same Eigen vector cannot correspond to two different Eigen values.
8. If X is an Eigen vector corresponding to the Eigen value λ for some matrix A, then kX is also an Eigen vector corresponding to λ, where k is a non-zero scalar.

- **Cayley-Hamilton Theorem:** Every square matrix satisfies its own characteristic equation.
- A matrix A is said to be **similar** to a matrix B, if there exists a non-singular
  matrix P such that $B = P^{-1}AP$.
- A matrix A is said to be **diagonalizable** if it is similar to a diagonal matrix.
- If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are different Eigen values and $X_1, X_2, \ldots, X_n$ are corresponding Eigen vectors, then the Eigen vectors form a linearly independent set of vectors.
- A matrix of order n is diagonalizable if all the roots of its characteristic equations are real and distinct.
- A matrix of order n is diagonalizable if and only if it has n linearly independent Eigen vectors.
- Here if A is similar to a diagonal matrix D with $P^{-1}AP$, then
  1. P is called the modal matrix whose columns are n linearly independent Eigen vectors of A.
  2. Matrix D has its diagonal elements as the Eigen values of A.
  3. Matrix D is called the spectral matrix.
- A square matrix A is said to be orthogonally diagonalizable if there exists an orthogonal matrix M such that $D = M^TAM$ is a diagonal matrix.
- A square matrix A is orthogonally diagonalizable if and only if it is a symmetric matrix.
- Symmetric matrices are always diagonalizable.