Contents

Chapter 1

Matrix and Determinant

1.1 Introduction

The study of matrices and determinants has wide applications in science, engineering, and computer science. A matrix is a rectangular array of numbers arranged in rows and columns, used to represent data or systems of linear equations in a compact form. Matrices allow operations such as addition, multiplication, and inversion, making them powerful tools for solving complex problems.

A determinant, associated with a square matrix, is a scalar value that provides key insights into the properties of the matrix. Determinants help in checking whether a system of linear equations has a unique solution, computing the inverse of a matrix, and understanding geometric concepts such as area, volume, and transformations. Together, matrices and determinants provide the foundation for higher concepts like vector spaces, linear transformations, eigenvalues, and eigenvectors, making this module essential for deeper studies in applied mathematics and related fields.

1.2 Applications of Matrices and Determinants in Engineering

- **Computer Graphics & Image Processing:** Rotation, scaling, translation, and pixel manipulation using transformation matrices.
- **Machine Learning, AI & Data Science:** Representation of datasets, regression

models, neural networks, and big data handling.

- **Cryptography & Network Security:** Message encoding/decoding and secure algorithms using invertible matrices and determinants.
- **Graph Theory & Network Analysis:** Adjacency and incidence matrices for analyzing computer and electrical networks.
- **Signal Processing & Communications:** Fourier transforms, filter design, and antenna array beam-forming using matrix methods.
- **Circuit Analysis:** Solving mesh and nodal equations through determinants (Cramer's Rule).
- **Control Systems:** State-space representation and stability analysis with eigenvalues and eigenvectors.
- **Power Systems & Electrical Machines:** Load flow analysis, stability assessment, and impedance matrix representation.
- **Structural & Mechanical Engineering:** Stress-strain relations, stiffness matrices, truss and frame analysis, and finite element method applications.
- **Vibrations & Robotics:** Solving equations of motion, kinematic transformations, and robotic motion planning using matrices.
- **Surveying & Civil Engineering Applications:** Coordinate transformations, mapping, and structural stability analysis.
- **Chemical Engineering:** Solving material/energy balance equations, reaction rate equations, and transport phenomena models using matrices.
- **Database Systems & IT Applications:** Data representation, indexing, and relation mapping with matrices.

Definition: A matrix is a rectangular array of numbers - in other words, numbers grouped into rows and columns. We use matrices to represent and solve systems of linear equations. For example, the system of equations

$$\begin{aligned} 8y + 16z &= 0 \\ x - 3z &= 1 \\ -4x + 14y + 2z &= 6 \end{aligned}$$

can be represented by what is called an augmented matrix as seen below:

$$\begin{array}{lcl} \text{Row 1 } (R_1) & \rightarrow & \begin{bmatrix} 0 & 8 & 16 & 0 \end{bmatrix} \\ \text{Row 2 } (R_2) & \rightarrow & \begin{bmatrix} 1 & 0 & -3 & 1 \end{bmatrix} \\ \text{Row 3 } (R_3) & \rightarrow & \begin{bmatrix} -4 & 14 & 2 & 6 \end{bmatrix} \end{array}$$

$$\begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ x & y & z & \text{constant} \end{array}$$

Solving a system of equations using a matrix means using row operations to get the matrix into the form called *reduced row echelon form* like the example below:

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

The last column can have any numbers.

1.2.1 Elementary row operation on a matrix

Let A be a matrix of size $m \times n$. An elementary row operation on A is an operation of the following three types:

(i) Interchanging any two rows of A.

(ii) Multiplication of a row of A by a non-zero number.

(iii) Addition of any row of A multiplied by any number to another row of A.

Similarly

1.2.2 Elementary column operation.

If B is obtained from A by applying a finite no. of elementary row operations the matrices A & B are called **'Row equivalent'**. We can perform elementary row operations on a matrix to solve the system of linear equations it represents. Following are some examples of these three types of row operations.

1.2.3 Interchanging two rows

Rows can be moved around by switching any two. In this case, R_1 and R_2 have been switched.

$$\begin{bmatrix} 0 & 8 & 16 & 0 \\ 1 & 0 & -3 & 1 \\ -4 & 14 & 2 & 6 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 0 & -3 & 1 \\ 0 & 8 & 16 & 0 \\ -4 & 14 & 2 & 6 \end{bmatrix}$$

1.2.4 Multiplying a row by a nonzero constant

We can multiply any row by any number except 0. When a row is multiplied by a number, every element in that row must be multiplied by the same number. Below, R_2 is multiplied by 2.

$$\begin{bmatrix} 1 & 0 & -3 & 1 \\ 0 & 8 & 16 & 0 \\ -4 & 14 & 2 & 6 \end{bmatrix} \xrightarrow{2R_2 \to R_2} \begin{bmatrix} 1 & 0 & -3 & 1 \\ 0 & 16 & 32 & 0 \\ -4 & 14 & 2 & 6 \end{bmatrix}$$

1.2.5 Adding a multiple of a row to another row

We may also multiple a row by any number except 0 and add the results to another row.

$$\begin{bmatrix} 1 & 0 & -3 & 1 \\ 0 & 8 & 16 & 0 \\ -4 & 14 & 2 & 6 \end{bmatrix} \xrightarrow{4R_1 + R_3 \to R_3} \begin{bmatrix} 1 & 0 & -3 & 1 \\ 0 & 8 & 16 & 0 \\ 0 & 14 & -10 & 10 \end{bmatrix}$$

1.3 Echelon matrix

A matrix A is said to be Echelon matrix if

(i) all zero-rows of A follow all non-zero rows of A.

(ii) the no. of zeros preceding the first non-zero element of a row increases as we pass from row to row downwards.

Example:

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

1.4 Rank of a matrix

Let A be a non-zero matrix. A natural number r is said to be rank of A if

(i) there is at least one r-th order non-singular square sub-matrix of A.

(ii) every square sub-matrix of A of order greater than r is singular.

Example 1. *Find the rank of* $A = \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}$ *using definition.*

Solution: $|A| = \begin{vmatrix} 2 & 3 \\ 2 & 3 \end{vmatrix} = -6 + 6 = 0$. So, the rank of A must be less than 2. Now, one 1st order submatrix (2) is non singular as $|2| = 2 \neq 0$. $\therefore$ Rank of A = 1.

Example 2. *Find the rank of the matrix* $\begin{pmatrix} 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ *using definition.*

Solution: All 3rd order submatrices are $({}^4C_3 = 4)$ $\begin{pmatrix} 1 & 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. All are singular. Now, one 2nd order sub-matrix is $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$, which is non singular. So, rank of $A = 2$ (by definition).

Theorem 1. *If an Echelon matrix has r number of non-zero rows, then the rank of this Echelon matrix is r.*

Example 3. *Find the rank of the matrix* $A = \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 2 & 6 & 2 & 6 & 2 \\ 3 & 9 & 1 & 10 & 6 \end{pmatrix}$.

Solution: Here A =

$$\begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 2 & 6 & 2 & 6 & 2 \\ 3 & 9 & 1 & 10 & 6 \end{pmatrix} \xrightarrow{R_3 \to R_3 - 2R_1} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & -2 & -2 & 0 \\ 3 & 9 & 1 & 10 & 6 \end{pmatrix} \xrightarrow{R_4 \to R_4 - 3R_1} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & -2 & -2 & 0 \\ 0 & 0 & -5 & -2 & 3 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 + R_2} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & -2 & 3 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_4} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & -5 & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{R_3 \to R_3 + \frac{5}{2} R_2} \begin{pmatrix} 1 & 3 & 2 & 4 & 1 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B(say)$$

, which is an Echelon matrix. Now, rank of A = No. of non zero rows in B = 3.

Example 4. *Find the rank of* $A = \begin{pmatrix} 0 & 0 & 1 & 2 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ -2 & 2 & 1 & 1 & 0 \end{pmatrix}$

Solution: Here,

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ -2 & 2 & 1 & 1 & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} 2 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ -2 & 2 & 1 & 1 & 0 \end{pmatrix} \xrightarrow{R_3 \to R_3 + R_1} \begin{pmatrix} 2 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 2 & 2 & 2 & 0 \end{pmatrix} \xrightarrow{R_3 \leftrightarrow R_2}$$

$$\begin{pmatrix} 2 & 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{pmatrix} = B(say)$$

which is an Echelon matrix. Therefore, the Rank of $A = 3$.

1.5 Nullity of a matrix

The nullity of a matrix refers to the dimension of its null space, which consists of all vectors that satisfy the homogeneous equation $Ax = 0$. In simpler terms, it represents the count of free variables in the system or, equivalently, the number of columns in the matrix that do not contain leading entries once the matrix is expressed in reduced row echelon form. To determine nullity, one often applies the **Rank-Nullity Theorem**, which states that the sum of a matrix's rank (dimension of its column space) and its nullity is equal to the total number of columns in the matrix. That means, for a matrix A of size $m \times n$ $rank(A) + nullity(A) = n =$ No. of Columns. For example, for the previous question, rank = 3, No. of Columns = 5 $\therefore$ Nullity = 5-3 = 2.

1.6 Exercises

1. Find the rank of

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

2. Find the rank of

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

3. For what value(s) of k does the matrix

$$C = \begin{bmatrix} 1 & k & 1 \\ 0 & 1 & k \\ k & 1 & 0 \end{bmatrix}$$

have rank less than 3?
4. Find the nullity of

$$D = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 8 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

5. Find x such that rank of

$$E = \begin{bmatrix} 2 & 1 & 4 \\ 1 & x & 2 \\ 4 & 0 & x+2 \end{bmatrix}.$$

is 2
6. A 4×6 matrix has rank 3. What is its nullity?
7. True or False: If a square matrix A of order n has $\det(A) \neq 0$, then

(a) $\text{rank}(A) = n$

(b) $\text{nullity}(A) = 0$

8. Suppose A is a 5×7 matrix and $\text{nullity}(A) = 4$. Find the rank of A.

1.7 System of Linear Equations

A **linear equation** in variables $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \ldots, a_n, b$ are constants. The constant a_i is called the *coefficient* of x_i; b is the *constant term*.

A **system of linear equations** is a collection of such equations in the same variables. For example, a system with m equations and n variables:

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

A *solution* is a tuple $(s_1, s_2, \ldots, s_n)$ that satisfies all equations.

$$\text{Let } A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \text{called the coefficient matrix,}$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Then, the above system of linear equations can be written in the form $AX = B$

1.7.1 System of homogeneous and non-homogeneous system of linear equations

If $b_1 = b_2 = b_3 = \cdots b_m = 0$ then $B = 0$ and the system $AX = B$ is known as homogeneous system of linear equation $AX = 0$. That is

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

If at least one of $b_1,\ b_2,\ b_3, \cdots, b_m$ is non-zero, then $B \neq 0$. Such a system of equations is called a non-homogeneous system of linear equations.

1.7.2 consistency and inconsistency of the system of linear equations

A system of linear equations is called an inconsistent system of linear equations if it has no solution. Again the system is called a consistent system of linear equations having one or more solution .

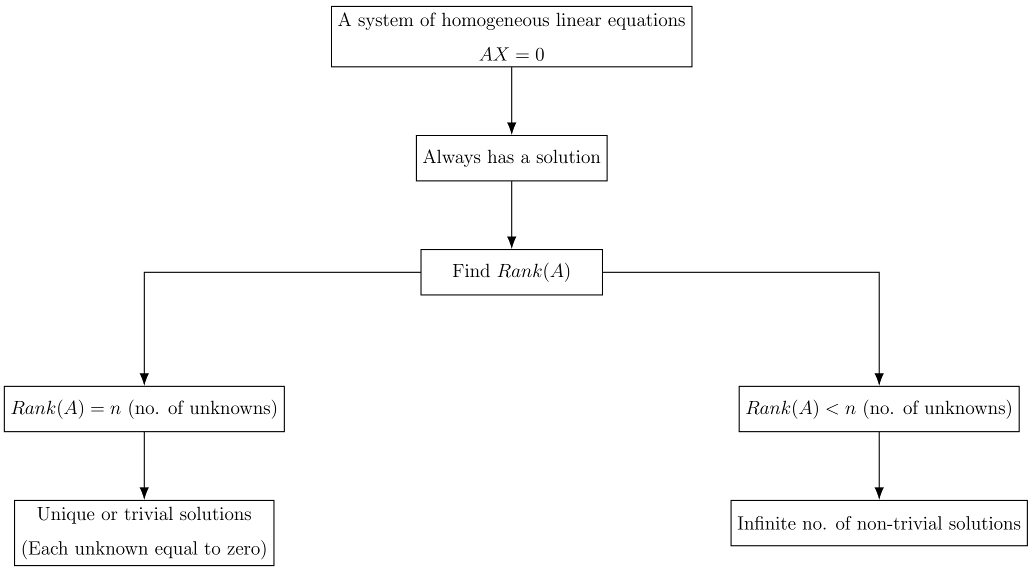
For a system of homogeneous linear equations

A homogeneous system $AX = 0$ is always consistent since $X = 0$ is always a solution. The solution in which every variable takes the value zero is known as the Null Solution or the Trivial Solution. Thus, a homogeneous system is always consistent.

A system of homogeneous linear equations has either the trivial solution or an infinite number of solutions.

(i) If $Rank(A) = number\ of\ unknowns$, the system has only the trivial solution.

(ii) If $Rank(A) < number of unknowns$, the system has an infinite number of non-trivial solutions.



Example 5. *Solve the following homogeneous system:*

$$x + y + 3z = 0$$

$$2x + y + z = 0$$

$$3x + 2y + 4z = 0$$

$$x + y + 3z = 0, \quad 2x + y + z = 0, \quad 3x + 2y + 4z = 0.$$

Solution. It is clear that it has a trivial (zero) solution $(0, 0, 0)$. Here the number of equations $m = 3$ and number of unknowns $n = 3$. The system is consistent.

Now we write the system as $AX = 0$, where

$$A = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 1 \\ 3 & 2 & 4 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad O = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Applying elementary row operations to A:

$$\begin{bmatrix}1&1&3\\2&1&1\\3&2&4\end{bmatrix}\xrightarrow{R_2\to R_2-2R_1,\ R_3\to R_3-3R_1}\begin{bmatrix}1&1&3\\0&-1&-5\\0&-1&-5\end{bmatrix}\xrightarrow{R_3\to R_3-R_2}\begin{bmatrix}1&1&3\\0&-1&-5\\0&0&0\end{bmatrix}.$$

Multiplying the second row by -1:

$$\begin{bmatrix}1&1&3\\0&1&5\\0&0&0\end{bmatrix}.$$

Thus the rank of A is $\rho(A)=2<3$ (number of unknowns). Therefore, a non-trivial solution exists.

So the given system is equivalent to:

$$\begin{bmatrix}1&0&-2\\0&1&5\\0&0&0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Hence the equations reduce to

$$x-2z=0,\qquad y+5z=0.$$

Let $z=t$ (a parameter). Then

$$x=2t,\quad y=-5t,\quad z=t,\quad t\in\mathbb{R}.$$

Thus, the system has infinitely many non-trivial solutions.

$$(x,y,z)=(2t,\ -5t,\ t),\quad t\in\mathbb{R}.$$

For a system of non-homogeneous linear equations

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

be a system of non-homogeneous linear equations with m equations and n unknowns.

Let $Rank(A)$ denote the rank of the coefficient matrix A and $Rank([A : B])$ denote the rank of the augmented matrix obtained by appending the column B to A. The **augmented matrix** of the system above is:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

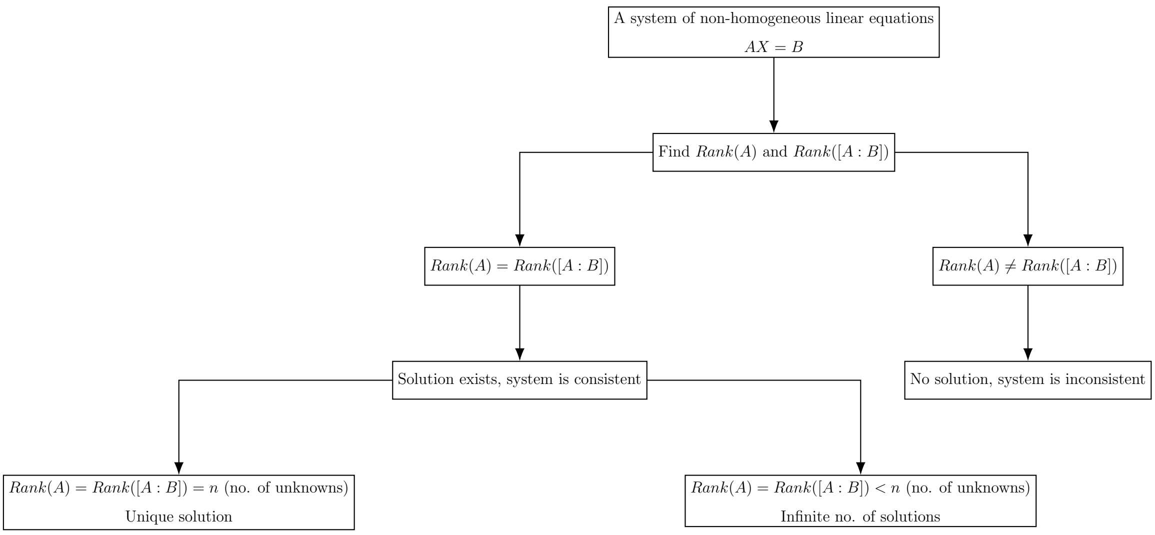
Any system of non-homogeneous linear equations has one of the following:

- No solution,
- A unique solution,
- Infinitely many solutions.

(i) If $Rank(A) \neq Rank([A : B])$, the system is inconsistent (no solution).

(ii) If $Rank(A) = Rank([A : B]) =$ number of unknowns, the system has a unique solution.

(iii) If $Rank(A) = Rank([A : B]) <$ number of unknowns, the system has infinitely many solutions.



Example 6. *Check the consistency of the following system of equations and solve if possible:*

$$x + y + z = 1$$

$$2x + y + 2z = 2$$

$$3x + 2y + 3z = 5$$

Solution. The system of linear equations can be written in matrix form as $AX = B$, i.e.,

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 2 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}.$$

The coefficient matrix of the system of equations is

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

The augmented matrix is

$$[A : B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 2 \\ 3 & 2 & 3 & 5 \end{array}\right].$$

Applying elementary row operations on the augmented matrix $[A : B]$, we have

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 2 \\ 3 & 2 & 3 & 5 \end{array}\right] \xrightarrow{R_2 \leftarrow R_2 - 2R_1,\ R_3 \leftarrow R_3 - 3R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 - 3R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right].$$

Here, $Rank([A : B]) = 3$ and $Rank(A) = 2$.
Since $Rank([A : B]) \neq Rank(A)$, the given system of equations is inconsistent.
In other words, the system does not have any solution.

Example 7. *Solve the following if possible*

$$x + y + z = 3,$$

$$x + 2y + 3z = 4,$$

$$x + 4y + 9z = 6.$$

Solution. The augmented matrix is

$$[A:B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 6 \end{array}\right].$$

Performing row operations (subtract Row1 from Row2 and Row3, then reduce) one convenient sequence is:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 6 \end{array}\right] \xrightarrow{R_2 \leftarrow R_2 - R_1,\ R_3 \leftarrow R_3 - R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 3 & 8 & 3 \end{array}\right]$$
$$\xrightarrow{R_3 \leftarrow R_3 - 3R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 0 \end{array}\right].$$

Thus $Rank(A) = Rank([A:B]) = 3$, so the system has a unique solution. Back-substitution gives

$$\begin{aligned} x + y + z &= 3, \\ y + 2z &= 1, \\ 2z &= 0. \end{aligned}$$

$$\text{Which gives} \quad z = 0, \quad y = 1, \quad x = 2.$$

Example 8. *Investigate for which values of parameters λ and μ the system*

$$x + y + z = 6,$$

$$x + 2y + 3z = 10,$$

$$x + 2y + \lambda z = \mu$$

has (i) no solution, (ii) a unique solution, (iii) infinitely many solutions.

Solution. The augmented matrix is

$$[A:B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 10 \\ 1 & 2 & \lambda & \mu \end{array}\right].$$

Subtracting the first row from rows 2 and 3 gives

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 1 & \lambda-1 & \mu-6 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & \lambda-3 & \mu-10 \end{array}\right].$$

We thus observe the following:

(i) If $\lambda \neq 3$, then $Rank(A) = 3 =$ No. of unknown. Then the system has a unique solution for any μ.

(ii) If $\lambda = 3$ and $\mu \neq 10$, then $Rank(A) = 2$ but $Rank([A : B]) = 3$, hence the system is inconsistent (no solution).

(iii) If $\lambda = 3$ and $\mu = 10$, then $Rank(A) = Rank([A : B]) = 2 < 3$ (No. of unknown), so the system has infinitely many solutions.

Example 9. *Investigate for what values of λ and μ the equations*

$$x + y + z = 6,$$

$$x + 2y + 4z = 10,$$

$$2x + 3y + \lambda z = \mu$$

have (1) no solution, (2) a unique solution, and (3) infinitely many solutions.

Solution. The augmented matrix is

$$[A : B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 1 & 2 & 4 & 10 \\ 2 & 3 & \lambda & \mu \end{array}\right].$$

At first performing the operations $R_{31} \to R_3 - 2R_1$ and $R_2 \to R_2 - R_1$, and then as follows

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 3 & 4 \\ 0 & 1 & \lambda-2 & \mu-12 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 6 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & \lambda-5 & \mu-16 \end{array}\right].$$

We thus observe the following:

(i) If $\lambda \neq 5$, then $Rank(A) = 3 =$ No. of unknown. Then the system has a unique solution for any μ.

(ii) If $\lambda = 5$ and $\mu \neq 16$, then $Rank(A) = 2$ but $Rank([A : B]) = 3$, hence the system is inconsistent (no solution).

(iii) If $\lambda = 5$ and $\mu = 16$, then $Rank(A) = Rank([A : B]) = 2 < 3$ (No. of unknown), so the system has infinitely many solutions.

Example 10. *Determine the values of a and b so that the system*

$$\begin{aligned} 2x + 3y + 4z &= 9 \\ x - 2y + az &= 5 \\ 3x + 4y + 7z &= b \end{aligned}$$

has (i) a unique solution, (ii) infinitely many solutions, and (iii) no solution.

Solution. The augmented matrix is

$$[A : B] = \left[\begin{array}{ccc|c} 2 & 3 & 4 & 9 \\ 1 & -2 & a & 5 \\ 3 & 4 & 7 & b \end{array}\right].$$

Swap $R_1 \leftrightarrow R_2$, then eas follows:

$$\left[\begin{array}{ccc|c} 1 & -2 & a & 5 \\ 2 & 3 & 4 & 9 \\ 3 & 4 & 7 & b \end{array}\right] \xrightarrow{R_2 \leftarrow R_2 - 2R_1,\ R_3 \leftarrow R_3 - 3R_1} \left[\begin{array}{ccc|c} 1 & -2 & a & 5 \\ 0 & 7 & 4-2a & -1 \\ 0 & 10 & 7-3a & b-15 \end{array}\right].$$

$$\left[\begin{array}{ccc|c} 1 & -2 & a & 5 \\ 0 & 7 & 4-2a & -1 \\ 0 & 10 & 7-3a & b-15 \end{array}\right] \xrightarrow{R_3 \leftarrow 7R_3 - 10R_2} \left[\begin{array}{ccc|c} 1 & -2 & a & 5 \\ 0 & 7 & 4-2a & -1 \\ 0 & 0 & 9-a & 7b-95 \end{array}\right].$$

We thus observe the following:

(i) If $a \neq 9$, then $Rank(A) = 3 =$ No. of unknown. Then the system has a unique solution for any b.

(ii) If $a = 9$ and $7b - 95 \neq 0 \implies b \neq \frac{95}{7}$, then $Rank(A) = 2$ but $Rank([A : B]) = 3$, hence the system is inconsistent (no solution).

(iii) If $a = 9$ and $b = \frac{95}{7}$, then $Rank(A) = Rank([A : B]) = 2 < 3$ (No. of unknown), so the system has infinitely many solutions.

Example 11. *Find out for what values of λ, the equations*

$$x + y + z = 1$$

$$x + 2y + 4z = \lambda$$

$$x + 4y + 10z = \lambda^2$$

have a solution and solve completely in each case.

Solution. The augmented matrix is

$$[A : B] = \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \lambda \\ 1 & 4 & 10 & \lambda^2 \end{array}\right].$$

By the operations $R_{21} \to R_{21} - R_{11}$ and $R_{31} \to R_{31} - R_{11}$, we get

$$[A|B] \sim \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \lambda - 1 \\ 0 & 3 & 9 & \lambda^2 - 1 \end{array}\right].$$

By $R_{32} \to R_{32} - 3R_{22}$, we have

$$[A|B] \sim \left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \lambda - 1 \\ 0 & 0 & 0 & \lambda^2 - 3\lambda + 2 \end{array}\right] \qquad (1)$$

Whence rank$(A) = 2$. The system of equations will be consistent if rank$[A|B] =$ rank$(A) = 2$.

Hence we must have

$$\lambda^2 - 3\lambda + 2 = 0, \quad \text{so that } \lambda = 1 \text{ or } 2.$$

From (1), an equivalent system of equations is

$$x + y + z = 1, \qquad y + 3z = \lambda - 1. \qquad (2)$$

Since $Rank(A) = Rank([A : B]) = 2 < n = 3$ (the number of unknowns), there will be an infinite number of solutions.

Taking $z = k$, where k is an arbitrary constant, and solving the equations (2), we obtain

$$x = 2k - \lambda + 2, \qquad y = \lambda - 1 - 3k, \qquad z = k.$$

So, for $\lambda = 1$, the solution is

$$x = 2k + 1, \qquad y = -3k, \qquad z = k.$$

Again, for $\lambda = 2$, the solution is

$$x = 2k, \qquad y = 1 - 3k, \qquad z = k.$$

1.7.3 Matrix Inversion method

Consider a non-homogeneous system of n- equations with n unknowns

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

Its matrix form is $AX = B$, where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

First find the determinant of the coefficient matrix A i.e. $|A|$ Now, If $|A| \neq 0$, then the system is consistent and has unique solution given by

$$X = A^{-1}B$$

Example 12. *Test the consistency of the following system of equations and, if consistent, find the solution:*

$$\begin{aligned} 2x - y + 3z &= 9 \\ x + y + z &= 6 \\ x - y + z &= 2. \end{aligned}$$

Solution. The matrix form of the given equations is

$$\begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 6 \\ 2 \end{bmatrix},$$

In matrix form, we can write $AX = B$.

$$\text{where} \quad A = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 9 \\ 6 \\ 2 \end{bmatrix}.$$

We have

$$|A| = 2(1+1) + 1(1-1) + 3(-1-1) = 2 \cdot 2 + 1 \cdot 0 + 3(-2) = 4 + 0 - 6 = -2 \neq 0.$$

Since $|A| \neq 0$, the given system of equations is consistent and has a unique solution given by

$$X = A^{-1}B \tag{1}$$

Now, we compute A^{-1}, so we find cofactors.

$$C_{11} = (-1)^{1+1} \det \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = (1)(1+1) = 2.$$

Similarly, one finds:

$$C_{12} = 0, \quad C_{13} = -2, \quad C_{21} = -2,\ C_{22} = -1,\ C_{23} = 1, \quad C_{31} = -4,\ C_{32} = 1,\ C_{33} = 3.$$

Thus,

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

Hence,

$$A^{-1} = \frac{1}{|A|} \operatorname{adj}(A) = \frac{1}{-2} \begin{bmatrix} 2 & -2 & -4 \\ 0 & -1 & 1 \\ -2 & 1 & 3 \end{bmatrix}.$$

Now, by (1),

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B = \frac{1}{-2} \begin{bmatrix} 2 & -2 & -4 \\ 0 & -1 & 1 \\ -2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 9 \\ 6 \\ 2 \end{bmatrix}.$$

Carrying out the multiplication,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{-2} \begin{bmatrix} 2(9) - 2(6) - 4(2) \\ 0(9) - 1(6) + 1(2) \\ -2(9) + 1(6) + 3(2) \end{bmatrix} \implies \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

Hence the required solution is

$$x = 1, \quad y = 2, \quad z = 3.$$

EXERCISE

(i) Solve the following equation

$$\begin{aligned} x + y + z &= 9, \\ 2x + 5y + 7z &= 52, \\ 2x + y - z &= 0. \end{aligned}$$

(ii) Examine the consistency of the following systems of equations, and if consistent find the complete solutions:

$$\begin{aligned} 2x + 4y - z &= 9, \\ 3x - y + 5z &= 5, \\ 8x + 2y + 9z &= 19. \end{aligned}$$

(iii) Show that the following systems of equations are consistent. Also solve them.

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

(iv) Examine the consistency of the following system of equations and solve, when possible:

(i)

$$\begin{aligned} 2x_1 - 2x_2 - 4x_3 &= 8, \\ 2x_1 + 3x_2 + 2x_3 &= 8, \\ -x_1 + x_2 - x_3 &= \tfrac{7}{2}. \end{aligned}$$

(ii)

$$\begin{aligned} 5x + 3y + 7z &= 4, \\ 3x + 26y + 2z &= 9, \\ 7x + 2y + 10z &= 5. \end{aligned}$$

(iii)

$$\begin{aligned} x_1 + x_2 + x_3 &= 6, \\ x_1 - x_2 - x_3 &= -4, \\ x_1 + x_2 - x_3 &= 0, \\ 2x_1 - 2x_3 &= 4. \end{aligned}$$

(iv)

$$\begin{aligned} x + 2y - z &= 10, \\ x - y - 2z &= -2, \\ 2x + y - 3z &= 8. \end{aligned}$$

(v) Find the values of a and b so that the following system of equations have (a) a unique solution, (b) no solution, and (c) infinitely many solutions.

(i)

$$
\begin{aligned}
2x + 3y + 5z &= 9,\\
7x + 3y - 2z &= 8,\\
2x + 3y + az &= b.
\end{aligned}
$$

(ii)

$$
\begin{aligned}
x + y + z &= b,\\
2x + y + 3z &= b + 1,\\
5x + 2y + az &= b^2.
\end{aligned}
$$

(iii)

$$
\begin{aligned}
3x - 2y + z &= b,\\
5x - 8y + 9z &= 3,\\
2x + y + az &= -1.
\end{aligned}
$$

1.8 Cayley-Hamilton Theorem

The origins of the Cayley-Hamilton Theorem trace back to the 19$^{\text{th}}$ century, a period marked by significant advancements in algebra and the study of linear transformations. In 1853, the British mathematician Arthur Cayley published a seminal paper in which he demonstrated that a square matrix satisfies its own characteristic equation, though his proof was initially restricted to matrices of order two. Around the same time, the Irish mathematician William Rowan Hamilton, renowned for his work on quaternions, had already explored related ideas concerning algebraic equations of linear operators. Their combined contributions laid the foundation for what later became formally known as the Cayley-Hamilton Theorem. The theorem was further refined and generalized by later mathematicians, who extended Cayley's original insight to matrices of arbitrary order and provided rigorous proofs using determinants and properties of linear transformations. Its recognition as a fundamental result in linear algebra gradually took shape during the late 19$^{\text{th}}$ and early 20$^{\text{th}}$ centuries, when the abstract theory of matrices began to crystallize as an independent field. Historically, the Cayley-Hamilton Theorem not only marked a milestone in matrix theory but also served as an essential stepping stone toward modern algebra. By establishing a direct connection between matrices and their characteristic polynomials, it opened avenues for new methods in solving linear equations, analyzing transformations, and developing computational techniques. Today, the theorem is regarded as both a classical achievement in mathematics and a continuing source of applications across science and engineering.

In essence, the Cayley-Hamilton Theorem asserts that every square matrix satisfies its own characteristic equation.

Let $A = (a_{ij})_{n\times n}$ be an $n \times n$ matrix over a filed F. Then $det(A - \lambda I_n)$ is said to be the characteristic polynomial of the matrix A, denoted as $P_A(\lambda)$. The equation $P_A(\lambda) = 0$ is said to be the characteristic equation of the matrix A.

Let the the characteristic polynomial $P_A(\lambda)$ of A be expressed as

$$P_A(\lambda) = det(A - \lambda I_n) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + ... + c_n. \tag{1.1}$$

Then, according to the Cayley–Hamilton theorem

$$P_A(A) = A^n + c_1A^{n-1} + c_2A^{n-2} + ... + c_nI_n = \mathbf{0}, \tag{1.2}$$

where $\mathbf{0}$ is null matrix of order $n \times n$.

- **Eigen Value:** A root of the characteristic equation of a square matrix A is said to be an eigen value (or a characteristic value) of A.

- **Eigen Vector:** A non-null vector X is said to be an eigen vector (or a characteristic vector) of a square matrix A if there exists a scalar λ such that $AX = \lambda X$ holds.

1.9 Applications of the Cayley-Hamilton Theorem

Some of the most important applications of the Cayley-Hamilton theorem are discussed below:

- **Computing Matrix Inverses:** For a non-singular square matrix A, the Cayley-Hamilton Theorem provides an elegant way to obtain its inverse without directly resorting to Gaussian elimination or adjoint methods. By expressing the characteristic polynomial of A as

  $$P_A(\lambda) = det(A - \lambda I_n) = \lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + ... + c_n.,$$

  and substituting A into the polynomial, we obtain

  $$P_A(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I = \mathbf{0}.$$

  Rearranging this expression allows one to write A^{-1} as a linear combination of lower powers of A and the identity matrix, provided that A is invertible. This approach is especially useful in theoretical proofs and symbolic computations.

- **Evaluating Powers of a Matrix:** We can find A^n for a given square matrix A, where n is a natural number.

  Let A be an $n \times n$ square matrix with characteristic polynomial

  $$P_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0.$$

  According to the Cayley–Hamilton theorem, substituting A into its own characteristic polynomial yields

  $$P_A(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I = 0.$$

  From this relation, we can express the highest power A^n as a linear combination of lower powers of A:

  $$A^n = -\big(c_{n-1}A^{n-1} + c_{n-2}A^{n-2} + \cdots + c_1A + c_0I\big).$$

  Thus, any power A^m for $m > n$ can be reduced recursively to a linear combination of $\{I, A, A^2, \ldots, A^{n-1}\}$. This makes the computation of large powers of A efficient, avoiding repeated multiplication.

- **Evaluating Matrix Functions:** It facilitates the computation of functions of matrices, such as matrix exponentials and logarithms, by reducing higher powers of the matrix to linear combinations of lower-order powers.

  For a square matrix A, many functions such as the exponential e^A, logarithm $\log(A)$, or trigonometric functions like $\sin(A)$ and $\cos(A)$ are defined through their power series expansions. For example, the matrix exponential is defined as

  $$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots .$$

  Direct computation of such infinite series is impractical. However, by the Cayley–Hamilton theorem, any power of A greater than or equal to n (where n is the order of the matrix) can be expressed as a linear combination of $\{I, A, A^2, \ldots, A^{n-1}\}$. Hence, the infinite series reduces to a finite linear combination of these lower-order matrices.

1.10 Solved Problems

Example 13. *Use Cayley-Hamilton theorem to find A^{-1}, where $A = \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix}$*

Solution: The characteristic polynomial of A is $det(A - \lambda I_2)$,

$$i.e., \quad \begin{vmatrix} 2-\lambda & 1 \\ 3 & 5-\lambda \end{vmatrix} = (2-\lambda)(5-\lambda) - 3 = \lambda^2 - 7\lambda + 7.$$

Thus, the characteristic equation of A is $\lambda^2 - 7\lambda + 7 = 0$.
By the Cayley-Hamilton theorem, we have

$$\begin{aligned} & A^2 - 7A + 7I_2 = \mathbf{0} \\ \text{or,} \quad & A(A - 7I_2) = -7I_2 \\ \text{or,} \quad & A^{-1} = -\frac{1}{7}(A - 7I_2) = \frac{1}{7}\begin{pmatrix} 5 & -1 \\ -3 & 2 \end{pmatrix}. \end{aligned}$$

Therefore, $\boxed{A^{-1} = \frac{1}{7}\begin{pmatrix} 5 & -1 \\ -3 & 2 \end{pmatrix}.}$

Example 14. *Use Cayley-Hamilton theorem to find A^{50}, where $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.*

Solution: The characteristic polynomial of A is $det(A - \lambda I_2)$,

$$i.e., \quad \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 1.$$

Therefore, the characteristic equation of A is $\lambda^2 - 2\lambda + 1 = 0$. By the Cayley-Hamilton theorem, we have

$$\begin{aligned} & A^2 - 2A + I_2 = \mathbf{0} \\ \text{or,} \quad & A^2 - A = A - I_2 \\ \text{or,} \quad & A^3 - A^2 = A^2 - A = A - I_2 \\ & \cdots\cdots\cdots\cdots\cdots\cdots \\ & \cdots\cdots\cdots\cdots\cdots\cdots \\ \text{or,} \quad & A^{50} - A^{49} = A - I_2 \end{aligned}$$

Adding, we have

$$A^{50} = 50A - 49I_2 = 50\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} - 49\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 50 \\ 0 & 1 \end{pmatrix}.$$

Therefore, $\boxed{A = \begin{pmatrix} 1 & 50 \\ 0 & 1 \end{pmatrix}}$.

Example 15. *Show that the matrix* $A = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix}$ *satisfies its own characteristic equation.*

Solution: The characteristic polynomial of A is $det(A - \lambda I_3)$,

$$i.e., \quad \begin{vmatrix} 0-\lambda & 0 & 1 \\ 3 & 1-\lambda & 0 \\ -2 & 1 & 4-\lambda \end{vmatrix} = -\lambda^3 + 5\lambda^2 - 6\lambda + 5.$$

Thus, the characteristic equation of A is $\lambda^3 - 5\lambda^2 + 6\lambda - 5 = 0$.

We have to show that the matrix A satisfies its characteristic equation, i.e., $A^3 - 5A^2 + 6A - 5I = \mathbf{0}$.

Now,

$$A^2 = A.A = \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 4 \\ 3 & 1 & 3 \\ -5 & 5 & 14 \end{pmatrix}$$

$$A^3 = A^2.A = \begin{pmatrix} -2 & 1 & 4 \\ 3 & 1 & 3 \\ -5 & 5 & 14 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} = \begin{pmatrix} -5 & 5 & 14 \\ -3 & 4 & 15 \\ -13 & 19 & 51 \end{pmatrix}$$

So,

$$\begin{aligned} A^3 - 5A^2 + 6A - 5I &= \begin{pmatrix} -5 & 5 & 14 \\ -3 & 4 & 15 \\ -13 & 19 & 51 \end{pmatrix} - 5\begin{pmatrix} -2 & 1 & 4 \\ 3 & 1 & 3 \\ -5 & 5 & 14 \end{pmatrix} + 6\begin{pmatrix} 0 & 0 & 1 \\ 3 & 1 & 0 \\ -2 & 1 & 4 \end{pmatrix} \qquad (1.3) \\ &- 5\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -5 & 5 & 14 \\ -3 & 4 & 15 \\ -13 & 19 & 51 \end{pmatrix} - \begin{pmatrix} -10 & 5 & 20 \\ 15 & 5 & 15 \\ -25 & 25 & 70 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 6 \\ 18 & 6 & 0 \\ -12 & 6 & 24 \end{pmatrix} \qquad (1.4) \\ &- \begin{pmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Therefore, we have $\boxed{A^3 - 5A^2 + 6A - 5I = \mathbf{0}}$., i.e. the matrix A satisfies its characteristic equation.

Example 16. *If* $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, *then verify that A satisfies Cayley-Hamilton theorem . Hence find A^{-1} and A^9.*

Solution: The characteristic polynomial of A is $det(A - \lambda I_3)$,

$$i.e., \begin{vmatrix} 1-\lambda & 0 & 2 \\ 0 & -1-\lambda & 1 \\ 0 & 1 & -\lambda \end{vmatrix} = \lambda^3 - 2\lambda + 1.$$

Thus, the characteristic equation of A is $\lambda^3 - 2\lambda + 1 = 0$.
We have to show the matrix A satisfies its characteristic equation, i.e., $A^3 - 2A + I = \mathbf{0}$.
Now,

$$A^2 = A.A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

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

Thus,

$$\begin{aligned} A^3 - 2A + I &= \begin{pmatrix} 1 & 0 & 4 \\ 0 & -3 & 2 \\ 0 & 2 & -1 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 4 \\ 0 & -3 & 2 \\ 0 & 2 & -1 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 4 \\ 0 & -2 & 2 \\ 0 & 2 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Therefore, A satisfies the Caley-Hamilton theorem.
Now, from $A^3 - 2A + I = \mathbf{0}$, we have

$$\begin{aligned} & A^3 - 2A = -I \\ \text{or,} \quad & 2A - A^3 = I \\ \text{or,} \quad & A(2I - A^2) = I \end{aligned}$$

Thus, we have

$$A^{-1} = 2I - A^2 = 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -2 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Again, from $A^3 - 2A + I = \mathbf{0}$, we have

$$\begin{aligned}
A^3 &= 2A - I \\
\text{or, } A^9 &= (2A - I)^3 = 8A^3 - 12A^2 + 6A - I \\
&= 8(2A - I) - 12A^2 + 6A - I \\
&= 22A - 9I - 12A^2 \\
&= 22\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} - 9\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 12\begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -24 & 20 \\ 0 & -55 & 34 \\ 0 & 34 & -21 \end{pmatrix} \quad (1.5)
\end{aligned}$$

Therefore, $\boxed{A^9 = \begin{pmatrix} 1 & -24 & 20 \\ 0 & -55 & 34 \\ 0 & 34 & -21 \end{pmatrix}}$.

Example 17. *Let* $A = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$. *Use the Cayley–Hamilton theorem to compute* e^A.

Solution: The characteristic polynomial of A is

$$P_A(\lambda) = (\lambda - 2)(\lambda - 3) = \lambda^2 - 5\lambda + 6.$$

Therefore, the eigen values of A are $\lambda = 2, 3$ (Obtained from $P_A(\lambda) = 0$).

By Cayley-Hamilton,

$$A^2 - 5A + 6I = \mathbf{0} \quad \Longrightarrow \quad A^2 = 5A - 6I.$$

Hence every power A^m (for $m \geq 2$) can be reduced to a linear combination of I and A. Therefore, the matrix exponential may be written as

$$e^A = \alpha_0 I + \alpha_1 A$$

for some scalars α_0, α_1. To determine α_0, α_1 we use interpolation on the eigenvalues $\lambda = 2, 3$. For each eigenvalue λ,

$$e^\lambda = \alpha_0 + \alpha_1 \lambda.$$

Thus,

$$\begin{cases} e^2 = \alpha_0 + 2\alpha_1, \\ e^3 = \alpha_0 + 3\alpha_1. \end{cases}$$

Subtracting yields $\alpha_1 = e^3 - e^2$, and then

$$\alpha_0 = e^2 - 2\alpha_1 = 3e^2 - 2e^3.$$

Therefore,

$$\boxed{e^A = \left(3e^2 - 2e^3\right)I + \left(e^3 - e^2\right)A}.$$

Example 18. *Let* $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. *Compute* $\log(I + A)$.

Solution: Note that $A^2 = O$ (so A is nilpotent of index 2). The principal logarithm is given by the power series

$$\log(I + A) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{A^k}{k}.$$

Because $N^2 = O$, all terms with $k \geq 2$ vanish, leaving only the $k = 1$ term. Therefore,

$$\boxed{\log(I + N) = N}.$$

Example 19. *Let* $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. *Use the Cayley–Hamilton theorem to find* $\cos(A)$ *and* $\sin(A)$.

Solution. The characteristic polynomial of A is

$$P_A(\lambda) = \lambda^2 + 1,$$

So, the Cayley–Hamilton gives

$$A^2 + I = \mathbf{0} \quad \Longrightarrow \quad A^2 = -I.$$

Therefore, we have $A^{2k} = (-1)^k I$ and $A^{2k+1} = (-1)^k A$. Using the Taylor series definitions

$$\cos(A) = \sum_{k=0}^{\infty} (-1)^k \frac{A^{2k}}{(2k)!}, \qquad \sin(A) = \sum_{k=0}^{\infty} (-1)^k \frac{A^{2k+1}}{(2k+1)!},$$

and substituting the cycling pattern yields

$$\cos(J) = \sum_{k=0}^{\infty} \frac{I}{(2k)!} = \cosh(1)\, I, \qquad \sin(J) = \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} A = \sinh(1)\, A.$$

Thus,

$$\boxed{\cos(J) = \cosh(1)\, I, \qquad \sin(J) = \sinh(1)\, J}.$$

.

Exercises

(i) Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 4 \end{pmatrix}$. Use the Cayley–Hamilton theorem to express A^{100}.

(ii) For $A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$, compute e^A by using the the Cayley–Hamilton theorem.
Hint: Use eigenvalues to set up interpolation equations for the coefficients.

(iii) Let $A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{pmatrix}$. Show that A satisfies its characteristic equation. Also find A^4.

(iv) Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Use the Cayley–Hamilton theorem to compute A^{-1}.

(v) Let $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. (i) Find the characteristic polynomial of A. (ii) Use Cayley-Hamilton to compute A^{-1}.

(vi) Consider $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 11 & -6 \end{pmatrix}$. Show that A satisfies its characteristic equation, and hence determine A^{-1}.

(vii) Let A be a 3×3 matrix with characteristic polynomial

$$P_A(\lambda) = \lambda^3 - 5\lambda^2 + 8\lambda - 4.$$

Assuming $\det(A) \neq 0$, derive a formula for A^{-1} in terms of I, E, E^2 using the Cayley–Hamilton theorem.

(viii) Let $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & 11 & -6 \end{pmatrix}$. Show that A satisfies its characteristic equation. Also find A^{-1}.

1.11 Eigenvalue, Eigenvector, and Diagonalization of Matrices

Let $A = [a_{ij}]_{n\times n}$ be an n-dimensional square matrix and X be a $1 \times n$ non-null column vector. Then the non-zero column vector X is called an eigenvector of A if the linear transformation $Y = AX$ only stretches or shrinks X, in the same or opposite direction. Mathematically:

$$AX = \lambda X \tag{1.6}$$

Where λ is a scalar.
The scalar λ is called the eigenvalue corresponding to the eigenvector X. It tells us how much the eigenvector is stretched or compressed. If $\lambda > 0$, then X maintains the same direction. In case $\lambda < 0$, X is transformed in its opposite direction.
However, if λ is a complex number, the eigenvector is rotated and stretched. Thus, if $\lambda = a + ib$, then the magnitude $|\lambda| = \sqrt{a^2 + b^2}$ tells us how much the vector is stretched, in addition to the angle $\theta = \arg(\lambda) = \tan^{-1}(b/a)$ representing how much it rotates.

1.11.1 Applications in various fields

(i) They help reduce high-dimensional problems into more manageable findings regarding the "essence" of a system's behavior.

(ii) In linear algebra, they reveal how matrices change space. An eigenvector shows a direction that remains unchanged by a transformation, and its eigenvalue shows how much it stretches or shrinks.

(iii) In physics and engineering, they're crucial for studying vibrations, system stability, and structural behavior.

(iv) In data science and machine learning, eigenvectors help extract the most important features of complex datasets to reduce dimensionality without losing essential information.

(v) In quantum physics, they represent measurable quantities, such as energy levels, and help solve Schrödinger's equation.

(vi) Used to rank nodes on a network (like in Google's PageRank algorithm) or detect communities.

An Intuitive Visual

Imagine stretching or rotating a sheet of rubber (Fig. 1.1). Most points move unpredictably, but some special directions stretch perfectly in line. Those are your eigenvectors, and how far they stretch gives you the eigenvalues.

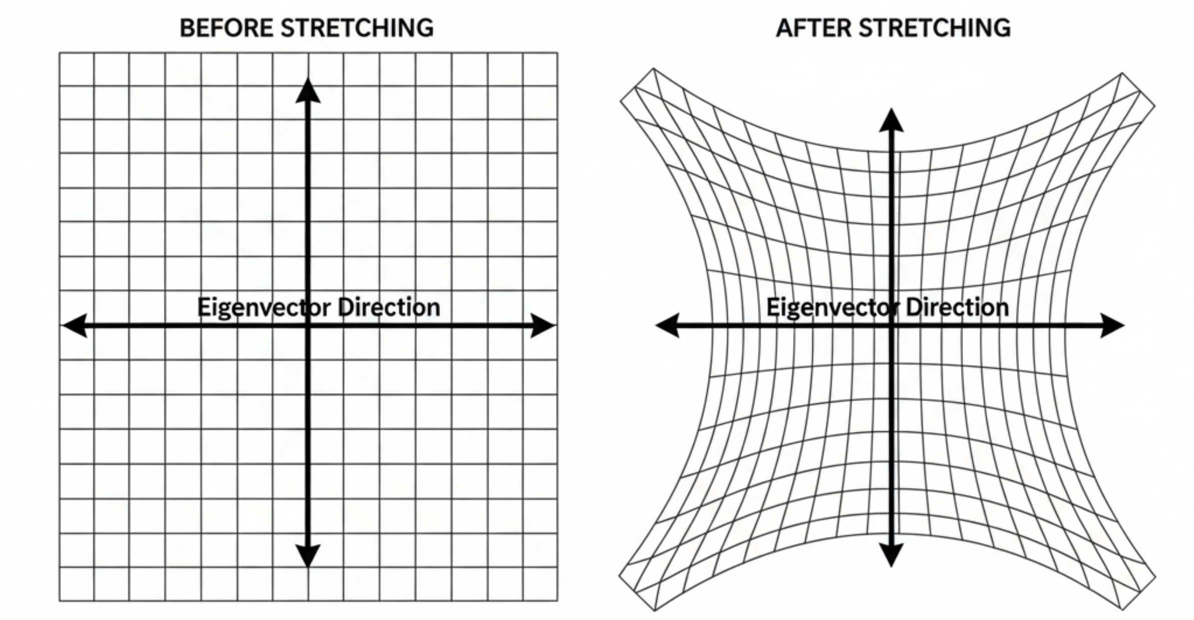


Figure 1.1: Visualizing the eigenvectors

In continuation of equation (1.6), we have, for $X \neq 0$,

$$AX = \lambda X \tag{1.7}$$

$$\Rightarrow (A - \lambda I)X = O \tag{1.8}$$

The equation (1.8) represents n homogeneous equations in n unknowns, which have non-trivial solutions if $(A - \lambda I)$ is a singular matrix, that is,

$$|A - \lambda I| = 0 \tag{1.9}$$

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \tag{1.10}$$

The LHS of equation (1.10), say $P(\lambda)$, is called the characteristic polynomial, and equation (1.10), that is, $P(\lambda) = 0$ itself, is called the characteristic equation. The roots $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ of this equation are called characteristic roots or eigenvalues. There is at least one and at most n distinct roots of the characteristic equation.

Properties of eigenvalues and eigenvectors

(a) The **Spectrum** of A is the set of all eigenvalues of A.

(b) If A is real, its eigenvalues are real or complex conjugates in pairs.

(c) From the characteristic equation

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0$$

we get an n^{th} degree polynomial in λ

$$\begin{aligned} |A - \lambda I| = (-1)^n \Big[&\lambda^n - \sigma_1 \lambda^{n-1} + \sigma_2 \lambda^{n-2} \dots \\ &+ (-1)^{n-1} \sigma_{n-1} \lambda + (-1)^n \sigma_n \Big] = 0 \end{aligned} \tag{1.11}$$

Here, **Trace:** σ_1 = sum of the diagonal elements of A
= trace of A
= sum of the roots of the polynomial equation
$= \lambda_1 + \lambda_2 + \cdots + \lambda_n$

Thus, trace(A) = sum of the eigen values of A.

$$\text{Determinant of } A = \sigma_n = |A| = \lambda_1 \cdot \lambda_2 \cdot \lambda_3 \cdots \lambda_n$$

Thus, the determinant of A = product of the eigenvalues of A.

(d) A and A^T have the same eigenvalues.

(e) If at least one eigenvalue is non-zero, then $|A| \neq 0$. Otherwise, A is singular, i.e $|A|$ is 0.

(f) If A is non-singular and $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ are the eigenvalues of A, then the eigenvalues of A^{-1} are $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{3}, \dots, \frac{1}{\lambda_n}$.

(g) "Spectral shift": The matrix $A \mp kI$ has eigenvalues $\lambda_i \mp k$ and has the same eigenvectors as A.

(h) If $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$ are the eigenvalues of A, then the eigenvalues of kA are $k\lambda_1, k\lambda_2, k\lambda_3, \dots, k$ k being a scalar.

(i) If λ is an eigenvalue of A, then λ^m is an eigenvalue of A^m, $m \in \mathbb{Z}$.

(j) An eigenvector cannot correspond to two distinct eigenvalues.

(k) The eigenvalues of a diagonal/upper triangular/lower triangular matrix are the diagonal elements of that matrix. If λ is an eigenvalue of an orthogonal matrix A, then $\frac{1}{\lambda}$ is also an eigenvalue of A. [A matrix A is said to be an orthogonal matrix if $A^T = A^{-1}$.]

1.11.2 Worked Examples on Eigenvalues and Eigenvectors

Prob. 1 Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 8 & -4 \\ 2 & 2 \end{pmatrix}$

Solution: The eigenvalues are the roots of the characteristic equation

$$\begin{vmatrix} 8-\lambda & -4 \\ 2 & 2-\lambda \end{vmatrix} = 0$$

i.e.,

$$(8-\lambda)(2-\lambda) + 8 = 0$$

or

$$\lambda^2 - 10\lambda + 24 = 0,$$
$$(\lambda - 4)(\lambda - 6) = 0$$

The two distinct eigenvalues are $\lambda = 4, 6$.
Eigen vector corresponding to eigen value $\lambda = 4$:
$(A - \lambda I)X = 0$

$$\begin{pmatrix} 8-4 & -4 \\ 2 & 2-4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

$$4x_1 - 4x_2 = 0$$
$$2x_1 - 2x_2 = 0 \quad \therefore x_1 = x_2$$

$$\bar{X}_1 = C_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$\bar{X}_2$ corresponding to $\lambda = 6$

$$\begin{pmatrix} 8-6 & -4 \\ 2 & 2-6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ 2 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$$

$$2x_1 - 4x_2 = 0 \quad \therefore x_1 = 2x_2$$

$$\bar{X}_2 = C_2 \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

Prob. 2 Find the sum and product of eigenvalues of A of **Prob. 1**.

Solution: Sum of the eigenvalues of $A = 4 + 6 = 10 = \text{trace}(A) = a_{11} + a_{22} = 8 + 2 = 10$.

Product of the eigenvalues of $A = 4.6 = 24 = |A| = 16 + 8 = 24$.

Prob. 3 Find the eigenvalues and eigen vectors of

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

Solution: For upper triangular, lower triangular, and diagonal matrices, the eigenvalues are given by the diagonal elements. The characteristic equation

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

i.e.,

$$(3-\lambda)(2-\lambda)(5-\lambda) = 0$$

So eigenvalues of A are $3, 2, 5$, which are the diagonal elements of A.

Eigen vector X_1 for $\lambda = 3$

$$\begin{bmatrix} 3-3 & 1 & 4 \\ 0 & 2-3 & 6 \\ 0 & 0 & 5-3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

i.e., $x_2 + 4x_3 = 0, \quad -x_2 + 6x_3 = 0, \quad 2x_3 = 0$
$x_2 = 0, x_3 = 0, x_1 =$ arbitrary.

$$X_1 = C_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Eigen vector X_2 for $\lambda = 2$

$$\begin{bmatrix} 3-2 & 1 & 4 \\ 0 & 2-2 & 6 \\ 0 & 0 & 5-2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

i.e., $x_1 + x_2 + 4x_3 = 0, \quad 6x_3 = 0, \quad 3x_3 = 0$

$\therefore x_3 = 0, x_1 = -x_2, \quad X_2 = C_1 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$

For $\lambda = 5$,

$$\begin{bmatrix} 3-5 & 1 & 4 \\ 0 & 2-5 & 6 \\ 0 & 0 & 5-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

i.e., $-2x_1 + x_2 + 4x_3 = 0, \quad -3x_2 + 6x_3 = 0$
$x_1 = 3x_3, x_2 = 2x_3$

$$X_3 = C_2 \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Prob. 4 $A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$. Determine the algebraic and geometric multiplicity.

Solution: The characteristic equation is

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

$$\lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2 = 0$$

So $\lambda = 1, 2, 2$ are eigenvalues with $\lambda = 2$ repeated twice (double root) of multiplicity 2. The algebraic multiplicity of the eigenvalue $\lambda = 2$ is 2.

For $\lambda = 1$,

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

The system of equations is:

$$\begin{aligned} 2x_2 + 2x_3 &= 0 \\ x_2 + x_3 &= 0 \\ -x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

From the first two equations, $x_2 = -x_3$. Substituting into the third gives $x_1 = -x_3$.

$$X_1 = C \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

For $\lambda = 2$,

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

The system of equations is:

$$\begin{aligned} -x_1 + 2x_2 + 2x_3 &= 0 \\ x_3 &= 0 \\ -x_1 + 2x_2 &= 0 \end{aligned}$$

From the second equation, $x_3 = 0$. From the third, $x_1 = 2x_2$.

$$X_2 = C \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Thus, only one eigenvector X_2 corresponds to the repeated eigenvalue $\lambda = 2$. The geometric multiplicity of the eigenvalue $\lambda = 2$ is one.

Prob. 5 If $A = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 2 & 1 \\ -1 & 2 & 2 \end{bmatrix}$, then determine the algebraic and geometric multiplicity.

Solution: The characteristic equation is

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

$$\lambda^3 - 5\lambda^2 + 8\lambda - 4 = (\lambda - 1)(\lambda - 2)^2 = 0$$

So $\lambda = 1, 2, 2$ are eigenvalues with $\lambda = 2$ repeated twice (double root) of multiplicity 2. The algebraic multiplicity of the eigenvalue $\lambda = 2$ is 2.

For $\lambda = 1$,

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

The system of equations is:

$$\begin{aligned} 2x_2 + 2x_3 &= 0 \\ x_2 + x_3 &= 0 \\ -x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

From the first two equations, $x_2 = -x_3$. Substituting into the third gives $x_1 = -x_3$.

$$X_1 = C \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

For $\lambda = 2$,

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

The system of equations is:

$$\begin{aligned} -x_1 + 2x_2 + 2x_3 &= 0 \\ x_3 &= 0 \\ -x_1 + 2x_2 &= 0 \end{aligned}$$

From the second equation, $x_3 = 0$. From the third, $x_1 = 2x_2$.

$$X_2 = C \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Thus, only one eigenvector X_2 corresponds to the repeated eigenvalue $\lambda = 2$. The geometric multiplicity of the eigenvalue $\lambda = 2$ is one.

Prob. 6 $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Determine the algebraic and geometric multiplicity.

Solution: Characteristic equation is

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)((2-\lambda)^2-1) = (1-\lambda)(1-\lambda)(3-\lambda) = 0$$

Thus, $\lambda = 1, 1, 3$ are the eigenvalues. So the algebraic multiplicity of the value $\lambda = 1$ is two. **For** $\lambda = 3$,

$$\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \implies x_3 = 0, x_1 = x_2, \quad X_1 = C \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

For $\lambda = 1$,

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

Here $n = 3$, $r = 1$, so $n - r = 3 - 1 = 2$ arbitrary variables. We have $x_1 + x_2 + x_3 = 0$ or $x_1 = -x_2 - x_3$ where x_2 and x_3 are arbitrary.

For a choice of $x_2 = 0, x_3 = 1$,

$$X_2 = C \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

For a choice of $x_2 = 1, x_3 = 0$,

$$X_3 = C \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Thus, for the repeated eigenvalue $\lambda = 1$, there correspond two linearly independent eigen vectors X_2 and X_3. So the geometric multiplicity of the eigenvalue $\lambda = 1$ is 2.

Prob. 7 Given that $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}$

Determine the algebraic and geometric multiplicity.

Solution: Characteristic equation is

$$\lambda^3 - 3\lambda^2 + 3\lambda - 1 = (\lambda - 1)^3 = 0$$

Thus $\lambda = 1, 1, 1$ is an eigenvalue of algebraic multiplicity 3. **For** $\lambda = 1$, The system of equations is:

$$\begin{aligned} -x_1 + x_2 &= 0 \qquad \therefore x_1 = x_2 \\ -x_2 + x_3 &= 0 \qquad \therefore x_2 = x_3 \\ x_1 - 3x_2 + 2x_3 &= 0 \end{aligned}$$

(The third equation is satisfied if $x_1 = x_2 = x_3$)

$$X = C \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Thus, only one eigenvector X corresponds to the thrice-repeated eigenvalue $\lambda = 1$, so the geometric multiplicity is one.

Prob. 8 Find the inverse transformation of

$$\begin{aligned} y_1 &= x_1 + 2x_2 + 5x_3 \\ y_2 &= -x_2 + 2x_3 \\ y_3 &= 2x_1 + 4x_2 + 11x_3 \end{aligned}$$

Solution: With $Y = \begin{bmatrix} y_1 & y_2 & y_3 \end{bmatrix}^T$, $X = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$, the coefficient matrix

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & -1 & 2 \\ 2 & 4 & 11 \end{bmatrix}. \quad \text{Its } |A| = -1$$

$$\text{Adj } A = \begin{bmatrix} -19 & -2 & 9 \\ 4 & 1 & -2 \\ 2 & 0 & -1 \end{bmatrix}$$

Thus, the inverse transformation is

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A^{-1}Y = \begin{bmatrix} 19 & 2 & -9 \\ -4 & -1 & 2 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

$$= \begin{bmatrix} 19y_1 & +2y_2 & -9y_3 \\ -4y_1 & -y_2 & +2y_3 \\ -2y_1 & & +y_3 \end{bmatrix}$$

1.11.3 Caley-Hamilton Theorem

Theorem 2. *Caley-Hamilton Theorem: Every square matrix satisfies its own characteristic equation.*

Prob. 1 Verify Cayley-Hamilton theorem for the matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$. Find A^{-1}. Determine A^8.

Solution: The characteristic equation is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 \\ 2 & -1-\lambda \end{vmatrix} = 0$$

or

$$\begin{aligned} (\lambda - 1)(1 + \lambda) - 4 &= 0 \\ \lambda^2 - 5 &= 0 \end{aligned}$$

so $A^2 - 5I = 0$

$$A^2 = A \cdot A = \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = 5I$$
$$\text{or } A^2 - 5I = 0$$

Thus A satisfies the characteristic equation. To find A^{-1}, multiply $A^2 - 5I = 0$ by A^{-1}.

$$\begin{aligned} A^{-1}A^2 - 5A^{-1}I &= 0 \\ A - 5A^{-1} &= 0 \\ \text{So } A^{-1} &= \frac{1}{5}A = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix} \end{aligned}$$

To find A^8, multiply $A^2 - 5I = 0$ by A^6.

$$\begin{aligned} A^6A^2 - 5IA^6 &= 0 \\ A^8 &= 5A^6 = 5 \cdot A^2 \cdot A^2 \cdot A^2 = 5 \cdot (5I)(5I)(5I) \\ A^8 &= 625I \end{aligned}$$

Prob. 2 Verify the Cayley-Hamilton theorem for the matrix

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

and hence find the inverse of A. Find A^4. Express $B = A^8 - 11A^7 - 4A^6 + A^5 + A^4 - 11A^3 - 3A^2 + 2A + I$ as a quadratic polynomial in A. Find B.

Solution: The characteristic equation of A is

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 2 & 4-\lambda & 5 \\ 3 & 5 & 6-\lambda \end{vmatrix} = (1-\lambda)[(4-\lambda)(6-\lambda) - 25]$$

$$-2[2(6-\lambda) - 15] + 3[10 - 3(4-\lambda)] = 0$$

$$\lambda^3 - 11\lambda^2 - 4\lambda + 1 = 0$$

Cayley-Hamilton theorem is verified if A satisfies the above characteristic equation, i.e.,

$$A^3 - 11A^2 - 4A + I = 0$$

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

$$= \begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix}$$

$$A^3 = A \cdot A^2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix}$$

$$= \begin{bmatrix} 157 & 283 & 353 \\ 283 & 510 & 636 \\ 353 & 636 & 793 \end{bmatrix}$$

verification

$$A^3 - 11A^2 - 4A + I$$

$$= \begin{bmatrix} 157 & 283 & 353 \\ 283 & 510 & 636 \\ 353 & 636 & 793 \end{bmatrix} - 11 \begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix} - 4 \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

To find A^{-1}: From characteristic equation $A^{-1} = -A^2 + 11A + 4I$. So

$$A^{-1} = -\begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix} + 11 \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} + 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -3 & -2 \\ -3 & 3 & 1 \\ -2 & 1 & -1 \end{bmatrix}$$

To find A^4: From Cayley-Hamilton theorem

$$A^3 - 11A^2 - 4A + I = 0$$

or

$$A^3 = 11A^2 + 4A - I$$

Multiplying both sides by A

$$A^4 = A \cdot A^3 = A(11A^2 + 4A - I) = 11A^3 + 4A^2 - A$$

$$= 11 \begin{bmatrix} 157 & 283 & 353 \\ 283 & 510 & 636 \\ 353 & 636 & 793 \end{bmatrix} + 4 \begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$$

$$= \begin{bmatrix} 1782 & 3211 & 4004 \\ 3211 & 5786 & 7215 \\ 4004 & 7215 & 8997 \end{bmatrix}$$

To find B: Rewrite

$$B = A^8 - 11A^7 - 4A^6 + A^5 + A^4 - 11A^3 - 3A^2 + 2A + I$$

$$= A^5(A^3 - 11A^2 - 4A + I) + A^2(A^3 - 11A^2 - 4A + I) + A^2 + A + I$$

since A satisfies the characteristic equation. Thus

$$B = A^2 + A + I$$

$$B = \begin{bmatrix} 14 & 25 & 31 \\ 25 & 45 & 56 \\ 31 & 56 & 70 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 16 & 27 & 34 \\ 27 & 50 & 61 \\ 34 & 61 & 77 \end{bmatrix}$$

1.12 Diagonalization of Matrices

Let A be a square matrix of order n.

Similar Matrix

A is said to be similar to B if there exists a non-singular matrix P such that $B = P^{-1}AP$. This transformation of A to B is known as a similarity transformation.

Invariant Eigen Values

Theorem 3. *Similar matrices A and B have the same eigenvalues.*

Proof: Suppose B is similar to A i.e., $B = P^{-1}AP$. Consider the characteristic polynomial of B:

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda I| \\ &= |P^{-1}AP - \lambda P^{-1}IP| \\ &= |P^{-1}(A - \lambda I)P| \\ &= |P^{-1}||A - \lambda I||P| \\ &= |A - \lambda I| \end{aligned}$$

since $|P^{-1}||P| = |P^{-1}P| = |I| = 1$. Thus, A and B have the same characteristic polynomial and therefore have the same eigenvalues.

Theorem 4. *If X is an eigenvector of A, then $Y = P^{-1}X$ is an eigenvector of the matrix B.*

Proof: Let X be an eigen vector of A so that $AX = \lambda X$. Consider $B = P^{-1}AP$. Post-multiplying by P^{-1}, we get

$$BP^{-1} = (P^{-1}AP)P^{-1} = (P^{-1}A)(PP^{-1}) = P^{-1}A$$

. Post-multiply by X:

$$\begin{aligned} B(P^{-1}X) &= P^{-1}AX = P^{-1}(AX) \\ &= P^{-1}\lambda X = \lambda(P^{-1}X) \end{aligned}$$

Thus, $P^{-1}X$ is an eigenvector of B corresponding to the eigenvalue λ.

Diagonalization

An n-square matrix A with n linearly independent eigenvectors is similar to a diagonal matrix D whose diagonal elements are the eigenvalues of A.

Proof

Let $X_1, X_2, \ldots, X_n$ be the n linearly independent eigen vectors of A corresponding to n eigen values $\lambda_1, \lambda_2, \ldots, \lambda_n$. Construct P, known as the modal matrix, having

$X_1, X_2, \ldots, X_n$ as the n column vectors, i.e., $P_{n\times n} = [X_1X_2 \ldots X_n]$. Since $X_1, X_2, \ldots, X_n$ are linearly independent, P^{-1} exists. Consider

$$\begin{aligned} AP &= A[X_1X_2 \ldots X_n] = [AX_1AX_2 \ldots AX_n] \\ &= [\lambda_1X_1\lambda_2X_2 \ldots \lambda_nX_n] \\ &= [X_1X_2 \ldots X_n] \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \ldots & 0 & \lambda_n \end{bmatrix} \end{aligned}$$

$AP = PD$ where D is the diagonal matrix with eigenvalues of A as the principal diagonal elements. D is known as the spectral matrix. Pre-multiplying by P^{-1} on both sides:

$$B = P^{-1}AP = P^{-1}PD = (P^{-1}P)D = ID = D$$

Powers of a Matrix A

Consider $D = P^{-1}AP$. Then

$$\begin{aligned} D^2 &= (P^{-1}AP)(P^{-1}AP) \\ &= P^{-1}A(PP^{-1})AP \\ &= P^{-1}A \cdot IAP = P^{-1}A^2P \end{aligned}$$

Similarly, $D^3 = P^{-1}A^3P$. Thus, $D^n = P^{-1}A^nP$.

To obtain A^n, pre-multiply by P and post-multiply by P^{-1},

$$\begin{aligned} PD^nP^{-1} &= P(P^{-1}A^nP)P^{-1} \\ &= (PP^{-1})A^n(PP^{-1}) \\ &= IA^nI = A^n \end{aligned}$$

$\therefore A^n = PD^nP^{-1}$.

Physical interpretation and Applications of Diagonalization

Diagonalization of a matrix is not just a mathematical procedure; it has a profound physical interpretation, representing a change of coordinate systems to one where a linear transformation or a physical system's behavior is as simple as possible.

The core physical interpretation is that diagonalization finds a special basis (the eigenvector basis) in which a complex linear transformation is equivalent to a simple scaling along the new basis axes.

Here is a breakdown of the physical interpretation with key examples:

The Fundamental Idea: Change of Basis

A matrix, say A, represents a linear transformation (e.g., rotation, stretching, shearing) in a standard coordinate system.

The process of diagonalization, $D = P^{-1}AP$, transforms matrix A into a new, simpler form D.

The columns of the matrix P are the eigenvectors of A. These eigenvectors form a new, special set of coordinate axes.

When we perform the similarity transformation, we are essentially looking at the original transformation A from the perspective of this new eigenvector basis.

In this new basis, the transformation is decoupled or uncoupled. The diagonal matrix D shows that each new basis vector (eigenvector) is simply stretched or shrunk by a factor given by the corresponding eigenvalue. There is no interaction, rotation, or shearing between these new basis axes.

Physical and Engineering Examples

(i) **Principal Component Analysis (Statistics and Data Science)** In data analysis, the covariance matrix of a dataset describes how different variables vary together.

- **Physical Interpretation:** Diagonalizing the covariance matrix identifies the principal components (the eigenvectors), which are new, uncorrelated variables. The eigenvalues represent the variance of the data along each of these principal components.
- **Result:** This process helps to reduce the dimensionality of data by identifying the directions of maximum variance. The largest eigenvalues correspond to the most significant features of the dataset.

(ii) **Rotational Dynamics (Physics and Engineering)** The moment of inertia tensor describes the rotational inertia of a rigid body. This tensor is a symmetric matrix.

- **Physical Interpretation:** Diagonalizing this matrix yields the principal moments of inertia (the eigenvalues) and the principal axes of rotation (the eigenvectors). These axes form the natural coordinate system for the rotating body.
- **Result:** In the principal axes frame, the rotational motion is simplified. If the body spins about a principal axis, it rotates smoothly without wobbling. The off-diagonal terms, which represent coupled motion, vanish.

(iii) **Vibrational Analysis (Mechanical and Civil Engineering)** In a system of coupled oscillators (like masses connected by springs), the motion can be described by a matrix.

- **Physical Interpretation:** Diagonalization reveals the natural frequencies (the square roots of the eigenvalues) and the normal modes of vibration (the eigenvectors). The normal modes represent the collective, synchronized motions of the masses, where they all oscillate at the same frequency.
- **Result:** A complex, coupled system can be understood as a superposition of these simple, independent normal modes. The diagonalization process "decouples" the system's differential equations, allowing for a much simpler analysis.

(iv) **Quantum Mechanics (Physics)** In quantum mechanics, a matrix operator represents a physical observable, such as energy, momentum, or spin.

- **Physical Interpretation:** The eigenvalues of the operator are the possible measurement outcomes for that observable. The eigenvectors are the corresponding quantum states in which the system is guaranteed to have that specific outcome.
- **Result:** Diagonalization finds the possible observable values for a quantum system, which is a fundamental concept in quantum theory.

1.12.1 Worked Examples on Diagonalization of Matrices

Prob. 1 Find a matrix P that diagonalizes the matrix

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

Verify that $P^{-1}AP = D$ where D is the diagonal matrix. Hence, find A^6.

Solution: A is diagonalizable by P whose columns are the linearly independent eigenvectors of A. The characteristic equation of A is

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

$$(4-\lambda)(3-\lambda) - 2 = \lambda^2 - 7\lambda + 10 = (\lambda - 2)(\lambda - 5) = 0,$$

so $\lambda = 2, 5$ are two distinct eigenvalues of A. For $\lambda = 2$, $2x_1 + x_2 = 0$, $x_2 = -2x_1$, $X_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ For $\lambda = 5$, $-x_1 + x_2 = 0$, $x_2 = x_1$, $X_2 = [1, 1]^T$

Thus, the matrix P, which diagonalizes A, is

$$P = [X_1, X_2] = \begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}$$

Verification:

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

$$\begin{aligned}
P^{-1}AP &= \frac{1}{3}\begin{pmatrix}1 & -1\\ 2 & 1\end{pmatrix}\begin{pmatrix}4 & 1\\ 2 & 3\end{pmatrix}\begin{pmatrix}1 & 1\\ -2 & 1\end{pmatrix}\\
&= \frac{1}{3}\begin{pmatrix}2 & -2\\ 10 & 5\end{pmatrix}\begin{pmatrix}1 & 1\\ -2 & 1\end{pmatrix}\\
&= \frac{1}{3}\begin{pmatrix}6 & 0\\ 0 & 15\end{pmatrix}\\
&= \begin{pmatrix}2 & 0\\ 0 & 5\end{pmatrix} = D = \text{diagonal matrix}
\end{aligned}$$

D contains eigenvalues 2, 5 as diagonal elements.

To find A^6:

$$A^6 = PD^6P^{-1} = \begin{pmatrix}1 & 1\\ -2 & 1\end{pmatrix}\begin{pmatrix}2^6 & 0\\ 0 & 5^6\end{pmatrix}\frac{1}{3}\begin{pmatrix}1 & -1\\ 2 & 1\end{pmatrix}$$

$$\begin{aligned}
A^6 &= \frac{1}{3}\begin{pmatrix}1 & 1\\ -2 & 1\end{pmatrix}\begin{pmatrix}64 & 0\\ 0 & 15625\end{pmatrix}\begin{pmatrix}1 & -1\\ 2 & 1\end{pmatrix}\\
&= \frac{1}{3}\begin{pmatrix}64 & 15625\\ -128 & 15625\end{pmatrix}\begin{pmatrix}1 & -1\\ 2 & 1\end{pmatrix}\\
&= \frac{1}{3}\begin{pmatrix}31314 & 15561\\ 31122 & 15753\end{pmatrix}\\
&= \begin{pmatrix}10438 & 5187\\ 10374 & 5251\end{pmatrix}
\end{aligned}$$

Prob. 2 Diagonalize

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

and hence find A^8. Find the modal matrix.

Solution: The non-singular square matrix P containing eigen vectors of A as columns, diagonalizes A. The characteristic equation of A is

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

i.e., $(\lambda+1)(\lambda-3)(\lambda-4) = 0$ so eigen values of A are $\lambda = -1, 3, 4$.

For $\lambda = -1$,

$$\begin{aligned}
2x_1 + 6x_2 + x_3 &= 0\\
x_1 + 3x_2 + 0 &= 0\\
4x_3 &= 0
\end{aligned}$$

Thus $x_3 = 0, x_1 = -3x_2$.

$$X_1 = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$$

For $\lambda = 3$,

$$\begin{aligned} -2x_1 + 6x_2 + x_3 &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$$

Thus $x_1 = x_2, x_3 = -4x_2$.

$$X_2 = \begin{bmatrix} 1 \\ 1 \\ -4 \end{bmatrix}$$

For $\lambda = 4$,

$$\begin{aligned} -3x_1 + 6x_2 + x_3 &= 0 \\ x_1 - 2x_2 &= 0 \\ -x_3 &= 0 \end{aligned}$$

Thus $x_3 = 0, x_1 = 2x_2$.

$$X_3 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Thus

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

is the modal matrix.

To find P^{-1}:

$$\left[\begin{array}{ccc|ccc} -3 & 1 & 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & -4 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ -3 & 1 & 2 & 1 & 0 & 0 \\ 0 & -4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 \to R_2 + 3R_1} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 4 & 5 & 1 & 3 & 0 \\ 0 & -4 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 4 & 5 & 1 & 3 & 0 \\ 0 & 0 & 5 & 1 & 3 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 \to \frac{1}{5}R_3]{R_2 \to \frac{1}{4}R_2} \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 5/4 & 1/4 & 3/4 & 0 \\ 0 & 0 & 1 & 1/5 & 3/5 & 1/5 \end{array}\right]$$

$$\xrightarrow[R_1 \to R_1 - R_3]{R_2 \to R_2 - \frac{5}{4}R_3} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & -1/5 & 2/5 & -1/5 \\ 0 & 1 & 0 & 0 & 0 & -1/4 \\ 0 & 0 & 1 & 1/5 & 3/5 & 1/5 \end{array}\right]$$

$$\xrightarrow{R_1 \to R_1 - R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/5 & 2/5 & 1/20 \\ 0 & 1 & 0 & 0 & 0 & -1/4 \\ 0 & 0 & 1 & 1/5 & 3/5 & 1/5 \end{array}\right]$$

Thus

$$P^{-1} = \frac{1}{20}\begin{bmatrix} -4 & 8 & 1 \\ 0 & 0 & -5 \\ 4 & 12 & 4 \end{bmatrix}$$

Diagonalization:

$$\begin{aligned} D &= P^{-1}AP \\ &= \frac{1}{20}\begin{bmatrix} -4 & 8 & 1 \\ 0 & 0 & -5 \\ 4 & 12 & 4 \end{bmatrix}\begin{bmatrix} 1 & 6 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} -3 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & -4 & 0 \end{bmatrix} \\ &= \frac{1}{20}\begin{bmatrix} 4 & -8 & -1 \\ 0 & 0 & -15 \\ 16 & 48 & 16 \end{bmatrix}\begin{bmatrix} -3 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & -4 & 0 \end{bmatrix} \\ &= \frac{1}{20}\begin{bmatrix} -20 & 0 & 0 \\ 0 & 60 & 0 \\ 0 & 0 & 80 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$

To find A^8:

$$A^8 = PD^8P^{-1}$$

$$A^8 = \begin{bmatrix} -3 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & -4 & 0 \end{bmatrix} \begin{bmatrix} (-1)^8 & 0 & 0 \\ 0 & 3^8 & 0 \\ 0 & 0 & 4^8 \end{bmatrix} \frac{1}{20} \begin{bmatrix} -4 & 8 & 1 \\ 0 & 0 & -5 \\ 4 & 12 & 4 \end{bmatrix}$$

$$= \frac{1}{20} \begin{bmatrix} -3 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & -4 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 6561 & 0 \\ 0 & 0 & 65536 \end{bmatrix} \begin{bmatrix} -4 & 8 & 1 \\ 0 & 0 & -5 \\ 4 & 12 & 4 \end{bmatrix}$$

$$= \frac{1}{20} \begin{bmatrix} -3 & 6561 & 131072 \\ 1 & 6561 & 65536 \\ 0 & -26244 & 0 \end{bmatrix} \begin{bmatrix} -4 & 8 & 1 \\ 0 & 0 & -5 \\ 4 & 12 & 4 \end{bmatrix}$$

$$= \frac{1}{20} \begin{bmatrix} 524300 & 1572840 & 491480 \\ 262140 & 786440 & 229340 \\ 0 & 0 & 131220 \end{bmatrix}$$

$$A^8 = \begin{bmatrix} 26215 & 78642 & 24574 \\ 13107 & 39322 & 11467 \\ 0 & 0 & 6561 \end{bmatrix}$$

1.13 Exercise: Eigenvalues & Eigenvectors

Find the eigenvalues and eigenvectors of:

(i) $\begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}$ **Ans:** $\lambda^2 + 7\lambda + 6 = 0, \lambda = -1, -6, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

(ii) $\begin{bmatrix} 6 & 8 \\ 8 & -6 \end{bmatrix}$ **Ans:** $10, -10, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$ **Ans:** $2, -1, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ **Ans:** $4, -1, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

(v) $\begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$ **Ans:** $5, -3, -3, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

(vi) $\begin{bmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{bmatrix}$ **Ans:** $\lambda^3 - 7\lambda^2 + 36 = 0, \lambda = -2, 3, 6, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

(vii) $\begin{bmatrix} -3 & -7 & -5 \\ 2 & 4 & 3 \\ 1 & 2 & 2 \end{bmatrix}$ **Ans:** $(\lambda - 1)^3 = 0, \lambda = 1, 1, 1, \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix}$

(viii) $\begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$ **Ans:** $\lambda^3 - 18\lambda^2 + 45\lambda = 0, \lambda = 0, 3, 15,$ $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$

(ix) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$ **Ans:** $\lambda^3 - 12\lambda^2 + 36\lambda - 32 = 0,\ \lambda = 2, 2, 8,$ $\begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, For $\lambda = 8,$ $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

(x) $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$ **Ans:** $(\lambda - 2)^3 = 0,\ \lambda = 2, 2, 2,$ $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

(xi) $\begin{bmatrix} 1 & -4 & -1 & -4 \\ 2 & 0 & 5 & -4 \\ -1 & 1 & -2 & 3 \\ -1 & 4 & -1 & 6 \end{bmatrix}$ **Ans:** $\lambda^4 - 5\lambda^3 + 9\lambda^2 - 7\lambda + 2 = 0,\ \lambda = 2, 1, 1, 1,$ For $\lambda = 2,$ $\begin{bmatrix} 2 \\ 3 \\ -2 \\ -3 \end{bmatrix}$, For $\lambda = 1,$ $\begin{bmatrix} 3 \\ 6 \\ -4 \\ -5 \end{bmatrix}$

(xii) $\begin{bmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{bmatrix}$ **Ans:** $\lambda^3 - 2\lambda^2 - 4\lambda + 8 = 0,\ \lambda = 2, 2, -2,$ For $\lambda = 2,$ $[0\ 1\ 1]^T$ For $\lambda = -2,$ $[-4\ \ -1\ 7]^T$

(xiii) $\begin{bmatrix} 3 & -2 & -5 \\ 4 & -1 & -5 \\ -2 & -1 & -3 \end{bmatrix}$ **Ans:** $(\lambda + 5)(\lambda - 2)^2 = 0,\ \lambda = 5, 2, 2,$ For $\lambda = 5,$ $X_1 = [3\ 2\ 4]^T$ For $\lambda = 2,$ $X_2 = [1\ 3\ \ -1]^T$

(xiv) Find the sum and product of the eigenvalues of $A = \begin{bmatrix} 2 & 3 & -2 \\ -2 & 1 & 1 \\ 1 & 0 & 2 \end{bmatrix}$

Ans: sum $= trace(A) = 2 + 1 + 2 = 5$, product $= |A| = 21$.

(xv) $\begin{pmatrix} 2 & 5 \\ 1 & -3 \end{pmatrix}$ **Ans:** Characteristic polynomial: $\lambda^2 + \lambda - 11$

(xvi) $\begin{pmatrix} 2 & -3 \\ 7 & -4 \end{pmatrix}$ **Ans:** Characteristic polynomial: $\lambda^2 + 2\lambda + 13$

(xvii) $\begin{bmatrix} 1 & 4 & -3 \\ 0 & 3 & 1 \\ 0 & 2 & -1 \end{bmatrix}$ **Ans:** Characteristic equation: $\lambda^3 - 3\lambda^2 - 3\lambda + 5 = 0$

1.14 Exercise: Eigenvalues & eigenvectors

(i) Verify Cayley-Hamilton theorem for $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$. Find A^{-1}. Find $B = A^5 - 4A^4 - 7A^3 + 11A^2 - A - 10I$ **Ans:** Characteristic equation: $\lambda^2 - 4\lambda - 5 = 0$, $A^{-1} = \frac{1}{5}\begin{pmatrix} -3 & 4 \\ 2 & -1 \end{pmatrix}$, $B = A + 5I = \begin{pmatrix} 6 & 4 \\ 2 & 8 \end{pmatrix}$

(ii) Use Cayley-Hamilton theorem to find A^{-1} if $A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & -3 \\ -2 & -4 & -4 \end{bmatrix}$ **Ans:** Characteristic equation: $\lambda^3 - 20\lambda + 8 = 0$, $A^{-1} = \frac{1}{4}\begin{bmatrix} 12 & 4 & 6 \\ -5 & -1 & -3 \\ -1 & -1 & -1 \end{bmatrix}$

(iii) Find A^{-1} and A^4 if $A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$ **Ans:** Characteristic equation: $\lambda^3 - 3\lambda^2 - \lambda + 9 = 0$, $A^{-1} = \frac{1}{9}\begin{bmatrix} 0 & 3 & 3 \\ 3 & 2 & -7 \\ 3 & -1 & -1 \end{bmatrix}$, $A^4 = \begin{bmatrix} 7 & -30 & 42 \\ 18 & -13 & 46 \\ -6 & -14 & 17 \end{bmatrix}$

(iv) Find A^{-1} for $A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$ **Ans:** Characteristic equation: $\lambda^3 - 6\lambda^2 + 9\lambda - 4 = 0$, $A^{-1} = \frac{1}{4}\begin{bmatrix} 3 & 1 & -1 \\ 1 & 3 & 1 \\ -1 & 1 & 3 \end{bmatrix}$

1.15 Exercise: Diagonalization of Matrices

Diagonalize the following matrices. Find the modal matrix P that diagonalizes (transforms) A.

(i) $\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ **Ans:** $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $D = \begin{pmatrix} 8 & 0 \\ 0 & 2 \end{pmatrix}$

(ii) $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ **Ans:** not diagonalizable since only one eigen vector $\begin{pmatrix} k \\ 0 \end{pmatrix}$ exists.

(iii) $\begin{pmatrix} 5 & 4 \\ 1 & 2 \end{pmatrix}$ **Ans:** $P = \begin{pmatrix} 4 & 1 \\ 1 & -1 \end{pmatrix}$, $D = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$

(iv) $\begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$ **Ans:** $P = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$, $D = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}$

(v) $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ **Ans:** $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $D = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$

(vi) $\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}$ **Ans:** $P = \begin{pmatrix} 1 & 2 \\ -1 & 5 \end{pmatrix}$, $D = \begin{pmatrix} -1 & 0 \\ 0 & 6 \end{pmatrix}$

(vii) $\begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$ hence find A^5 **Ans:** $P = \begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix}$,
$D = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$, $A^5 = \begin{pmatrix} 2344 & 781 \\ 2343 & 782 \end{pmatrix}$

(viii) $\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ **Ans:** no real eigen values, $\lambda = 1 \pm i$ so not diagonalizable over real.
Modal matrix over complex $P = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$, $D = \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}$.

Chapter 2

Differential Calculus and Integral Calculus

2.1 Introduction

Calculus is one of the most fundamental branches of mathematics, dealing with the study of change and accumulation. It provides essential tools for analyzing real-world problems in science, engineering, and technology. Differential Calculus focuses on the concept of a derivative, which measures the rate of change of a function with respect to its variables. It is widely used in analyzing motion, optimization, curve sketching, and in understanding the behavior of physical systems. Integral Calculus deals with the concept of integration, which represents the accumulation of quantities and the area under curves. It is applied in computing areas, volumes, work, probability distributions, and in solving differential equations. Together, differential and integral calculus form the foundation of mathematical analysis and play a vital role in engineering, physics, computer science, economics, and many other fields, making this module indispensable for B.Tech students.

When a function is differentiated more than once in quick succession, this is called successive differentiation. The rate of change of a function with respect to its independent variable is given by conventional differentiation; however, this idea can be extended to derive higher-order rates of change using successive differentiation.

2.2 Applications of Differential & Integral Calculus in Engineering

- Optimization of algorithms, processes, and design parameters.
- Gradient-based methods in artificial intelligence and machine learning.
- Signal and waveform analysis, including Fourier transforms and power calculations.
- Analysis of current, voltage, magnetic flux, and inductance in electrical systems.
- Velocity, acceleration, and motion analysis in kinematics and dynamics.
- Work done by variable forces, heat transfer, and vibration analysis.
- Slope, curvature, stress–strain analysis, and load distribution in structures.
- Area and volume calculations for irregular shapes and physical bodies.
- Reaction rate kinetics, material and energy balance, and reactor modeling.
- Diffusion, mass transfer, and accumulation processes in physical systems.
- Probability distributions and statistical applications in data science.
- Image processing, computer graphics, and optimization techniques.

2.3 Basic Concept of Derivatives

Theorem 5. *The derivative of a function measures* ***its rate of change with respect to its independent variable****. In simple words, it tells us how a function changes as its input changes.*

Mathematically, if $y = f(x)$ be a function of 'x', then the derivative of 'f' with respect to 'x' is defined as:

$$y_1 = \frac{dy}{dx} = f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{2.1}$$

This limit, if it exists, ***gives the instantaneous rate of change of 'f' at the point 'x'.***

Geometrical Meaning

- The derivative at a point on a curve represents **the slope of the tangent line to the curve at that point**.
- Positive derivative $\rightarrow$ curve is **increasing**
- Negative derivative $\rightarrow$ curve is **decreasing**.
- Zero derivative $\rightarrow$ **stationary point (could be maxima, minima, or point of inflection)**.

Physical Meaning

In physics, If $s = f(t)$ is displacement as a function of time 't' then:

- **First derivative (velocity):** $v = \frac{ds}{dt}$
- **Second derivative (acceleration):** $a = \frac{d^2s}{dt^2}$

Thus, the derivative relates directly to **speed, acceleration, and other rates of change in physical systems**.

Applications

The applications of the derivatives are:

- Finding **slopes and tangents**.
- Calculating **maxima and minima** for optimization.
- **Motion analysis in physics**.
- Rate of change in **economics and life sciences**.
- Developing **Taylor and Maclaurin series expansions**.

Basic rules of differentiation

If 'f' and 'g' are differentiable functions of 'x', then

1. $\frac{d}{dx}(kf) = k\frac{df}{dx}$, where '$k$' is a constant.
2. $\frac{d}{dx}(k) = 0$, where 'k' is a constant.
3. $\frac{d}{dx}(f \pm g) = \frac{df}{dx} \pm \frac{dg}{dx}$

4. $\frac{d}{dx}(f.g) = f\frac{dg}{dx} + g\frac{df}{dx}$ [Product rule]

5. $\frac{d}{dx}(\frac{f}{g}) = \frac{g\frac{df}{dx} - f\frac{dg}{dx}}{g^2}$ [Quotient rule]

6. $\frac{d}{dx}[f\{g(x)\}] = \frac{d}{dv}[f\{g(x)\}]\frac{dv}{dx}$ [Composite function rule]

Some well-known functions and their first-order derivatives

	$y = f(x)$	$\frac{dy}{dx} = f'(x)$		$y = f(x)$	$\frac{dy}{dx} = f'(x)$
1.	x^n	nx^{n-1}	17.	$(ax+b)^n$	$na(ax+b)^{n-1}$
2.	$\frac{1}{x}$	$-\frac{1}{x^2}$	18.	$\frac{1}{(ax+b)}$	$-\frac{a}{(ax+b)^2}$
3.	$\frac{1}{x^n}$	$-\frac{n}{x^{n+1}}$	19.	$\frac{1}{(ax+b)^n}$	$-\frac{na}{(ax+b)^{n+1}}$
4.	$log_e x$	$\frac{1}{x}$	20.	$log_e(ax+b)$	$\frac{a}{ax+b}$
5.	e^x	e^x	21.	$e^{(ax+b)}$	$ae^{(ax+b)}$
6.	$sin\,x$	$cos\,x$	22.	$sin\,(ax+b)$	$a\,cos\,(ax+b)$
7.	$cos\,x$	$-sin\,x$	23.	$cos\,(ax+b)$	$-a\,sin\,(ax+b)$
8.	$tan\,x$	$sec^2\,x$	24.	$tan\,(ax+b)$	$a\,sec^2\,(ax+b)$
9.	$cot\,x$	$-cosec^2\,x$	25.	$cot\,(ax+b)$	$-a\,cosec^2\,(ax+b)$
10.	$sec\,x$	$sec\,x.tan\,x$	26.	$sec\,(ax+b)$	$a\,sec\,(ax+b).tan\,(ax+b)$
11.	$cosec\,x$	$-cosec\,x.cot\,x$	27.	$cosec\,(ax+b)$	$-a\,cosec\,(ax+b).cot\,(ax+b)$
12.	$sin^{-1}x$	$\frac{1}{\sqrt{1-x^2}}$	28.	$sin^{-1}(ax+b)$	$\frac{a}{\sqrt{1-(ax+b)^2}}$
13.	$cos^{-1}x$	$-\frac{1}{\sqrt{1-x^2}}$	29.	$cos^{-1}(ax+b)$	$-\frac{a}{\sqrt{1-(ax+b)^2}}$
14.	$tan^{-1}x$	$\frac{1}{1+x^2}$	30.	$tan^{-1}(ax+b)$	$\frac{a}{1+(ax+b)^2}$
15.	$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$	31.	$\sqrt{(ax+b)}$	$\frac{a}{2\sqrt{(ax+b)}}$
16.	a^x	$a^x\,log_e a$	32.	$a^{(px+q)}$	$p\,a^{(px+q)}\,log_e a$

Successive Differentiation

Let $y = f(x)$ be a differentiable function of x. The derivative $\frac{dy}{dx}$ is called the first derivative of y with respect to x. In general, it is a function of x. The derivative of $\frac{dy}{dx}$

is called the second derivative of y with respect to x and it is denoted as $\frac{d^2y}{dx^2}$.
Thus,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right).$$

Similarly, the derivative of $\frac{d^2y}{dx^2}$ is called the third derivative of y with respect to x and it is denoted as $\frac{d^3y}{dx^3}$.

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left(\frac{d^2y}{dx^2}\right)$$

and so on.

The n^{th} differential coefficient of y with respect to x is denoted by

$$\frac{d^ny}{dx^n} = \frac{d}{dx}\left(\frac{d^{n-1}y}{dx^{n-1}}\right).$$

Thus, the successive derivatives of y are

$$\frac{dy}{dx}, \frac{d^2y}{dx^2}, \frac{d^3y}{dx^3}, \ldots, \frac{d^ny}{dx^n}.$$

These derivatives are also denoted by

- $y_1, y_2, y_3, \cdots y_n$ or
- $Dy,\ D^2y,\ D^3y, \ldots, D^ny$, where $D = \frac{d}{dx}$, $D^2 = \frac{d^2}{dx^2}$, $D^n = \frac{d^n}{dx^n}$. or
- $f'(x), f''(x), f'''(x), \ldots, f^{(n)}(x)$ or
- $y', y'', y''', \ldots, y^{(n)}$.

The process of finding second and higher order derivatives is called **successive differentiation**.

For example: If $y = x^7$, then

$$\frac{dy}{dx} = 7x^6, \qquad \frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{dx}\right) = 7 \cdot 6x^5 = 42x^5, \qquad \frac{d^3y}{dx^3} = 42 \cdot 5x^4 = 210x^4,$$

$$\frac{d^4y}{dx^4} = 210 \cdot 4x^3 = 840x^3, \qquad \frac{d^5y}{dx^5} = 840 \cdot 3x^2 = 2520x^2, \qquad \frac{d^6y}{dx^6} = 2520 \cdot 2x = 5040x,$$

$$\frac{d^7y}{dx^7} = 5040, \qquad \frac{d^8y}{dx^8} = 0,$$

and other higher derivatives vanish.

Note. Let $y = x^3$,

$$\frac{dy}{dx} = 3x^2, \quad \frac{d^2y}{dx^2} = 6x, \quad \frac{d^3y}{dx^3} = 6.$$

Now,

$$\left(\frac{dy}{dx}\right)^2 = (3x^2)^2 = 9x^4, \quad \text{and} \quad \left(\frac{dy}{dx}\right)^3 = (3x^2)^3 = 27x^6.$$

Hence, we find that

$$\frac{d^2y}{dx^2} \neq \left(\frac{dy}{dx}\right)^2, \quad \text{and} \quad \frac{d^3y}{dx^3} \neq \left(\frac{dy}{dx}\right)^3.$$

In general,

$$\frac{d^ny}{dx^n} \neq \left(\frac{dy}{dx}\right)^n.$$

2.4 Worked Examples

Example 1. Find the first two differential coefficients with respect to x of $x^3 \cos x$.
Solution. Let

$$y = x^3 \cos x.$$

Differentiating with respect to x,

$$\frac{dy}{dx} = x^2[-\sin x] + \cos x \cdot 2x = -x^2 \sin x + 2x \cos x.$$

Again differentiating with respect to x,

$$\frac{d^2y}{dx^2} = [-x^2 \cos x - \sin x \cdot 2x] + 2[(x(-\sin x) + \cos x \cdot 1)]$$

$$= -x^2 \cos x - 2x \sin x - 2x \sin x + 2 \cos x$$

$$= (2 - x^2) \cos x - 4x \sin x.$$

Example 2. If $y = e^{ax} \sin bx$, then show that

$$\frac{d^2y}{dx^2} - 2a\frac{dy}{dx} + (a^2 + b^2)y = 0.$$

.
Solution. Given

$$y = e^{ax} \sin bx.$$

Differentiating with respect to x,

$$\frac{dy}{dx} = e^{ax} \cos bx \cdot b + \sin bx \cdot ae^{ax} = e^{ax}(b \cos bx + a \sin bx).$$

Again differentiating with respect to x,

$$\frac{d^2y}{dx^2} = e^{ax} \left[-b^2 \sin bx + ab \cos bx + ab \cos bx + a^2 \sin bx\right]$$

$$= e^{ax} \left[(a^2 - b^2) \sin bx + 2ab \cos bx\right].$$

Therefore,

$$LHS = \frac{d^2y}{dx^2} - 2a\frac{dy}{dx} + (a^2 + b^2)y$$

$$= e^{ax} \left[(a^2 - b^2) \sin bx + 2ab \cos bx\right] - 2ae^{ax}[b \cos bx + a \sin bx] + (a^2 + b^2)e^{ax} \sin bx$$

$$= e^{ax} \left[(a^2 \sin bx - b^2 \sin bx + 2ab \cos bx) - 2ab \cos bx - 2a^2 \sin bx + a^2 \sin bx + b^2 \sin bx\right]$$

$$= e^{ax} \times 0 = 0 = RHS.$$

Example 3. If $y = \cos^{-1} x$, then show that

$$(1 - x^2)y_2 - xy_1 = 0.$$

.
Solution. Given

$$y = \cos^{-1} x.$$

Differentiating with respect to x,

$$\frac{dy}{dx} = -\frac{1}{\sqrt{1 - x^2}}.$$

Squaring,

$$\left(\frac{dy}{dx}\right)^2 = \frac{1}{1-x^2},$$

thus,

$$y_1^2 = \frac{1}{1-x^2} \quad \Rightarrow \quad (1-x^2)y_1^2 = 1.$$

Again differentiating with respect to x,

$$(1-x^2)2y_1y_2 + y_1^2(-2x) = 0.$$

Dividing by $2y_1$,

$$(1-x^2)y_2 - xy_1 = 0.$$

Example 4. If $y = \frac{\log x}{x^2}$, then show that

$$x^3\frac{d^3y}{dx^3} + 6x^2\frac{d^2y}{dx^2} + 4x\frac{dy}{dx} - 4y = 0.$$

Solution. Given

$$y = \frac{\log x}{x^2}.$$

Differentiating with respect to x,

$$\frac{dy}{dx} = \frac{x^2 \cdot \frac{1}{x} - \log x \cdot 2x}{x^4} = \frac{x - 2x\log x}{x^4} = \frac{x[1-2\log x]}{x^4} = \frac{1-2\log x}{x^3}.$$

Again differentiating with respect to x,

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{1-2\log x}{x^3}\right).$$

Using quotient rule,

$$= \frac{x^3 \cdot \left(\frac{-2}{x}\right) - (1-2\log x)3x^2}{x^6}$$

$$= \frac{-2x^2 - 3x^2(1-2\log x)}{x^6}$$

$$= \frac{-2x^2 - 3x^2 + 6x^2\log x}{x^6}$$

$$= \frac{x^2[-5+6\log x]}{x^6}$$

$$= \frac{6\log x - 5}{x^4}.$$

Again differentiating with respect to x,

$$\frac{d^3y}{dx^3} = \frac{d}{dx}\left(\frac{6\log x - 5}{x^4}\right).$$

Using quotient rule,

$$= \frac{x^4 \cdot \frac{6}{x} - (6\log x - 5)4x^3}{x^8}$$

$$= \frac{6x^3 - 24x^3\log x + 20x^3}{x^8}$$

$$= \frac{x^3[26 - 24\log x]}{x^8}$$

$$= \frac{26 - 24\log x}{x^5}.$$

Therefore,

$$LHS = x^3\frac{d^3y}{dx^3} + 6x^2\frac{d^2y}{dx^2} + 4x\frac{dy}{dx} - 4y$$

$$= x^3 \cdot \frac{26 - 24\log x}{x^5} + 6x^2 \cdot \frac{6\log x - 5}{x^4} + 4x \cdot \frac{1 - 2\log x}{x^3} - 4 \cdot \frac{\log x}{x^2}$$

$$= \frac{26 - 24\log x}{x^2} + \frac{36\log x - 30}{x^2} + \frac{4 - 8\log x}{x^2} - \frac{4\log x}{x^2}.$$

Thus,

$$LHS = \frac{[26 - 24\log x + 36\log x - 30 + 4 - 8\log x - 4\log x]}{x^2}$$

$$= \frac{[26 - 30 + 4 + (-24\log x + 36\log x - 8\log x - 4\log x)]}{x^2}$$

$$= \frac{[0]}{x^2} = 0 = RHS.$$

Example 5. If $x = a\cos^3\theta,\ y = b\sin^3\theta$, then find $\frac{d^2y}{dx^2}$.
Solution. Given

$$x = a\cos^3\theta, \quad y = b\sin^3\theta,$$

which are parametric equations.
Differentiating with respect to θ, we get,

$$\frac{dx}{d\theta} = 3a\cos^2\theta(-\sin\theta) = -3a\cos^2\theta\sin\theta,$$

and

$$\frac{dy}{d\theta} = 3b\sin^2\theta\cos\theta.$$

Therefore,

$$\frac{dy}{dx} = \frac{dy/d\theta}{dx/d\theta} = \frac{3b\sin^2\theta\cos\theta}{-3a\cos^2\theta\sin\theta} = -\frac{b}{a}\tan\theta.$$

Again differentiating,

$$\frac{d^2y}{dx^2} = \frac{d}{d\theta}\left(-\frac{b}{a}\tan\theta\right)\cdot\frac{d\theta}{dx}.$$

We have,

$$\frac{d}{d\theta}\left(-\frac{b}{a}\tan\theta\right) = -\frac{b}{a}\sec^2\theta,$$

and

$$\frac{d\theta}{dx} = \frac{1}{dx/d\theta} = \frac{1}{-3a\cos^2\theta\sin\theta}.$$

Therefore,

$$\frac{d^2y}{dx^2} = \frac{-\frac{b}{a}\sec^2\theta}{-3a\cos^2\theta\sin\theta} = \frac{b}{3a^2}\sec^4\theta\csc\theta.$$

Example 6. If $x = \sin t$ and $y = \sin pt$, then prove that

$$(1-x^2)y_2 - xy_1 + p^2y = 0.$$

Solution. Given

$$x = \sin t, \quad y = \sin pt.$$

Differentiating with respect to t,

$$\frac{dx}{dt} = \cos t, \quad \frac{dy}{dt} = p\cos pt.$$

Therefore,

$$y_1 = \frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{p\cos pt}{\cos t}.$$

Squaring,

$$y_1^2 = \frac{p^2\cos^2 pt}{\cos^2 t} = p^2\frac{1-\sin^2 pt}{1-\sin^2 t} = p^2\frac{1-y^2}{1-x^2}.$$

Therefore,

$$(1-x^2)y_1^2 = p^2(1-y^2).$$

Differentiating with respect to x,

$$(1-x^2)2y_1y_2 + y_1^2(-2x) = -2p^2yy_1.$$

Dividing by $2y_1$,

$$(1-x^2)y_2 - xy_1 + p^2y = 0.$$

2.5 Exercise:

1. If $y = e^x\cos x$, then prove that

$$\frac{d^4y}{dx^4} + 4y = 0.$$

2. If $y = \sin(\sin x)$, then prove that

$$\frac{d^2y}{dx^2} + \tan x\frac{dy}{dx} + y\cos^2 x = 0.$$

3. If $y = Ax^{n+1} + Bx^{-n}$, then prove that

$$x^2\frac{d^2y}{dx^2} = n(n+1)y.$$

4. If $y = Ae^{-hx}\cos(pt+e)$, then show that

$$\frac{d^2y}{dt^2} + 2h\frac{dy}{dt} + n^2y = 0,$$

where $n^2 = p^2 + k^2$.

5. If $y = a\cos(\log x) + b\sin(\log x)$, then show that

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + y = 0.$$

6. If $y = \frac{ax^2+bx+c}{1-x}$, then show that

$$(1-x)y_3 = 3y_2.$$

7. If $x = \left(t + \frac{1}{t}\right),\ y = \frac{1}{2}\left(t - \frac{1}{t}\right)$, then prove that

$$\frac{dy}{dx} = \frac{2t}{(1-t^2)}.$$

8. If $y = (\sin^{-1} x)^2$, then show that

$$(1-x^2)\frac{d^2y}{dx^2} + x\frac{dy}{dx} = 2.$$

9. If $y = (\tan^{-1} x)^n$, then prove that

$$(x^2+1)y_2 + 2x(x^2+1)y_1 - 2 = 0.$$

10. If $xy = ae^x + be^{-x}$, then prove that

$$x\frac{d^2y}{dx^2} + 2\frac{dy}{dx} - xy = 0.$$

11. If $y^2 = ax^2 + 2bx + c$, where a, b, c are constants, then show that

$$\frac{d^2x}{dy^2} = \frac{b^2 - ac}{(ax+b)^3}.$$

12. If $x^2 + y^2 = 3axy$, prove that

$$D^2y = \frac{2a^2xy}{(ax-y)^3}.$$

13. If $y = e^t(ax+b)$, then prove that

$$\frac{d^2y}{dx^2} - 6\frac{dy}{dx} + 9y = 0.$$

14. If $x = (a+bt)e^{-nt}$, then show that

$$\frac{d^2x}{dt^2} + 2n\frac{dx}{dt} + n^2x = 0.$$

15. If $y = x^{-n}\log x$, show that

$$x\frac{d^2y}{dx^2} - (n-2)\frac{dy}{dx} - (n-1)nx^{-n-1} = 0.$$

16. If $xy = ae^x + be^{-x}$, then prove that

$$x\frac{d^2y}{dx^2} + 2\frac{dy}{dx} - xy = 0.$$

17. If $x = 2\cos t - \cos 2t,\ y = 2\sin t - \sin 2t$,
then prove that the value of $\frac{d^3y}{dx^3} = -3$ when $t = \frac{\pi}{2}$.

The nth Derivative of Standard Functions

1. Find the n^{th} derivative of e^{ax}.
Solution. Let

$$y = e^{ax}.$$

Differentiating with respect to x successively, we get

$$y_1 = ae^{ax}, \quad y_2 = a^2e^{ax}, \quad y_3 = a^3e^{ax}, \ldots$$

Thus,

$$y_n = a^n e^{ax}.$$

Note:

(1) Putting $a = 1$ in the above, we get $y = e^x$ and $y_n = e^x$.
(2) Since $a = e^{\log a}$, we have

$$a^x = e^{x \log a}.$$

Therefore,

$$\frac{d^n}{dx^n}(a^x) = (\log a)^n e^{x \log a} = (\log a)^n a^x.$$

So, if $y = a^x$, then

$$y_n = (\log a)^n a^x.$$

2. Find the n^{th} derivative of $(ax+b)^m$.
Solution.
Let

$$y = (ax+b)^m.$$

Differentiating with respect to x successively:

$$y_1 = m(ax+b)^{m-1} \cdot a = am(ax+b)^{m-1},$$

$$y_2 = a \cdot m(m-1)(ax+b)^{m-2},$$

$$y_3 = a^2 m(m-1)(m-2)(ax+b)^{m-3},$$

$$\vdots$$

$$y_n = a^n m(m-1)(m-2)\dots(m-n+1)(ax+b)^{m-n}.$$

This result is true for all values of m.

Particular Cases

Case 1: If m is a positive integer and equal to n, then

$$y_n = a^n n(n-1)(n-2)\dots 1(ax+b)^{m-n},$$

$$= a^n n!,$$

and

$$\frac{d^n}{dx^n}(ax+b)^n = n!a^n.$$

Since y_n is a constant, $y_{n+1} = 0, y_{n+2} = 0$ and so on.
Case 2: If m is an integer and $m < n$, then the nth derivative of $(ax+b)^m$ is 0.
Case 3: If $a = 1, b = 0$, then

$$y = x^m,$$

$$y_n = m(m-1)(m-2)\dots(m-n+1)x^{m-n}.$$

If m is a positive integer and $m = n$, then

$$y_n = n(n-1)(n-2)\dots 1 = n!.$$

3. Find the n^{th} derivative of $\frac{1}{ax+b}$.

Solution.
Let

$$y = \frac{1}{ax+b} = (ax+b)^{-1}.$$

Differentiating successively:

$$y_1 = (-1)(ax+b)^{-2} \cdot a = \frac{-a}{(ax+b)^2},$$

$$y_2 = a(-1)(-2)(ax+b)^{-3} \cdot a = \frac{(-1)^2 2! a^2}{(ax+b)^3},$$

$$y_3 = a^2(-1)(-2)(-3)(ax+b)^{-4} \cdot a = \frac{(-1)^3 3! a^3}{(ax+b)^4},$$

$$\vdots$$

$$y_n = a^{n-1}(-1)(-2)(-3)\ldots(-n)(ax+b)^{-(n+1)} \cdot a,$$

$$y_n = \frac{(-1)^n n! a^n}{(ax+b)^{n+1}}.$$

Particular case

If $a = 1, b = 0$, then $y = \log x$. Thus,

$$y_n = (-1)^{n-1}(n-1)! x^{-n},$$

where n is a positive integer.

—

4. Find the n^{th} derivative of $\log(ax+b)$.

Solution.

Let

$$y = \log_e(ax+b).$$

Differentiating with respect to x,

$$y_1 = \frac{1}{ax+b} \cdot a = \frac{a}{ax+b}.$$

Again differentiating,

$$y_2 = a(-1)(ax+b)^{-2} \cdot a = \frac{-a^2}{(ax+b)^2},$$

$$y_3 = a^2(-1)(-2)(ax+b)^{-3} \cdot a = \frac{2a^3}{(ax+b)^3},$$

$$\vdots$$

$$y_n = (-1)^{n-1}(n-1)! \frac{a^n}{(ax+b)^n}.$$

—

5. Find the n^{th} derivative of $\sin(ax+b)$.

Solution.
Let

$$y = \sin(ax+b).$$

Differentiating successively,

$$y_1 = \cos(ax+b)\cdot a = a\cos(ax+b),$$

$$y_2 = -a^2\sin(ax+b),$$

$$y_3 = -a^3\cos(ax+b),$$

$$y_4 = a^4\sin(ax+b),$$

$$\vdots$$

Thus,

$$y_n = a^n \sin\left(ax+b+n\frac{\pi}{2}\right).$$

Particular case
If $a=1, b=0$, then $y=\sin x$. Therefore,

$$y_n = \sin\left(x+n\frac{\pi}{2}\right).$$

6. Find the n^{th} derivative of $\cos(ax+b)$.

Solution.
Let

$$y = \cos(ax+b).$$

Differentiating w.r.t. x successively, we get

$$y_1 = -\sin(ax+b)\cdot a = -a\sin(ax+b) = a\cos\left(ax+b+\frac{\pi}{2}\right),$$

$$y_2 = -a\sin\left(ax+b+\frac{\pi}{2}\right)\cdot a = a^2\cos\left(ax+b+2\frac{\pi}{2}\right),$$

$$y_3 = a^3\cos\left(ax+b+3\frac{\pi}{2}\right),$$

$$\vdots$$

$$y_n = a^n \cos\left(ax + b + n\frac{\pi}{2}\right).$$

—

Particular case:
If $a = 1, b = 0$, then $y = \cos x$. Thus,

$$y_n = \cos\left(x + n\frac{\pi}{2}\right).$$

—

7. Find the n^{th} derivative of $e^{ax}\sin(bx + c)$.

Solution.
Let

$$y = e^{ax}\sin(bx + c).$$

Differentiating successively,

$$y_1 = e^{ax}[b\cos(bx + c) + a\sin(bx + c)],$$

If $a = r\cos\theta,\ b = r\sin\theta$, then

$$a^2 + b^2 = r^2(\cos^2\theta + \sin^2\theta) = r^2, \quad r = \sqrt{a^2 + b^2},$$

and

$$\tan\theta = \frac{b}{a}.$$

Thus,

$$y_1 = re^{ax}\sin(bx + c + \theta),$$

$$y_2 = r^2e^{ax}\sin(bx + c + 2\theta),$$

$$y_3 = r^3e^{ax}\sin(bx + c + 3\theta),$$

$$\vdots$$

$$y_n = r^ne^{ax}\sin\left(bx + c + n\tan^{-1}\frac{b}{a}\right).$$

Therefore,

$$y_n = (a^2+b^2)^{n/2} e^{ax} \sin\left(bx + c + n\tan^{-1}\frac{b}{a}\right).$$

8. Find the n^{th} derivative of $e^{ax}\cos(bx+c)$

Solution.

Let

$$y = e^{ax}\cos(bx+c).$$

Differentiating w.r.t. x successively, we get

$$y_1 = e^{ax}[-\sin(bx+c)\cdot b + \cos(bx+c)\cdot a],$$

$$y_1 = e^{ax}[a\cos(bx+c) - b\sin(bx+c)].$$

Put $a = r\cos\theta,\ b = r\sin\theta$, then $r^2 = a^2+b^2 \Rightarrow r = (a^2+b^2)^{1/2}$ and

$$\tan\theta = \frac{b}{a} \quad \Rightarrow \quad \theta = \tan^{-1}\left(\frac{b}{a}\right).$$

Therefore,

$$y_1 = e^{ax}[r\cos\theta\cos(bx+c) - r\sin\theta\sin(bx+c)],$$

$$= re^{ax}[\cos\theta\cos(bx+c) - \sin\theta\sin(bx+c)],$$

$$= re^{ax}\cos(bx+c+\theta).$$

Similarly,

$$y_2 = r^2e^{ax}\cos(bx+c+2\theta),$$

$$y_3 = r^3e^{ax}\cos(bx+c+3\theta),$$

$$\vdots$$

$$y_n = (a^2+b^2)^{n/2}e^{ax}\cos\left(bx+c+n\tan^{-1}\frac{b}{a}\right).$$

2.6 Worked Examples:

1. The n^{th} derivative using partial fractions

Example 1. Find the n^{th} derivative of (i) $\frac{x}{(x-2)^2}$, (ii) $\frac{2x}{(2x+1)(x-1)}$.

Solution.

(i) $\frac{x}{(x-2)^2}$

Let $y = \frac{x}{(x-2)^2}$.

Split the RHS into partial fractions to reduce it to the form $\frac{1}{(ax+b)^m}$.

Let

$$\frac{x}{(x-2)^2} = \frac{A}{x-2} + \frac{B}{(x-2)^2}$$
$$\implies x = A(x-2) + B.$$

Put $x = 2$, then:

$$2 = A \cdot 0 + B \implies B = 2.$$

Equating the coefficients of x, $A = 1$.

$$\therefore \frac{x}{(x-2)^2} = \frac{1}{x-2} + \frac{2}{(x-2)^2}$$
$$\implies y = \frac{1}{x-2} + \frac{2}{(x-2)^2}$$

then

$$y_n = \frac{(-1)^n n!}{(x-2)^{n+1}} + \frac{(-1)^n (n+1)!}{(x-2)^{n+2}} \quad \text{[Refer formula 2]}$$

(ii) $\dfrac{2x}{(2x+1)(x-1)}$

Let

$$y = \frac{2x}{(2x+1)(x-1)}$$

Split the R.H.S into partial fractions.

Let

$$\frac{2x}{(2x+1)(x-1)} = \frac{A}{2x+1} + \frac{B}{x-1}$$
$$\implies 2x = A(x-1) + B(2x+1)$$

Put $x = \frac{-1}{2}$, then:

$$2\left(-\frac{1}{2}\right) = A\left(-\frac{1}{2} - 1\right) \implies -1 = A\left(-\frac{3}{2}\right) \implies A = \frac{2}{3}$$

Put $x = 1$, then:

$$2 \cdot 1 = B(2 \cdot 1 + 1) \implies 2 = 3B \implies B = \frac{2}{3}$$

$$\therefore \frac{2x}{(2x+1)(x-1)} = \frac{2/3}{2x+1} + \frac{2/3}{x-1}$$
$$\implies y = \frac{2}{3} \cdot \frac{1}{2x+1} + \frac{2}{3} \cdot \frac{1}{x-1}$$
$$\text{Hence,} \quad y_n = \frac{2}{3} \cdot \frac{(-1)^n n! 2^n}{(2x+1)^{n+1}} + \frac{2}{3} \cdot \frac{(-1)^n n!}{(x-1)^{n+1}}$$
$$= \frac{(-1)^n n! 2^{n+1}}{3(2x+1)^{n+1}} + \frac{2(-1)^n n!}{3(x-1)^{n+1}} \quad \text{[Refer formula 3]}$$

Example 2 Find the n^{th} derivative of $\dfrac{x^4}{(x-1)(x-2)}$.

Solution.

$$\text{Let} \quad y = \frac{x^4}{(x-1)(x-2)}$$

First split the R.H.S into partial fractions. Since the degree of the numerator is greater than the degree of denominator, it is an improper fraction.

We divide $\dfrac{x^4}{(x-1)(x-2)} = x^2 - 3x + 2$ and then split the proper fraction into partial fractions.

$$\therefore \frac{x^4}{(x-1)(x-2)} = x^2 + 3x + 7 + \frac{15x-14}{(x-1)(x-2)}$$
$$\text{Let} \quad \frac{15x-14}{(x-1)(x-2)} = \frac{A}{x-1} + \frac{B}{x-2}$$
$$\implies 15x - 14 = A(x-2) + B(x-1)$$

Put $x = 1$, then:

$$15 \cdot 1 - 14 = A(1-2) + B \cdot 0 \implies 1 = -A \implies A = -1$$

Put $x = 2$, then:

$$15 \cdot 2 - 14 = B(2-1) \implies 16 = B \implies B = 16$$

$$\therefore \frac{15x-14}{(x-1)(x-2)} = \frac{-1}{x-1} + \frac{16}{x-2}$$
$$\therefore \frac{x^4}{(x-1)(x-2)} = x^2 + 3x + 7 - \frac{1}{x-1} + \frac{16}{x-2}$$
$$\implies y = x^2 + 3x + 7 - \frac{1}{x-1} + \frac{16}{x-2}$$
$$\text{Hence,} \quad y_n = -\frac{(-1)^n n!}{(x-1)^{n+1}} + \frac{16(-1)^n n!}{(x-2)^{n+1}} \qquad n > 2$$

2. The n^{th} derivatives of trigonometric functions

If y is a simple power of sine or cosine or product of sine and cosine, then they can be expressed as a sum of sines and cosines of multiple angles and n^{th} derivative can be found.

Example 3 Find the n^{th} derivative of (i) $\sin^2 x$ (ii) $\sin^3 x$.

Solution. (i) $\sin^2 x$

$$\text{Let} \quad y = \sin^2 x = \frac{1-\cos 2x}{2} = \frac{1}{2} - \frac{\cos 2x}{2}$$

$$\therefore y_n = -\frac{1}{2} \cdot 2^n \cos\left(2x + n\frac{\pi}{2}\right) = -2^{n-1}\cos\left(2x + \frac{n\pi}{2}\right) \qquad \text{[Refer the formula 5]}$$

(ii) **Let**

$$y = \sin^3 x \implies \frac{1}{4}(3\sin x - \sin 3x) = \frac{3}{4}\sin x - \frac{1}{4}\sin 3x \quad [\text{since } \sin 3x = 3\sin x - 4\sin^3 x]$$

$$\text{Hence,} \quad y_n = \frac{3}{4}\sin\left(x + n\frac{\pi}{2}\right) - \frac{1}{4}\sin\left(3x + n\frac{\pi}{2}\right) \quad \text{[Refer the formula 5]}$$

Example 4 Find the n^{th} derivative of (i) $\cos^4 x$ (ii) $\sin x \sin 2x \sin 3x$.

Solution. (i) **Let**

$$y = \cos^4 x = (\cos^2 x)^2 = \left(\frac{1+\cos 2x}{2}\right)^2$$

$$= \frac{1}{4}\left(1 + 2\cos 2x + \cos^2 2x\right)$$

$$= \frac{1}{4}\left(1 + 2\cos 2x + \frac{1+\cos 4x}{2}\right) = \frac{1}{4} + \frac{1}{2}\cos 2x + \frac{1}{8}\cos 4x$$

$$\implies y = \frac{3}{8} + \frac{1}{2}\cos 2x + \frac{1}{8}\cos 4x$$

$$\text{Hence,} \quad y_n = \frac{1}{2}2^n\cos\left(2x + \frac{n\pi}{2}\right) + \frac{1}{8}4^n\cos\left(4x + \frac{n\pi}{2}\right)$$

(ii) **Let** $y = \sin x \sin 2x \sin 3x$

We know that

$$\sin A \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)] \quad \text{and} \quad \sin A \cos B = \frac{1}{2}[\sin(A+B) + \sin(A-B)]$$

$$y = \frac{1}{2}\sin x[\cos x - \cos 5x] = \frac{1}{2}(\sin x \cos x - \sin x \cos 5x)$$

$$= \frac{1}{2}\left[\frac{1}{2}\sin 2x - \frac{1}{2}(\sin 6x - \sin 4x)\right] = \frac{1}{4}\sin 2x - \frac{1}{4}(\sin 6x - \sin 4x)$$

$$\text{Hence,} \quad y_n = \frac{1}{4}2^n \sin\left(2x + \frac{n\pi}{2}\right) - \frac{1}{4}6^n \sin\left(6x + \frac{n\pi}{2}\right) + \frac{1}{4}4^n \sin\left(4x + \frac{n\pi}{2}\right)$$

$$\implies y_n = 2^{n-2}\sin\left(2x + \frac{n\pi}{2}\right) - 3^n \cdot 2^{n-2}\sin\left(6x + \frac{n\pi}{2}\right) + 4^{n-2}\sin\left(4x + \frac{n\pi}{2}\right) \quad \text{[Refer formula 5]}$$

Example 5 Find the n^{th} differential coefficient of $\sin^3\theta\cos^5\theta$.
Solution. Let

$$y = \sin^3\theta\cos^5\theta \quad \text{and let} \quad x = \cos\theta + i\sin\theta = e^{i\theta}$$

$$\frac{1}{x} = \frac{1}{\cos\theta + i\sin\theta} = \frac{1}{e^{i\theta}} = e^{-i\theta} = \cos\theta - i\sin\theta$$

$$\therefore \quad x + \frac{1}{x} = 2\cos\theta \quad \text{and} \quad x - \frac{1}{x} = 2i\sin\theta$$

By De Moivre's theorem,

$$x^n = \cos n\theta + i\sin n\theta \quad \text{and} \quad \frac{1}{x^n} = \cos n\theta - i\sin n\theta$$

$$x^n + \frac{1}{x^n} = 2\cos n\theta \quad \text{and} \quad x^n - \frac{1}{x^n} = 2i\sin n\theta$$

$$\therefore \quad 2^5\cos^5\theta = \left(x + \frac{1}{x}\right)^5 \quad \text{and} \quad 2^3 i\sin^3\theta = \left(x - \frac{1}{x}\right)^3$$

$$\therefore 2^7 i\sin^3\theta \cdot 2^5\cos^5\theta = \left(x - \frac{1}{x}\right)^3\left(x + \frac{1}{x}\right)^5$$

$$\implies -i2^8\sin^3\theta\cos^5\theta = \left[\left(x - \frac{1}{x}\right)\left(x + \frac{1}{x}\right)\right]^2\left(x + \frac{1}{x}\right)^3$$

$$= \left(x^2 - \frac{1}{x^2}\right)^2\left(x^2 + 2 + \frac{1}{x^2}\right)$$

$$= \left(x^4 + 2x^2 + 1 + \frac{2}{x^2} + \frac{1}{x^4}\right) - 6$$

$$= x^8 + 2x^6 + x^4 - 3x^6 - 6x^2 - 3 + \frac{6}{x^2} + \frac{1}{x^4}$$

$$= \left(x^8 + \frac{1}{x^8}\right) + 2\left(x^6 + \frac{1}{x^6}\right) - 3\left(x^4 + \frac{1}{x^4}\right) - 6$$

$$\implies -i2^8 \sin^3\theta\cos^5\theta = 2i\sin 8\theta + 2\cdot 2i\sin 6\theta - 2\cdot 2i\sin 4\theta - 6\cdot 2i\sin 2\theta$$

$$\implies -2^7 \sin^3\theta\cos^5\theta = \sin 8\theta + 2\sin 6\theta - 2\sin 4\theta - 6\sin 2\theta \quad [\text{dividing by } 2i]$$

$$\implies \sin^3\theta\cos^5\theta = \frac{1}{2^7}\left[\sin 8\theta + 2\sin 6\theta - 2\sin 4\theta - 6\sin 2\theta\right]$$
$$= \frac{1}{2^7}\left[6\sin 2\theta + 2\sin 4\theta - 2\sin 6\theta - \sin 8\theta\right]$$

$$y = \frac{1}{2^7}\left(6\sin 20^\circ + 2\sin 40^\circ - 2\sin 60^\circ - \sin 80^\circ\right)$$

$$y_n = \frac{1}{2^7}\left[6\cdot 2^n \sin\left(20 + \frac{n\pi}{2}\right) + 2\cdot 4^n \sin\left(40 + \frac{n\pi}{2}\right) - 2\cdot 6^n \sin\left(60 + \frac{n\pi}{2}\right)\right.$$

$$\left. -8^n \sin\left(80 + \frac{n\pi}{2}\right)\right]$$

Example 6 If $y = \dfrac{\log x}{x}$, then prove that

$$y_n = \frac{(-1)^n n!}{x^{n+1}}\left[\log x - 1 - \frac{1}{2} - \frac{1}{3} - \cdots - \frac{1}{n}\right].$$

Solution. Given

$$y = \frac{\log x}{x}$$

Differentiating w.r.t. x, we get

$$y_1 = \frac{1}{x^2} - \frac{\log x}{x^2} = -\frac{1}{x^2}[\log x - 1].$$

Again differentiating w.r.t. x, we get

$$y_2 = (-1)\left[\frac{1}{x^3} + (\log x - 1)\left(-\frac{2}{x^3}\right)\right]$$
$$= (-1)\left[\frac{1}{x^3} - \frac{2}{x^3}(\log x - 1)\right]$$
$$= (-1)^2 2! \frac{1}{x^3}\left[\log x - 1 - \frac{1}{2}\right]$$

Again differentiating w.r.t. x, we get

$$y_3 = (-1)^2 \cdot 2! \left[\frac{1}{x^4} + \left(\log x - 1 - \frac{1}{2}\right)\left(-\frac{3}{x^4}\right)\right]$$

$$= (-1)^3 3! \frac{1}{x^4}\left[\log x - 1 - \frac{1}{2} - \frac{1}{3}\right]$$

Proceeding in this way or by induction, we get:

$$y_n = \frac{(-1)^n n!}{x^{n+1}}\left[\log x - 1 - \frac{1}{2} - \frac{1}{3} - \cdots - \frac{1}{n}\right]$$

2.7 Exercise:

1. Find the n^{th} derivative of (i) $x^n \sin 3x$ (ii) $x^2 \cos 4x$ (iii) $\sin 3x \cdot \sin 2x$
2. Find the n^{th} derivative of $x^4 + \log_e(3x^3 + 5x - 2)$.
3. Find the n^{th} derivative of $\dfrac{x^2}{2x^2 + 7x + 6}$.
4. Find the n^{th} derivative of $\dfrac{1}{x^2 - 6x + 8}$.
5. Find the n^{th} derivative of $\sin^3 x \cos^2 x$.
6. Find the n^{th} derivative of $e^{3x} \cos 6x + \log 4x$.

Answers to Exercise

1. (i) $x^2 \cdot 3^n \sin\left(\frac{n\pi}{2} + 3x\right) + 2nx \cdot 3^{n-1} \sin\left((n-1)\frac{\pi}{2} + 3x\right) + n(n-1)3^{n-2} \sin\left((n-2)\frac{\pi}{2} + 3x\right)$

   (ii) $x^2 \cdot 4^n \cos\left(\frac{\pi}{2} + 4x\right) + 2nx \cdot 4^{n-1} \cos\left((n-1)\frac{\pi}{2} + 4x\right) + n(n-1)4^{n-2} \cos\left((n-2)\frac{\pi}{2} + 4x\right)$

   (iii) $\frac{1}{2}\cos\left(\frac{\pi}{2} + x\right) - \frac{1}{2}5^n \cos\left(\frac{\pi}{2} + 5x\right)$

2. $(-1)^n (n-1)! \left[\dfrac{3^n}{(3x-1)^{n+1}} + \dfrac{1}{(x+2)^n}\right]$
3. $(-1)^n n! \left[\dfrac{9 \cdot 2^{n-1}}{(2x+3)^{n+1}} - \dfrac{4}{(x+2)^{n+1}}\right]$
4. $\dfrac{1}{2}(-1)^n n! \left[\dfrac{1}{(x-2)^{n+1}} + \dfrac{1}{(x-4)^{n+1}}\right]$
5. $\dfrac{1}{24} 5^n \sin\left(\dfrac{n\pi}{2} + 5\theta\right) - 3^n \sin\left(\dfrac{n\pi}{2} + 3\theta\right) - 2\sin\left(\dfrac{n\pi}{2} + \theta\right)$
6. $(45)^n e^{3x} \cos\left(\dfrac{\pi}{2} + 6x\right) + (-1)^n (n-1)! \dfrac{1}{x^n}$

Theorem: Leibnitz's Theorem

Statement: If $y = uv$ is the product of two differentiable functions u and v of x, then the n^{th} derivative of y is

$$y_n = uv_n + {}^nC_1 u_1 v_{n-1} + {}^nC_2 u_2 v_{n-2} + \cdots + {}^nC_r u_r v_{n-r} + \cdots + u_n v$$

where u_r and v_r are the r^{th} derivatives of u and v respectively.
Proof: Given $y = uv$.

Let

$$P(n): \quad y_n = uv_n + {}^nC_1 u_1 v_{n-1} + {}^nC_2 u_2 v_{n-2} + \cdots + {}^nC_r u_r v_{n-r} + \cdots + u_n v \qquad (1)$$

We prove the theorem by induction on n.

Basic step:

$$P(1): \quad y_1 = uv_1 + u_1 v$$

Which is true from product rule. $\therefore P(1)$ is true.

Inductive Step

We assume that the theorem is true for $n = k$ (> 1). i.e., $p(k)$ is true $\implies y_k = uv_k + {}^kC_1 u_1 v_{k-1} + \cdots + {}^kC_r u_r v_{k-r} + \cdots + u_k v$ is true.

To prove $p(k+1)$ is true. That is to prove:

$$y_{k+1} = uv_{k+1} + {}^{k+1}C_1 u_1 v_k + {}^{k+1}C_2 u_2 v_{k-1} + \cdots + {}^{k+1}C_r u_r v_{k-r+1}$$

Differentiating (2) w.r.t. x, we get:

$$y_{k+1} = uv_{k+1} + u_1 v_k + {}^kC_1[u_1 v_k + u_2 v_{k-1}] + \cdots + {}^kC_r[u_r v_{k-r+1} + u_{r+1} v_{k-r}] + \cdots + u_k v_1 + u_{k+1} v$$

$$\implies y_{k+1} = uv_{k+1} + {}^{k+1}C_1 u_1 v_k + {}^{k+1}C_2 u_2 v_{k-1} + \cdots + {}^{k+1}C_r u_r v_{k-r+1} + \cdots + u_{k+1} v$$

Since ${}^kC_r + {}^kC_{r-1} = {}^{k+1}C_r$.
$\therefore \quad P(k+1)$ is true.

Thus, $P(k)$ is true $\implies P(k+1)$ is true. $\therefore$ By induction $p(n)$ is true for all values of $n \in \mathbb{N}$. Hence, the theorem is true for all values of $n \in \mathbb{N}$.

$$y_n = uv_n + {}^nC_1 u_1 v_{n-1} + {}^nC_2 u_2 v_{n-2} + \cdots + {}^nC_r u_r v_{n-r} + \cdots + u_n v$$

Worked Examples

Example 1

If $y = \cos(m \sin^{-1} x)$, then prove that $(1 - x^2)y'' - xy' + m^2 y = 0$. Hence prove that $(1 - x^2)y_{n+2} - (2n+1)xy_{n+1} + (m^2 - n^2)y_n = 0$.

Solution.
Given: $y = \cos(m \sin^{-1} x) \implies \cos^{-1} y = m \sin^{-1} x$
Differentiating (1) w.r.t. x, we get:

$$\frac{-1}{\sqrt{1-y^2}} \cdot \frac{dy}{dx} = \frac{m}{\sqrt{1-x^2}} \implies \frac{-1}{\sqrt{1-y^2}} y_1 = \frac{m}{\sqrt{1-x^2}}$$

Squaring,

$$\frac{y_1^2}{1-y^2} = \frac{m^2}{1-x^2} \implies (1-x^2)y_1^2 = m^2(1-y^2)$$

Differentiating w.r.t. x, we get:

$$(1-x^2)2y_1y_2 - 2xy_1^2 = -2m^2yy_1$$

$$\implies (1-x^2)y_2 - xy_1 + m^2y = 0$$

Dividing by $2y_1$ on both sides,

$$(1-x^2)y_2 - xy_1 = -m^2y \implies (1-x^2)y_2 - xy_1 + m^2y = 0 \tag{2}$$

Differentiating (2) w.r.t. x, n times by Leibnitz's theorem, we get

$$(1-x^2)y_{n+2} + {}^1C_1(-2x)y_{n+1} + {}^2C_2(-2)y_n - [xy_{n+1} + {}^1C_1y_n] + m^2y = 0$$

$$\implies (1-x^2)y_{n+2} - 2nxy_{n+1} - \frac{n(n-1)}{1\cdot 2}2y_n - xy_{n+1} - ny_n + m^2y = 0$$

$$\implies (1-x^2)y_{n+2} - (2n+1)xy_{n+1} + [m^2 - n(n-1) - n]y_n = 0$$

$$\implies (1-x^2)y_{n+2} - (2n+1)xy_{n+1} + (m^2-n^2)y_n = 0$$

Example 2

If $y = a\cos(\log x) + b\sin(\log x)$, then show that $x^2y_{n+2} + (2n+1)xy_{n+1} + (n^2+1)y_n = 0$.
Solution.
Given $y = a\cos(\log x) + b\sin(\log x)$
Differentiating (1) w.r.t. x, we get:

$$y_1 = a(-\sin(\log x)) \cdot \frac{1}{x} + b\cos(\log x) \cdot \frac{1}{x} = \frac{-1}{x}[a\sin(\log x) - b\cos(\log x)]$$

$$\implies xy_1 = -a\sin(\log x) + b\cos(\log x)$$

Again differentiating w.r.t. x, we get

$$xy_2 + y_1 = -a\cos(\log x) \cdot \frac{1}{x} - b(-\sin(\log x)) \cdot \frac{1}{x}$$

$$= \frac{-1}{x}[a\cos(\log x) + b\cos(\log x)] = \frac{-y}{x}$$

$$\implies x^2y_2 + xy_1 = -y \implies x^2y_2 + xy_1 + y = 0 \tag{2}$$

Differentiating (2) w.r.t. x, n times by Leibnitz's theorem, we get:

$$x^2y_{n+2} + {}^1C_1 2xy_{n+1} + {}^2C_2 2y_n + xy_{n+1} + {}^1C_1 y_n = 0$$

$$\implies x^2y_{n+2} + 2nxy_{n+1} + \frac{n(n-1)}{1\cdot 2}2y_n + xy_{n+1} + ny_n = 0$$

$$\implies x^2y_{n+2} + (2n+1)xy_{n+1} + [n(n-1) + n + 1]y_n = 0$$

$$\implies x^2y_{n+2} + (2n+1)xy_{n+1} + (n^2+1)y_n = 0$$

Example 3

If $y = \left(x + \sqrt{1+x^2}\right)^m$, then prove that
$(1+x^2)y_{n+2} + (2n+1)xy_{n+1} + (n^2 - m^2)y_n = 0.$

Solution.

Given

$$y = \left(x + \sqrt{1+x^2}\right)^m \tag{1}$$

Differentiating (1) w.r.t. x, we get:

$$y_1 = m\left(x + \sqrt{1+x^2}\right)^{m-1} \cdot \left[1 + \frac{1}{2\sqrt{1+x^2}} \cdot 2x\right]$$

$$= m\left(x + \sqrt{1+x^2}\right)^{m-1} \cdot \left(1 + \frac{x}{\sqrt{1+x^2}}\right)$$

$$= m\left(x + \sqrt{1+x^2}\right)^{m-1} \cdot \frac{x + \sqrt{1+x^2}}{\sqrt{1+x^2}}$$

$$= m\frac{\left(x + \sqrt{1+x^2}\right)^m}{\sqrt{1+x^2}} = \frac{my}{\sqrt{1+x^2}}$$

Squaring:

$$y_1^2 = \frac{m^2y^2}{1+x^2} \implies (1+x^2)y_1^2 = m^2y^2$$

Again differentiating w.r.t. x:

$$(1+x^2)2y_1y_2 + 2xy_1^2 = 2m^2yy_1 \implies (1+x^2)y_2 + xy_1 = m^2\frac{y}{1}$$

$$\implies (1+x^2)y_2 + xy_1 - m^2y = 0 \tag{2}$$

Differentiating (2) w.r.t. x, n times by Leibnitz's theorem, we get:

$$(1+x^2)y_{n+2} + {}^1C_1 2xy_{n+1} + {}^2C_2 2y_n + xy_{n+1} + {}^1C_1 y_n - m^2y = 0$$

$$\implies (1+x^2)y_{n+2} + 2nxy_{n+1} + \frac{n(n-1)}{2}2y_n + xy_{n+1} + ny_n - m^2y = 0$$

$$\implies (1+x^2)y_{n+2} + (2n+1)xy_{n+1} + (n(n-1) + n - m^2)y_n = 0$$

$$\implies (1+x^2)y_{n+2} + (2n+1)xy_{n+1} + (n^2 - m^2)y_n = 0$$

Problems to Find $y_n(0)$

Example 4

If $y = (\sin^{-1} x)^2$, then prove that
$(1-x^2)y_{n+2} - xy_{n+1} - 2 = 0$. Hence show that
$(1-x^2)y_{n+2} - (2n+1)xy_{n+1} - n^2y_n = 0$ and find $y_n(0)$.

Solution.
Given

$$y = (\sin^{-1} x)^2 \tag{1}$$

Differentiating (1) w.r.t. x, we get

$$y_1 = \frac{2\sin^{-1} x}{\sqrt{1-x^2}} \tag{2}$$

Squaring,

$$y_1^2 = \frac{4(\sin^{-1} x)^2}{(1-x^2)} \implies (1-x^2)y_1^2 = 4y \tag{3}$$

Differentiating (3) w.r.t. x, we get

$$(1-x^2)2y_1y_2 - 2xy_1^2 = 4y_1 \implies (1-x^2)y_2 - xy_1 = 2 \tag{4}$$

[Dividing by $2y_1$]

Differentiating (4) w.r.t. x, n times using Leibnitz's theorem, we get

$$(1-x^2)y_{n+2} + nC_1(-2x)y_{n+1} + nC_2(-2)y_n + xy_{n+1} + nC_1y_n = 0$$

$$\implies (1-x^2)y_{n+2} - 2nxy_{n+1} - n(n-1)y_n + xy_{n+1} + ny_n = 0$$

$$\implies (1-x^2)y_{n+2} + (2n+1)xy_{n+1} - [n(n-1) - n]y_n = 0$$

$$\implies (1-x^2)y_{n+2} + (2n+1)xy_{n+1} - n^2y_n = 0 \tag{5}$$

To find $y_n(0)$

Put $x = 0$ in (5), we get

$$y_{n+2}(0) - n^2 y_n(0) = 0 \implies y_{n+2}(0) = n^2 y_n(0)$$

Put $x = 0$ in (2), we get

$$y_1(0) = \frac{2\sin^{-1} 0}{\sqrt{1-0}} = 0$$

Put $x = 0$ in (4), we get

$$y_2(0) - 2 = 0 \implies y_2(0) = 2$$

We have

$$y_{n+2}(0) = n^2 y_n(0) \implies y_{n+2}(0) = n^2 y_n(0)$$

If n is odd, $n = 1, 3, 5, 7, \ldots$ then we have

$$y_1(0) = y_1(0) = 0, \quad y_3(0) = 3^2 y_1(0) = 0, \quad y_5(0) = 5^2 y_3(0) = 0$$

$$\therefore \text{If } n \text{ is odd}, \quad y_n(0) = 0.$$

If n is even, $n = 2, 4, 6, 8, \ldots$ then we have

$$y_2(0) = 2, \quad y_4(0) = 2^2 \cdot y_2(0) = 2^2 \cdot 2,$$

$$y_6(0) = 4^2 \cdot y_4(0) = 2^2 \cdot 4^2, \quad y_8(0) = 6^2 \cdot y_6(0) = 2 \cdot 2^2 \cdot 4^2 \cdot 6^2 \text{ and so on.}$$

$$\therefore \text{If } n \text{ is even}, \quad y_n(0) = 2 \cdot 2^2 \cdot 4^2 \cdot 6^2 \cdots (n-2)^2$$

$$\therefore y_n(0) = \begin{cases} 2 \cdot 2^2 \cdot 4^2 \cdot 6^2 \cdots (n-2)^2, & \text{if } n \text{ is even} \\ 0, & \text{if } n \text{ is odd} \end{cases}$$

Example 5

Find the value of the n^{th} derivative of $e^{a\sin^{-1} x}$ for $x = 0$.

Solution. Let

$$y = e^{a\sin^{-1} x} \tag{1}$$

Differentiating w.r.t x, we get

$$y_1 = e^{a\sin^{-1} x} \cdot \frac{a}{\sqrt{1-x^2}} = \frac{ay}{\sqrt{1-x^2}} \tag{2}$$

$$\implies (1-x^2)y_1^2 = a^2 y^2$$

Again differentiating w.r.t x, we get

$$(1-x^2)2y_1y_2 - 2xy_1^2 = a^2 \cdot 2yy_1$$

$$\implies (1-x^2)y_2 - xy_1 = a^2y \tag{3}$$

Differentiating (3) w.r.t x, n times using Leibnitz's theorem, we get

$$(1-x^2)y_{n+2} + {}^nC_1(-2x)y_{n+1} + {}^nC_2(-2)y_n - [xy_{n+1} + {}^nC_1y_n] - a^2y_n = 0$$

$$\implies (1-x^2)y_{n+2} - 2nxy_{n+1} - \frac{n(n-1)}{1\cdot 2}2y_n - xy_{n+1} - ny_n - a^2y_n = 0$$

$$\implies (1-x^2)y_{n+2} - (2n+1)xy_{n+1} - [n(n-1) + n + a^2]y_n = 0$$

$$\implies (1-x^2)y_{n+2} - (2n+1)xy_{n+1} - (n^2+a^2)y_n = 0 \tag{4}$$

Put $x = 0$ in (4), then

$$y_{n+2}(0) - (n^2+a^2)y_n(0) = 0$$

$$\implies y_{n+2}(0) = (n^2+a^2)y_n(0) \tag{5}$$

Put $x = 0$ in (1), then $\quad y(0) = 1$
Put $x = 0$ in (2), then $\quad y_1(0) = a$
Put $x = 0$ in (3), then $\quad y_2(0) = a^2y(0) \implies y_2(0) = a^2 \cdot 1 = a^2$
Putting $n = 1, 3, 5, 7, \ldots$ in (5) we get

$$y_3(0) = (1^2+a^2)y_1(0) = (1^2+a^2)a = a(1^2+a^2)$$

$$y_5(0) = (3^2+a^2)y_3(0) = a(1^2+a^2)(3^2+a^2)$$

$$y_7(0) = (5^2+a^2)y_5(0) = a(1^2+a^2)(3^2+a^2)(5^2+a^2)$$

and so on.

If n is odd,

$$y_n(0) = a(1^2+a^2)(3^2+a^2)\ldots[(n-2)^2+a^2]$$

Putting $n = 2, 4, 6, 8, \ldots$ in (5), we get

$$y_4(0) = (2^2+a^2)y_2(0) = a^2(2^2+a^2)$$

$$y_6(0) = (4^2+a^2)y_4(0) = a^2(2^2+a^2)(4^2+a^2)$$

$$y_8(0) = (6^2+a^2)y_6(0) = a^2(2^2+a^2)(4^2+a^2)(6^2+a^2)$$

and so on.

If n is even,

$$y_n(0) = a^2(2^2+a^2)(4^2+a^2)(6^2+a^2)\cdots[(n-2)^2+a^2].$$

$$y_n(0) = \begin{cases} a^2(2^2+a^2)(4^2+a^2)\cdots[(n-2)^2+a^2], & \text{if } n \text{ is even} \\ a(1^2+a^2)(3^2+a^2)\cdots[(n-2)^2+a^2], & \text{if } n \text{ is odd} \end{cases}$$

Example 6

Considering $x^n x^n = x^{2n}$ and using Leibnitz's theorem, prove that

$$1+\frac{n^2}{1^2}+\frac{n^2(n-1)^2}{1^2\cdot 2^2}+\frac{n^2(n-1)^2(n-2)^2}{1^2\cdot 2^2\cdot 3^2}+\cdots=\frac{(2n)!}{(n!)^2}$$

Solution.
Let

$$y = x^{2n}$$

Differentiating w.r.t x, n times we get

$$\begin{aligned} y_n &= 2n(2n-1)(2n-2)\cdots(2n-n+1)x^{2n-n} \\ &= 2n(2n-1)(2n-2)\cdots(n+1)x^n \\ &= \frac{2n(2n-1)(2n-2)\cdots(n+1)n\cdot 2\cdot 1}{1\cdot 2\cdots n}x^n \\ \implies y_n &= \frac{(2n)!}{(n)!}x^n \end{aligned} \tag{1}$$

Also

$$y = x^n\cdot x^n \implies u = x^n, \quad v = x^n \implies y = uv$$

By Leibnitz's theorem,

$$y_n = uv_n + nC_1u_1v_{n-1} + nC_2u_2v_{n-2} + \cdots + nC_ru_rv_{n-r} + \cdots + u_nv$$

$$u_1 = nx^{n-1}, \quad u_2 = n(n-1)x^{n-2}$$

$$u_3 = n(n-1)(n-2)x^{n-3}$$

$$u_{n-2} = n(n-1)(n-2)\cdots[n-(n-3)]x^{n-(n-2)}$$

$$= n(n-1)(n-2)\cdots 4\cdot 3\cdot 2\cdot 1\cdot x^2 = \frac{n!}{1\cdot 2}x^2$$

$$u_{n-1} = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1\cdot x = n!x$$

$$u_n = n!$$

Similarly, if $v = x^n$, then $v_n = n!$

$$\therefore \quad y_n = x^n n! + nC_1 n! x^{n-1}\cdot n!x + nC_2 n(n-1)x^{n-2}\cdot\frac{n!}{1\cdot 2}x^2 + \cdots$$

$$+nC_3 n(n-1)(n-2)x^{n-3}\cdot\frac{n!}{1\cdot 2\cdot 3}x^3 + \cdots + n!x^n$$

$$y_n = x^n n! + nC_1 n! x^{n-1}\cdot n!x + nC_2 n(n-1)x^{n-2}\cdot\frac{n!}{1\cdot 2}x^2 + \cdots$$

$$+nC_3 n(n-1)(n-2)x^{n-3}\cdot\frac{n!}{1\cdot2\cdot3}x^3+\cdots+n!x^n$$
$$=n!x^n\left[1+\frac{n^2}{1^2}+\frac{n^2(n-1)^2}{1^2\cdot2^2}+\frac{n^2(n-1)^2(n-2)^2}{1^2\cdot2^2\cdot3^2}+\cdots\right]$$

From (1) and (2), we get

$$n!x^n\left[1+\frac{n^2}{1^2}+\frac{n^2(n-1)^2}{1^2\cdot2^2}+\frac{n^2(n-1)^2(n-2)^2}{1^2\cdot2^2\cdot3^2}+\cdots\right]=\frac{(2n)!}{n!}x^n$$
$$\implies 1+\frac{n^2}{1^2}+\frac{n^2(n-1)^2}{1^2\cdot2^2}+\frac{n^2(n-1)^2(n-2)^2}{1^2\cdot2^2\cdot3^2}+\cdots=\frac{(2n)!}{(n!)^2}$$

Exercise: Successive Differentiation

1. If $y=\left(x+\sqrt{1+x^2}\right)^m$, then prove that $(1+x^2)y_{n+2}+(2n+1)xy_{n+1}+(n^2-m^2)y_n=0$.
2. If $y=(1-x)^n e^{-ax}$, then prove that $(1-x)\frac{dy}{dx}=axy$ and $(1-x)y_{n+1}-(n+ax)y_n-nay=0$.
3. If $y=\tan^{-1}x$, then prove that $(1+x^2)y_{n+2}+(2n+1)xy_{n+1}+n(n+1)y_n=0$.
4. If $y=(\sin^{-1}x)^2$, then prove that $(1-x^2)y_{n+2}-(2n+1)xy_{n+1}-n^2y_n=0$.
5. If $y=(\cos^{-1}x)^2$, then prove that $(1-x^2)y_{n+2}-(2n+1)xy_{n+1}-n^2y_n=0$.
6. If $y=\sin^{-1}x$, then prove that
(i) $(1-x^2)y_2-xy_1=0$
(ii) $(1-x^2)y_{n+2}-(2n+1)xy_{n+1}-n^2y_n=0$.
7. If $\cos^{-1}\left(\frac{y}{b}\right)=\log\left(\frac{x}{n}\right)$, then prove that $x^2y_{n+2}+(2n+1)xy_{n+1}+2n^2y_n=0$.
8. If $y=\left[\log\left(x+\sqrt{1+x^2}\right)\right]^{\frac{1}{2}}$, then prove that $(1+x^2)y_{n+2}+(2n+1)xy_{n+1}+n^2y_n=0$.
9. If $y=(x^2+y^2)^{-m/2}\cdot 2x$, then prove that $(x^2-1)y_{n+2}+(2n+1)xy_{n+1}+(n^2-m^2)y_n=0$.
10. If $x=\sin t, y=\sin pt$, then prove that
(i) $(1-x^2)y''-xy'+p^2y=0$
(ii) $(1-x^2)y_{n+2}-(2n+1)xy_{n+1}-(n^2-p^2)y_n=0$.

Rolle's Theorem

Theorem. Let f be continuous on $[a,b]$ and differentiable on (a,b). If $f(a)=f(b)$, then there is a number $c\in(a,b)$ such that $f'(c)=0$.

Geometrical Interpretation:

If $f(x)$ satisfies all the conditions of Rolle's theorem, then there exists a point $c\in(a,b)$ such that the tangent at c is parallel to the x-axis (see Fig. 2.1).

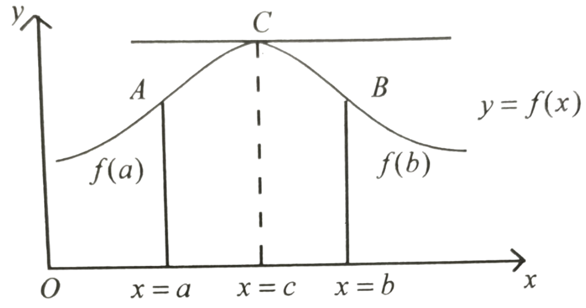


Figure 2.1:

Example 1. Verify Rolle's theorem for $f(x) = x^2 - 4x + 3$ on $[1, 3]$.

Solution: Given the function $f(x) = 2x^3 + x^2 - 4x - 3$ on the interval $[-\sqrt{2}, \sqrt{2}]$. Since $f(x)$ is a polynomial, it is continuous on the closed interval $[-\sqrt{2}, \sqrt{2}]$ and differentiable on the open interval $(-\sqrt{2}, \sqrt{2})$.
Now $f(-\sqrt{2}) = 2(-\sqrt{2})^3 + (-\sqrt{2})^2 - 4(-\sqrt{2}) - 3 = -1$, and $f(\sqrt{2}) = 2(\sqrt{2})^3 + (\sqrt{2})^2 - 4(\sqrt{2}) - 3 = -1$.
Hence, $f(-\sqrt{2}) = f(\sqrt{2})$.
Again, $f'(x) = 6x^2 + 2x - 4$. Let $f'(c) = 0$, So,

$$6c^2 + 2c - 4 = 0. \implies 3c^2 + c - 2 = 0.$$

Therefore,

$$c = \frac{-1 \pm \sqrt{1^2 - 4 \times 3 \times (-2)}}{2 \times 3} = \frac{-1 \pm \sqrt{1 + 24}}{6} = \frac{-1 \pm 5}{6}.$$

Thus,

$$c = \frac{4}{6} = \frac{2}{3} \quad \text{or} \quad c = \frac{-6}{6} = -1.$$

Both $c = -1$ and $c = \frac{2}{3}$ lie within $(-\sqrt{2}, \sqrt{2})$.
Hence, all conditions of Rolle's theorem are satisfied, and there exist points

$$c = -1 \quad \text{and} \quad c = \frac{2}{3}$$

such that

$$f'(c) = 0$$

on the interval $(-\sqrt{2}, \sqrt{2})$.

Example 2. Verify Rolle's theorem for $f(x) = \cos x$ in $[-\pi/2, \pi/2]$.

Solution: Given the function $f(x) = \cos x$ on the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The function $f(x) = \cos x$ is continuous on the closed interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and differentiable on the open interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ since cosine is a standard trigonometric function with these properties. Now $f\left(-\frac{\pi}{2}\right) = \cos\left(-\frac{\pi}{2}\right) = 0$, and $f\left(\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right) = 0$.
Thus,

$$f\left(-\frac{\pi}{2}\right) = f\left(\frac{\pi}{2}\right).$$

Again,

$$f'(x) = -\sin x.$$

Let $f'(c) = 0$, so $-\sin c = 0 \implies \sin c = 0$. The solutions for $\sin c = 0$ in the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ is $c = 0$.
Since all conditions of Rolle's theorem are satisfied and there exists $c = 0 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ such that $f'(c) = 0$, Rolle's theorem holds for $f(x) = \cos x$ on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Exercise 1. Show that Rolle's theorem is not applicable to $f(x) = \tan x$ in $[0, \pi]$, although $f(0) = f(\pi)$.
Exercise 2. Verify Rolle's theorem for $f(x) = x^3 - 6x^2 + 11x - 6$ in $1 \leq x \leq 3$.

The Mean Value Theorem (Lagrange's Theorem)

Theorem. If f is continuous on $[a, b]$ and differentiable on (a, b), then there exists a number $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Remark. Rolle's theorem is a special case of the Lagrange's MVT. Indeed, if $f(a) = f(b)$, then, by MVT,

$$f'(c) = \frac{f(b) - f(a)}{b - a} = \frac{0}{b - a} = 0$$

for some $c \in (a, b)$.

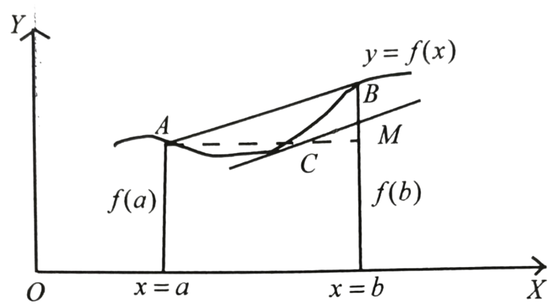


Figure 2.2:

Geometrical Interpretation:

If $f(x)$ satisfies all the conditions of Lagrange's MVT theorem, then there exists a point $c \in (a, b)$ such that the tangent at c is parallel to the straight line passing through the points $A(a, f(a))$ and $B(b, f(b))$ as shown in Fig. 2.2.

Another form of Lagrange's MVT

Let $f(x)$ be a function defined on a finite closed interval $[a, a+h]$ such that

(i) $f(x)$ is continuous on $[a, a+h]$,

(ii) $f(x)$ is differentiable on $(a, a+h)$,

Then $f(a+h) = f(a) + hf'(a+\theta h), \quad 0 < \theta < 1$. Now, in the interval $[0, h]$, we get

$$f(h) = f(0) + hf'(\theta h), \quad 0 < \theta < 1.$$

Example 3. Verify Lagrange's MVT for $f(x) = x(x-1)(x-2)$ in $0 \le x \le 1/2$.

Solution: Since $f(x)$ is a polynomial, hence continuous on $[0, \frac{1}{2}]$ and differentiable on $(0, \frac{1}{2})$.

Now, $f(0) = 0 \cdot (-1) \cdot (-2) = 0$, and $f\left(\frac{1}{2}\right) = \frac{1}{2}\left(\frac{1}{2} - 1\right)\left(\frac{1}{2} - 2\right) = \frac{3}{8}$. Therefore, $f(0) \neq f(1/2)$.

In addition, $\frac{f\left(\frac{1}{2}\right) - f(0)}{\frac{1}{2} - 0} = \frac{\frac{3}{8} - 0}{\frac{1}{2}} = \frac{3}{8} \times \frac{2}{1} = \frac{3}{4} = 0.75$.

Here $f'(x) = 3x^2 - 6x + 2$.

Applying Lagrange's MVT, there exists $c \in (0, \frac{1}{2})$ such that

$$f'(c) = \frac{3}{4} \implies 3c^2 - 6c + 2 = \frac{3}{4},$$

which gives, $3c^2 - 6c + 2 - \frac{3}{4} = 0 \implies 3c^2 - 6c + \frac{5}{4} = 0$. Multiply entire equation by 4 to clear fractions:

$$12c^2 - 24c + 5 = 0.$$

Therefore,

$$c = \frac{24 \pm \sqrt{(-24)^2 - 4 \cdot 12 \cdot 5}}{2 \cdot 12} = \frac{24 \pm \sqrt{576 - 240}}{24} = \frac{24 \pm \sqrt{336}}{24}.$$

Simplify $\sqrt{336} = \sqrt{16 \times 21} = 4\sqrt{21}$:

$$c = \frac{24 \pm 4\sqrt{21}}{24} = 1 \pm \frac{\sqrt{21}}{6}.$$

Check which root lies in $(0, \frac{1}{2})$:

$$1 - \frac{\sqrt{21}}{6} \approx 1 - 0.76 = 0.24 \in (0, 0.5),$$

$$1 + \frac{\sqrt{21}}{6} > 1,$$

so the valid c is

$$c = 1 - \frac{\sqrt{21}}{6} \approx 0.24.$$

Conclusion: There exists a $c \in (0, \frac{1}{2})$ such that

$$f'(c) = \frac{f\left(\frac{1}{2}\right) - f(0)}{\frac{1}{2} - 0},$$

verifying Lagrange's MVT for the function $f(x) = x(x-1)(x-2)$ in the interval $[0, \frac{1}{2}]$.

Example 4. Use MVT to prove that the following inequalities:

(i) $0 < \frac{1}{x} \log \frac{e^x - 1}{x} < 1$.

(ii) $\frac{x}{1+x} < \log(1+x) < x$ if $x > 0$.

Solutions:

(i) Let $f(x) = e^x$. Then from MVT $f(x) = f(0) + xf'(\theta x),\ 0 < \theta < 1$, gives $e^x = e^0 + xe^{\theta x}$.

$$\therefore e^x - 1 = xe^{\theta x} \implies e^{\theta x} = \frac{e^x - 1}{x}, \implies \theta = \frac{1}{x} \log \frac{e^x - 1}{x}.$$

$$0 < \frac{1}{x} \log \frac{e^x - 1}{x} < 1. \ \ [\because 0 < \theta < 1.]$$

(ii) Let $f(x) = \log(1+x)$. Then from MVT,

$$f(x) = f(0) + xf'(\theta x),\ 0 < \theta < 1,$$

we have $\log(1+x) = \log 1 + x\frac{1}{1+\theta x}$,

$$\therefore \log(1+x) = \frac{x}{1+\theta x},$$

$$\because 0 < \theta < 1, \text{ and } x > 0 \implies 0 < \theta x < x.$$

or

$$1 < 1 + \theta x < 1 + x \implies 1 > \frac{1}{1+\theta x} > \frac{1}{1+x}$$

or

$$\frac{x}{1+x} < \frac{x}{1+\theta x} < x \implies \frac{x}{1+x} < \log(1+x) < x.$$

Cauchy's MVT: Let f and g be two functions defined on $[a, b]$ such that

(i) f and g are continuous on $[a, b]$,

(ii) f and g are differentiable on (a, b), and

(iii) $g'(x) \neq 0$ for any x in (a, b).

Then there exist atleast one $c \in (a, b)$, such that

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Note: If $g(x) = x$, then Cauchy's MVT is equivalent to Lagrange's MVT.

Exercise 3. In Cauchy's MVT, if $f(x) = e^x$ and $g(x) = e^{-x}$, show that θ is independent of both x and h and is equal to $1/2$.

Taylor's theorem: Let $f(x)$ be a function defined on $[a, b]$ such that

(i) the $(n-1)$th derivative $f^{(n-1)}$ is continuous on $[a, b]$, and

(ii) the nth derivative $f^{(n)}$ exists on (a, b).

Then there exist at least one value c, $a < c < b$, such that

$$f(b) = f(a)+(b-a)f'(a)+\frac{1}{2!}(b-a)^2 f''(a)+\cdots+\frac{1}{(n-1)!}(b-a)^{n-1}f^{(n-1)}(a)+\frac{1}{n!}(b-a)^n f^{(n)}(c).$$

Note: If we replace the interval $[a, b]$ by $[a, a+h]$, $h > 0$. Then there exist atleast one number θ, $0 < \theta < 1$, such that

$$f(a+h) = f(a) + hf'(a) + \frac{1}{2!}h^2 f''(a) + \cdots + \frac{1}{(n-1)!}h^{n-1}f^{(n-1)}(a) + \frac{1}{n!}h^n f^{(n)}(a+\theta h),$$

The last term ($R_n = \frac{1}{n!}h^n f^{(n)}(a+\theta h)$) in the above expression is called the Lagrange's form of remainder.

Maclaurin's theorem: Let $f(x)$ be a function defined on $[-h, h]$ such that

(i) the $(n-1)$th derivative $f^{(n-1)}$ is continuous on $[-h, h]$, and

(ii) the nth derivative $f^{(n)}$ exists on $(-h, h)$.

Then there exists at least one value θ, $0 < \theta < 1$, such that

$$f(x) = f(0) + xf'(0) + \frac{1}{2!}x^2 f''(0) + \cdots + \frac{1}{(n-1)!}x^{n-1} f^{(n-1)}(0) + \frac{1}{n!}x^n f^{(n)}(\theta x).$$

Maclaurin's infinite series: Suppose that $f(x)$ has derivatives of every order in $[-h, h]$, for some $h > 0$. Then for $x \in [-h, h]$

$$f(x) = f(0) + xf'(0) + \frac{1}{2!}x^2 f''(0) + \cdots + \frac{1}{(n-1)!}x^{n-1} f^{(n-1)}(0) + \frac{1}{n!}x^n f^{(n)}(0) + \cdots \infty.$$

or,

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

if $\lim_{n\to\infty} R_n \to 0$, where R_n is the remainder after the n terms.

Example 5. Expand $f(x) = \sin x$ in power of x using Maclaurin's series.

Solution: Take $f(x) = \sin x$. Compute derivatives at $x = 0$:

$$\begin{aligned} f(x) &= \sin x, & f(0) &= 0, \\ f'(x) &= \cos x, & f'(0) &= 1, \\ f''(x) &= -\sin x, & f''(0) &= 0, \\ f^{(3)}(x) &= -\cos x, & f^{(3)}(0) &= -1, \\ f^{(4)}(x) &= \sin x, & f^{(4)}(0) &= 0, \end{aligned}$$

and the pattern repeats every four derivatives.

Only the odd derivatives at 0 are nonzero, and they alternate in sign. Substituting into the Maclaurin formula yields

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

Writing out the first terms,

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots.$$

2.8 Improper Integrals

An integral is said to be an **improper integral** if one or both of the following conditions hold:

(i) The interval of integration is infinite. For example,

$$\int_a^\infty f(x)\,dx \quad \text{or} \quad \int_{-\infty}^\infty f(x)\,dx.$$

(ii) The integrand becomes unbounded (i.e., has a discontinuity) at one or more points in the interval of integration. For example,

$$\int_0^1 \frac{1}{x}\,dx \quad \text{or} \quad \int_{-1}^1 \frac{1}{\sqrt{1-x^2}}\,dx.$$

Formal Definition:

- If $f(x)$ is continuous on $[a, \infty)$, then

$$\int_a^\infty f(x)\,dx = \lim_{t\to\infty} \int_a^t f(x)\,dx,$$

  provided the limit exists and is finite.

- If $f(x)$ is continuous on $(a, b]$ but has a discontinuity at $x = a$, then

$$\int_a^b f(x)\,dx = \lim_{\epsilon\to 0^+} \int_{a+\epsilon}^b f(x)\,dx,$$

  provided the limit exists and is finite.

If the limit exists, the improper integral is said to be **convergent**; otherwise, it is **divergent**. They are vital in many areas of science and engineering, including probability theory, signal processing, and differential equations.

2.9 Applications of Improper Integrals in Different Branches of Engineering

Improper integrals play an essential role in mathematical modeling across various engineering branches. They are used to analyze phenomena extending to infinity, evaluate singular functions, and compute integrals involving probability distributions and transforms.

Applications

- **Mechanical Engineering (ME):** Improper integrals arise in Fourier analysis for vibration problems and in heat conduction in infinite rods and plates.

- **Chemical Engineering (CHE):** Used in modeling reaction kinetics, mass transfer, and diffusion processes over semi-infinite domains.

- **Civil Engineering (CE):** Applied in hydrology (rainfall distribution, flood probability) and in computing deflections in beams where load extends infinitely.

- **Electrical Engineering (EE):** Essential in transient analysis of circuits using Laplace and Fourier transforms, both defined as improper integrals.

- **Applied Electronics & Instrumentation Engineering (AEIE):** Signal filtering and frequency response analysis require Fourier integrals over infinite time domains.

- **Biotechnology (BT):** Growth models, population kinetics, and enzyme activity studies often involve integrals over infinite time or concentration ranges.

- **Food Technology (FT):** Heat penetration in food preservation and diffusion in packaging materials can be modeled using improper integrals.

- **Agriculture Engineering (AE):** Soil-water infiltration models and crop growth analysis use improper integrals for probability and diffusion-related models.

- **Electronics & Communication Engineering (ECE):** Fourier transform and Laplace transform methods (defined via improper integrals) are central in signal transmission and communication theory.

- **Computer Science & Engineering (CSE):** Improper integrals are used in algorithm analysis (e.g., evaluating continuous approximations of discrete sums) and numerical integration for infinite domains.

- **Cyber Security:** Probability density functions in cryptography and security models often require improper integrals for normalization.

- **Data Science:** Many probability distributions (Gaussian, exponential, Cauchy) require improper integrals in evaluating expectations, variances, and likelihoods.

- **Artificial Intelligence & Machine Learning (AI–ML):** Integrals over infinite domains are used in Bayesian inference, kernel methods, and probabilistic models.

- **Information Technology (IT):** Involves improper integrals in performance analysis of networks, error probability estimation, and stochastic process modeling.

2.2 Convergence Criteria

- Comparison Test
- Limit Comparison Test
- Absolute Convergence
- p-Test: $\int_1^\infty \frac{1}{x^p} dx$ converges if and only if $p > 1$

2.10 Types of Improper Integrals

Improper integrals are of two major types:

2.10.1 Type I: Infinite Limits

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx$$

$$\int_{-\infty}^b f(x)\,dx = \lim_{a\to-\infty} \int_a^b f(x)\,dx$$

$$\int_{-\infty}^\infty f(x)\,dx = \int_{-\infty}^c f(x)\,dx + \int_c^\infty f(x)\,dx$$

2.10.2 Type II: Discontinuous Integrands

$$\int_a^b f(x)\,dx \quad \text{where } f(x) \to \infty \text{ as } x \to a^+ \text{ or } x \to b^-$$

For example:

$$\int_0^1 \frac{1}{x^p} dx \quad \text{diverges for } p \geq 1,\ \text{converges for } p < 1$$

2.11 Convergence of Improper Integrals

The convergence of an improper integral depends on the behavior of the function near the problematic point (either infinity or a singularity).

Key Rule

If the limit defining the improper integral exists and is finite, then the integral converges. Otherwise, it diverges.

2.12 Tests for Convergence

2.12.1 1. p-test

For $p > 0$,

$$\int_1^\infty \frac{1}{x^p} dx \begin{cases} \text{converges} & \text{if } p > 1 \\ \text{diverges} & \text{if } p \leq 1 \end{cases}$$

2.12.2 2. Comparison Test

Let $0 \leq f(x) \leq g(x)$ for $x \geq a$.

- If $\int_a^\infty g(x)\,dx$ converges, then so does $\int_a^\infty f(x)\,dx$
- If $\int_a^\infty f(x)\,dx$ diverges, then so does $\int_a^\infty g(x)\,dx$

2.12.3 3. Limit Comparison Test

If $\lim_{x\to\infty} \frac{f(x)}{g(x)} = c > 0$, then $\int_a^\infty f(x)dx$ and $\int_a^\infty g(x)dx$ either both converge or both diverge.

2.13 Absolute and Conditional Convergence

- An integral $\int_a^\infty f(x)dx$ is said to be absolutely convergent if $\int_a^\infty |f(x)|dx$ converges.
- If $\int_a^\infty f(x)dx$ converges but $\int_a^\infty |f(x)|dx$ diverges, it is conditionally convergent.

Improper Integrals

Illustrative Examples

Example 1. Examine the convergence of the improper integral

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}}.$$

Solution: Here $f(x) = \dfrac{1}{\sqrt{1-x^2}}$, $x \in (0, 1]$ has a point of infinite discontinuity at $x = 1$.

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \lim_{\epsilon\to 0^+} \int_0^{1-\epsilon} \frac{dx}{\sqrt{1-x^2}} = \lim_{\epsilon\to 0^+} \left[\sin^{-1} x\right]_0^{1-\epsilon} = \sin^{-1}(1) = \frac{\pi}{2}.$$

Thus, the given improper integral converges with value $\pi/2$.

Example 4. Examine the convergence of

$$\int_0^\infty \frac{dx}{(1+x)\sqrt{x}}.$$

Solution:

$$\int_0^\infty \frac{dx}{(1+x)\sqrt{x}} = \lim_{X\to\infty}\int_0^X \frac{dx}{(1+x)\sqrt{x}}.$$

Let $x = t^2 \Rightarrow dx = 2tdt$.

$$= \lim_{X\to\infty}\int_0^{\sqrt{X}} \frac{2t}{(1+t^2)t}dt = 2\lim_{X\to\infty}\tan^{-1}\sqrt{X}.$$

$$= 2\cdot\frac{\pi}{2} = \pi.$$

Hence, the integral is convergent with value π.

—

Example 5. Evaluate

$$\int_{-\infty}^\infty \frac{dx}{x^2+1}.$$

Solution: (Steps from page 80...) Final answer: π.

—

Example 6. Evaluate, if possible, the improper integral

$$\int_{-1}^1 \frac{1+x}{\sqrt{1-x}}dx.$$

Solution: (Steps from page 81...) Final answer: π.

—

Example 7. Examine for convergence

$$\int_0^\infty \cos x\,dx.$$

Solution:

$$\int_0^\infty \cos x\,dx = \lim_{X\to\infty}\sin X,$$

which does not exist. Hence divergent.

—

Example 8. Evaluate

$$\int_{-\infty}^\infty \frac{x}{x^4+1}dx.$$

Solution: (Steps from page 82...) Final answer: 0.

—

Example 9. Evaluate, if possible,

$$\int_0^\pi \frac{dx}{1-\cos x}.$$

2.14 Examples

Example 1

Evaluate $\int_1^\infty \frac{1}{x^2}dx$

Solution:

$$\int_1^\infty \frac{1}{x^2}dx = \lim_{b\to\infty}\left[-\frac{1}{x}\right]_1^b = 1$$

Converges.

Example 2

Evaluate $\int_1^\infty \frac{1}{x}dx$

$$= \lim_{b\to\infty} \ln x\Big|_1^b = \infty \quad \textbf{Diverges}$$

Example 3

Evaluate $\int_0^1 \frac{1}{\sqrt{x}}dx$

$$= \lim_{a\to 0^+}\left[2\sqrt{x}\right]_a^1 = 2 \quad \textbf{Converges}$$

Example 4

Evaluate $\int_0^1 \frac{1}{x^2}dx$

$$= \lim_{a\to 0^+}\left[-\frac{1}{x}\right]_a^1 = \infty \quad \textbf{Diverges}$$

Example 5

Use the comparison test to determine convergence:

$$\int_1^\infty \frac{1}{x^2+1}dx$$

Since $\frac{1}{x^2+1} < \frac{1}{x^2}$ and $\int_1^\infty \frac{1}{x^2}dx$ converges, this also **converges**.

Example 6

Evaluate $\int_1^\infty \frac{\sin x}{x^2}dx$

Use comparison: $|\frac{\sin x}{x^2}| \leq \frac{1}{x^2}$, so it **converges absolutely**.

Example 7

Evaluate $\int_1^\infty \frac{1}{x(\ln x)^2}dx$

Use substitution: Let $u = \ln x$, $du = \frac{1}{x}dx$, then:

$$\int_1^\infty \frac{1}{x(\ln x)^2}dx = \int_0^\infty \frac{1}{u^2}du = 1 \quad \textbf{Converges}$$

Example 8

$$\int_0^1 \ln x \, dx$$

Improper at 0:

$$= \lim_{a \to 0^+} [x \ln x - x]_a^1 = -1 \quad \textbf{Converges}$$

Example 9

$$\int_0^\infty e^{-x}dx = \left[-e^{-x}\right]_0^\infty = 1 \quad \textbf{Converges}$$

Example 10

$$\int_0^1 \frac{\ln x}{\sqrt{x}}dx$$

Use substitution: $x = t^2 \Rightarrow dx = 2tdt$, transform limits, result converges.

Example 11

$$\int_0^1 \frac{1}{x^p}dx \text{ converges if } p < 1$$

Example 12

$$\int_1^\infty \frac{1}{x^p}dx \text{ converges if } p > 1$$

Example 13

$$\int_1^\infty \frac{1}{x\ln x}dx \quad \textbf{Diverges}$$

Example 14

$$\int_2^\infty \frac{1}{x(\ln x)^2}dx \quad \textbf{Converges}$$

Example 15

$$\int_1^\infty \frac{1}{x^{1.01}}dx = \textbf{Converges (p ¿ 1)}$$

3. Solved Examples

(i) $\int_1^\infty \frac{1}{x^2}dx$

(ii) $\int_0^1 \frac{1}{\sqrt{x}}dx$

(iii) $\int_1^\infty \frac{1}{x}dx$

(iv) $\int_0^\infty e^{-x}dx$

(v) $\int_0^1 \ln(x)dx$

(vi) $\int_0^\infty \frac{1}{1+x^2}dx$

(vii) $\int_0^1 \frac{1}{x^p}dx$ for different p

(viii) $\int_0^\infty \frac{\sin x}{x}dx$

(ix) $\int_1^\infty \frac{\ln x}{x^2}dx$

(x) $\int_0^\infty xe^{-x^2}dx$

(xi) $\int_1^\infty \frac{1}{x(\ln x)^2}dx$

(xii) $\int_0^1 \frac{1}{1-x}dx$

(xiii) $\int_0^1 \frac{1}{x \ln x} dx$

(xiv) $\int_0^\infty \frac{x}{(1+x^2)^2} dx$

(xv) $\int_1^\infty \frac{\arctan x}{x^2} dx$

4. Exercise Problems

(i) Evaluate $\int_0^\infty e^{-2x} dx$

(ii) Determine convergence of $\int_1^\infty \frac{\ln x}{x} dx$

(iii) Evaluate $\int_1^\infty \frac{1}{x^3+1} dx$

(iv) Check convergence: $\int_0^1 \frac{1}{x^2} dx$

(v) Evaluate $\int_0^\infty \frac{x}{e^x-1} dx$

(vi) Test convergence: $\int_0^1 \frac{1}{x(\ln x)^2} dx$

(vii) Evaluate $\int_0^\infty \frac{\cos x}{1+x^2} dx$

(viii) Convergence of $\int_2^\infty \frac{1}{x(\ln x)} dx$

(ix) Evaluate $\int_1^\infty \frac{1}{x^{3/2}} dx$

(x) Find $\int_0^\infty x e^{-x} dx$

5. Multiple Choice Questions (MCQs)

1. $\int_1^\infty \frac{1}{x^p} dx$ converges for:
   (a) $p < 1$
   (b) $p = 1$
   (c) $p > 1$
   (d) None

2. Improper integral $\int_0^1 \frac{1}{\sqrt{x}} dx$ is:
   (a) Convergent
   (b) Divergent
   (c) Infinite

(d) Undefined

3. $\int_0^\infty e^{-x}dx$ equals:
   (a) 0
   (b) 1
   (c) ∞
   (d) -1

4. Which of the following tests is not used for improper integrals?
   (a) Ratio test
   (b) Comparison test
   (c) Limit comparison test
   (d) p-test

5. $\int_1^\infty \frac{\ln x}{x^2}dx$ is:
   (a) Convergent
   (b) Divergent
   (c) Oscillating
   (d) Not defined

6. $\int_1^\infty \frac{1}{x}dx$ diverges because:

7. $\int_0^\infty \frac{1}{x^2+1}dx = ?$

8. The integral $\int_1^\infty \frac{1}{(x\ln x)^2}dx$ is:

9. Which of the following is a convergent integral?

10. True or False: All improper integrals over infinite intervals diverge.

11. $\int_0^1 \ln x dx = ?$

12. For which value of p, the integral $\int_0^1 \frac{1}{x^p}dx$ converges?

13. $\int_1^\infty \frac{\sin x}{x}dx$ is:

14. Improper integrals are used in:

15. $\int_0^\infty xe^{-x^2}dx = ?$

2.15 Gamma Function

The improper integral

$$\Gamma(n) = \int_0^\infty e^{-t} t^{n-1}\, dt$$

is called Gamma function.

2.15.1 Properties of Gamma function

Property 1: $\int_0^\infty e^{-at} t^{n-1} dt = \frac{\Gamma(n)}{a^n}$ for $n > 0, a > 0$.
Proof: Let start from the Gamma definition $\Gamma(n) = \int_0^\infty e^{-u} u^{n-1}\, du$ for $n > 0$. In the integral

$$I = \int_0^\infty e^{-at} t^{n-1}\, dt \qquad (a > 0,\ n > 0),$$

substitute $u = at$. Then $t = u/a$ and $dt = du/a$, so

$$I = \int_0^\infty e^{-u} \Big(\frac{u}{a}\Big)^{n-1} \frac{du}{a} = \frac{1}{a^n} \int_0^\infty e^{-u} u^{n-1}\, du = \frac{\Gamma(n)}{a^n}.$$

Thus, $\displaystyle\int_0^\infty e^{-at} t^{n-1}\, dt = \frac{\Gamma(n)}{a^n}$ for $n > 0,\ a > 0$.

Property 2: $\Gamma(n+1) = n\Gamma(n)$ for $n > 0$.
Proof: The Gamma function is defined for $n > 0$ by

$$\Gamma(n) = \int_0^\infty e^{-t} t^{n-1}\, dt.$$

To find $\Gamma(n+1)$, write

$$\Gamma(n+1) = \int_0^\infty e^{-t} t^n\, dt.$$

Use integration by parts:
Let $u = t^n \quad \Rightarrow \quad du = nt^{n-1} dt$, Let $dv = e^{-t} dt \quad \Rightarrow \quad v = -e^{-t}$.
Then

$$\Gamma(n+1) = \int_0^\infty t^n e^{-t} dt = \big[- t^n e^{-t} \big]_0^\infty + n \int_0^\infty t^{n-1} e^{-t} dt.$$

The boundary term $\big[- t^n e^{-t} \big]_0^\infty$ is zero (because $t^n e^{-t} \to 0$ as $t \to \infty$ and is zero at $t = 0$). The remaining integral is exactly $n\,\Gamma(n)$.
Thus:

$$\Gamma(n+1) = n\,\Gamma(n).$$

Property 3: $\Gamma(1) = 1$.
Proof: We prove directly from the definition of the Gamma function:

$$\Gamma(1) = \int_0^\infty e^{-t} t^{1-1}\, dt = \int_0^\infty e^{-t} \cdot 1\, dt = \int_0^\infty e^{-t}\, dt.$$

This is a standard exponential integral:

$$\int_0^\infty e^{-t}\, dt = \big[-e^{-t} \big]_0^\infty = (0) - (-1) = 1.$$

Thus

$$\Gamma(1) = 1$$

Property 4: $\Gamma(n+1) = n!$, when n is a positive integer.
Proof: From the previous result:

$$\Gamma(n+1) = n\,\Gamma(n), n > 0.$$

If n is a positive integer, write:

$$\Gamma(n+1) = n\Gamma(n) = n(n-1)\Gamma(n-1) = n(n-1)(n-2)\Gamma(n-2) \quad \cdots \quad = n(n-1)(n-2)\cdots 2\Gamma(2).$$

We already showed $\Gamma(1) = 1$. Using the recurrence once: $\Gamma(2) = 1.\Gamma(1) = 1$. Thus, the product stops at:

$$\Gamma(n+1) = n(n-1)(n-2)\cdots 2 \cdot 1 \cdot \Gamma(1) = n!.$$

2.16 Beta Function

The Beta function $B(m, n)$ is defined for real numbers $m > 0$ and $n > 0$ as:

$$B(m,n) = \int_0^1 t^{m-1}(1-t)^{n-1}\, dt$$

2.16.1 Properties of Beta Function

Property 1: $B(m, n) = B(n, m); m, n > 0$.
Proof: Start with

$$B(m,n) = \int_0^1 t^{m-1}(1-t)^{n-1} dt.$$

Substitute $t = 1 - u$, so $dt = -du$. When $t = 0 \to u = 1$ and $t = 1 \to u = 0$.

$$B(m,n) = \int_1^0 (1-u)^{m-1}u^{n-1}(-du) = \int_0^1 u^{n-1}(1-u)^{m-1}du.$$

This is exactly $B(n,m)$.

$$B(m,n) = B(n,m)$$

.

Property 2:

$$B(m,n) \ = \int_0^\infty \frac{x^{n-1}}{(1+x)^{m+n}}\,dx.$$

and

$$B(m,n) = \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}\,dx$$

.

Proof: Start from the definition on $[0,1]$

$$B(m,n) = \int_0^1 t^{m-1}(1-t)^{n-1}\,dt.$$

Use the substitution $x = \dfrac{t}{1-t}$. Then $t = \frac{x}{1+x}$, $1-t = \frac{1}{1+x}$, and

$$dt = \frac{1}{(1+x)^2}\,dx.$$

. As t runs from 0 to 1, x runs from 0 to ∞. Compute the transformed integrand:

$$t^{m-1}(1-t)^{n-1}\,dt = \Big(\frac{x}{1+x}\Big)^{m-1}\Big(\frac{1}{1+x}\Big)^{n-1}\frac{dx}{(1+x)^2} = \frac{x^{m-1}}{(1+x)^{m+n}}\,dx.$$

Therefore,

$$B(m,n) = \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}\,dx.$$

The Beta function is symmetric: $B(m,n) = B(n,m)$. Applying symmetry, we have

$$B(m,n) = B(n,m) = \int_0^\infty \frac{x^{n-1}}{(1+x)^{n+m}}\,dx,$$

which is exactly

$$B(m,n) = \int_0^\infty \frac{x^{n-1}}{(1+x)^{m+n}}\,dx.$$

.

Property 3: $B(m,n) = 2\int_0^{\pi/2} sin^{2m-1}\theta cos^{2n-1}\theta d\theta; m,n > 0.$

Proof: Start from the Beta function's definition on $[0,1]$:

$$B(m,n) = \int_0^1 t^{m-1}(1-t)^{n-1}\,dt, \qquad m>0,\ n>0.$$

Use the substitution

$$t = \sin^2\theta \quad \Longrightarrow \quad dt = 2\sin\theta\cos\theta\,d\theta.$$

When t goes from 0 to 1, θ goes from 0 to $\frac{\pi}{2}$. Also,

$$t^{m-1} = (\sin^2\theta)^{m-1} = \sin^{2m-2}\theta, \qquad (1-t)^{n-1} = (\cos^2\theta)^{n-1} = \cos^{2n-2}\theta.$$

Therefore,

$$\begin{aligned} B(m,n) &= \int_0^1 t^{m-1}(1-t)^{n-1}\,dt \\ &= \int_0^{\pi/2} \sin^{2m-2}\theta\,\cos^{2n-2}\theta\,\big(2\sin\theta\cos\theta\,d\theta\big) \\ &= 2\int_0^{\pi/2} \sin^{2m-1}\theta\,\cos^{2n-1}\theta\,d\theta. \end{aligned}$$

Property 4: $B(\frac{1}{2},\frac{1}{2}) = \pi$.
Proof: A neat proof uses the substitution $t = \sin^2\theta$.
Start from the definition:

$$B\left(\frac{1}{2},\frac{1}{2}\right) = \int_0^1 t^{-1/2}(1-t)^{-1/2}\,dt = \int_0^1 \frac{dt}{\sqrt{t(1-t)}}.$$

Let $t=\sin^2\theta$. Then $dt = 2\sin\theta\cos\theta\,d\theta$, and as t goes $0\to 1$, θ goes $0 \to \frac{\pi}{2}$. Also,

$$\sqrt{t(1-t)} = \sqrt{\sin^2\theta\,\cos^2\theta} = \sin\theta\cos\theta$$

Therefore,

$$\int_0^1 \frac{dt}{\sqrt{t(1-t)}} = \int_0^{\pi/2} \frac{2\sin\theta\cos\theta}{\sin\theta\cos\theta}\,d\theta = 2\int_0^{\pi/2} d\theta = \pi.$$

Hence,

$$B\left(\frac{1}{2},\frac{1}{2}\right) = \pi.$$

.

2.16.2 Relation between Gamma Function and Beta Function

$$B(m,n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}.$$

Proof: Start with product of two Gamma functions

$$\Gamma(m)\,\Gamma(n) = \int_0^\infty x^{m-1}e^{-x}dx \int_0^\infty y^{n-1}e^{-y}dy.$$

This can be written as a double integral:

$$\Gamma(m)\,\Gamma(n) = \int_0^\infty \int_0^\infty x^{m-1}y^{n-1}e^{-(x+y)}\,dx\,dy.$$

Introduce new variables: Let

$$u = \frac{x}{x+y}, \quad v = x+y.$$

* As $x > 0, y > 0$: $0 < u < 1$ and $0 < v < \infty$. * Solve for x, y:

$$x = uv, \quad y = (1-u)v.$$

* The Jacobian determinant is:

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = v.$$

Now the double integral becomes:

$$\Gamma(m)\,\Gamma(n) = \int_0^1 \int_0^\infty (uv)^{m-1}\big((1-u)v\big)^{n-1}e^{-v}\,v\,dv\,du.$$

Combine powers of v:

$$(uv)^{m-1}((1-u)v)^{n-1}v = u^{m-1}(1-u)^{n-1}v^{m+n-1}.$$

So the integral becomes:

$$\Gamma(m)\Gamma(n) = \int_0^1 u^{m-1}(1-u)^{n-1}\,du \int_0^\infty v^{m+n-1}e^{-v}\,dv.$$

The first integral is:

$$\int_0^1 u^{m-1}(1-u)^{n-1}du = B(m,n).$$

The second integral is:

$$\int_0^\infty v^{m+n-1}e^{-v}dv = \Gamma(m+n).$$

Thus

$$B(m,n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}.$$

2.16.3 Some Standard Results

Result 1: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.
Proof: Here's a clean beta–gamma proof.
Use the beta–gamma relation

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

With $x = y = \frac{1}{2}$ and $\Gamma(1) = 1$,

$$\Gamma\left(\frac{1}{2}\right)^2 = B\left(\frac{1}{2}, \frac{1}{2}\right).$$

Now evaluate $B(\frac{1}{2}, \frac{1}{2})$ from its integral form:

$$B\left(\frac{1}{2}, \frac{1}{2}\right) = \int_0^1 t^{-1/2}(1-t)^{-1/2}\,dt.$$

Let $t = \sin^2\theta$ so $dt = 2\sin\theta\cos\theta\,d\theta$, and when $t : 0 \to 1$ we have $\theta : 0 \to \frac{\pi}{2}$. Then

$$\begin{aligned} B\left(\frac{1}{2}, \frac{1}{2}\right) &= \int_0^{\pi/2} (sin^2\theta)^{-1/2}(\cos^2\theta)^{-1/2} 2\sin\theta\cos\theta d\theta \\ &= 2\int_0^{\pi/2} d\theta \\ &= \pi \end{aligned}$$

Hence $\Gamma(\frac{1}{2})^2 = \pi$. Since $\Gamma(\frac{1}{2}) > 0$, we conclude

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Result 2: $\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2}$.
Proof: Using the Gamma function,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-1/2}e^{-t}\,dt = \sqrt{\pi}.$$

Let $t = x^2$ so $dt = 2x\,dx$ and $x = t^{1/2}$. Then

$$\int_0^\infty e^{-x^2}\,dx = \int_0^\infty e^{-t}\,\frac{dt}{2\,t^{1/2}} = \frac{1}{2}\int_0^\infty t^{-1/2}e^{-t}\,dt = \frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

Result 3: $\int_0^{\pi/2} sin^p\theta cos^q\theta d\theta = \frac{1}{2}\frac{\Gamma(\frac{p+1}{2})(\frac{q+1}{2})}{(\frac{p+q+2}{2})}; p > -1, q > -1.$
Proof: Assume $p > -1$ and $q > -1$ so the integral converges. Put

$$I := \int_0^{\pi/2} \sin^p\theta\cos^q\theta\,d\theta.$$

Use the substitution $t = \sin^2\theta$. Then $dt = 2\sin\theta\cos\theta\,d\theta$, so

$$d\theta = \frac{dt}{2\sqrt{t(1-t)}}.$$

Also $\sin^p\theta = (\sin^2\theta)^{p/2} = t^{p/2}$ and $\cos^q\theta = (1-t)^{q/2}$. Hence

$$\begin{aligned} I &= \int_0^1 t^{p/2}(1-t)^{q/2}\,\frac{dt}{2\sqrt{t(1-t)}} \\ &= \frac{1}{2}\int_0^1 t^{(p-1)/2}(1-t)^{(q-1)/2}\,dt. \end{aligned}$$

Recognize the Beta integral:

$$I = \frac{1}{2}\,B\Big(\frac{p+1}{2}, \frac{q+1}{2}\Big).$$

Using the Beta–Gamma relation $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ gives

$$\int_0^{\pi/2} \sin^p\theta\cos^q\theta\,d\theta = \frac{1}{2}\frac{\Gamma\big(\frac{p+1}{2}\big)\,\Gamma\big(\frac{q+1}{2}\big)}{\Gamma\big(\frac{p+q+2}{2}\big)}, \qquad p > -1,\ q > -1.$$

This completes the proof.
Result 5: $\Gamma(m)\Gamma(1-m) = \frac{\pi}{sinm\pi}; 0 < m < 1.$

2.16.4 Worked Out Examples

Example 1: Evaluate $\int_0^{\frac{\pi}{2}} sin^4xcos^5xdx.$
Solution: We use the formula

$$\int_0^{\pi/2} \sin^p x\cos^q x\,dx = \frac{1}{2}\frac{\Gamma\big(\frac{p+1}{2}\big)\Gamma\big(\frac{q+1}{2}\big)}{\Gamma\big(\frac{p+q+2}{2}\big)} \qquad (p > -1, q > -1).$$

Here $p = 4,\ q = 5$. So,

$$I = \int_0^{\pi/2} \sin^4 x\cos^5 x\,dx = \frac{1}{2}\frac{\Gamma(\frac{5}{2})\,\Gamma(3)}{\Gamma(\frac{11}{2})}.$$

Compute the Gamma values: $\Gamma(3) = 2! = 2$, $\Gamma(\frac{5}{2}) = \frac{3}{2}\Gamma(\frac{3}{2}) = \frac{3}{2} \cdot \frac{1}{2}\sqrt{\pi} = \frac{3}{4}\sqrt{\pi}$, and

$$\Gamma(\frac{11}{2}) = \frac{9}{2} \cdot \frac{7}{2} \cdot \frac{5}{2} \cdot \frac{3}{2} \cdot \frac{1}{2}\sqrt{\pi} = \frac{945}{32}\sqrt{\pi}.$$

Thus

$$I = \frac{1}{2}\frac{\frac{3}{4}\sqrt{\pi} \cdot 2}{\frac{945}{32}\sqrt{\pi}} = \frac{1}{2}\frac{\frac{3}{2}}{\frac{945}{32}} = \frac{1}{2} \cdot \frac{3}{2} \cdot \frac{32}{945} = \frac{8}{315}.$$

So

$$\int_0^{\pi/2} \sin^4 x \cos^5 x \, dx = \frac{8}{315}.$$

Example 2: Evaluate $\int_0^{\pi/2} \sqrt{tanx}dx$.
Solution: Write the integral

$$I = \int_0^{\pi/2} \sqrt{\tan x} \, dx = \int_0^{\pi/2} \sin^{1/2} x \cos^{-1/2} x \, dx.$$

Compare exponents with the integrand in the beta formula:

$$\sin^{1/2} x \cos^{-1/2} x = \sin^{2m-1} x \cos^{2n-1} x$$

gives $2m - 1 = \frac{1}{2}$ and $2n - 1 = -\frac{1}{2}$. Hence

$$m = \frac{3}{4}, \qquad n = \frac{1}{4},$$

which are positive, so the formula applies. Therefore

$$I = \frac{1}{2}B\Big(\frac{3}{4}, \frac{1}{4}\Big) = \frac{1}{2}\,\Gamma\Big(\frac{3}{4}\Big)\Gamma\Big(\frac{1}{4}\Big).$$

Using the trigonometric (reflection) identity $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$ with $z = \frac{3}{4}$ (so $\sin(3\pi/4) = \frac{\sqrt{2}}{2}$), we get

$$\Gamma\Big(\frac{3}{4}\Big)\Gamma\Big(\frac{1}{4}\Big) = \frac{\pi}{\sin(3\pi/4)} = \pi\sqrt{2}.$$

Thus

$$I = \frac{1}{2}\pi\sqrt{2} = \frac{\pi}{\sqrt{2}}.$$

$$\int_0^{\pi/2} \sqrt{\tan x} \, dx = \frac{1}{2}B\Big(\frac{3}{4}, \frac{1}{4}\Big) = \frac{\pi}{\sqrt{2}}.$$

Example 3: Prove that $\int_0^\infty e^{-x^4} dx \times \int_0^\infty x^2 e^{-x^4} dx = \frac{\pi}{8\sqrt{2}}$.
Proof: We need to evaluate

$$I_1 = \int_0^\infty e^{-x^4}\, dx, \qquad I_2 = \int_0^\infty x^2 e^{-x^4}\, dx$$

and show that $I_1 I_2 = \frac{\pi}{8\sqrt{2}}$.

Step 1. Evaluate I_1
Let $t = x^4 \;\Rightarrow\; dt = 4x^3 dx$. We need dx in terms of dt:

$$dx = \frac{dt}{4x^3} = \frac{dt}{4t^{3/4}}.$$

Thus

$$I_1 = \int_0^\infty e^{-x^4} dx = \int_0^\infty e^{-t} \frac{dt}{4t^{3/4}} = \frac{1}{4}\int_0^\infty t^{-3/4} e^{-t}\, dt.$$

This integral is the Gamma function:

$$\int_0^\infty t^{a-1} e^{-t} dt = \Gamma(a).$$

Here $a - 1 = -3/4 \;\Rightarrow\; a = 1/4$. So

$$I_1 = \frac{1}{4}\Gamma\Big(\frac{1}{4}\Big).$$

Step 2. Evaluate I_2
Again set $t = x^4$, $dx = \frac{dt}{4t^{3/4}}$, $x^2 = t^{\frac{1}{2}}$:

$$I_2 = \int_0^\infty x^2 e^{-x^4} dx = \int_0^\infty t^{1/2} e^{-t} \frac{dt}{4t^{3/4}} = \frac{1}{4}\int_0^\infty t^{-1/4} e^{-t}\, dt.$$

Now $a - 1 = -\frac{1}{4} \;\Rightarrow\; a = \frac{3}{4}$.

$$I_2 = \frac{1}{4}\Gamma\Big(\frac{3}{4}\Big).$$

Step 3. Product

$$I_1 I_2 = \frac{1}{4}\Gamma\Big(\frac{1}{4}\Big) \cdot \frac{1}{4}\Gamma\Big(\frac{3}{4}\Big) = \frac{1}{16}\Gamma\Big(\frac{1}{4}\Big)\Gamma\Big(\frac{3}{4}\Big).$$

Step 4. Use Gamma reflection formula
The reflection formula:

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

Take $z = \frac{1}{4}$, so $1 - z = \frac{3}{4}$ and $\sin(\pi/4) = \frac{\sqrt{2}}{2}$:

$$\Gamma\Big(\frac{1}{4}\Big)\Gamma\Big(\frac{3}{4}\Big) = \frac{\pi}{\sin(\pi/4)} = \frac{\pi}{\sqrt{2}/2} = \pi\sqrt{2}.$$

Thus,

$$I_1 I_2 = \frac{1}{16}\pi\sqrt{2} = \frac{\pi}{16/\sqrt{2}} = \frac{\pi}{8\sqrt{2}}.$$

$$\int_0^\infty e^{-x^4}\,dx \;\times\; \int_0^\infty x^2 e^{-x^4}\,dx \;=\; \frac{\pi}{8\sqrt{2}}.$$

Example 4: Prove that $\int_0^1 \sqrt{1-x^4}dx = \frac{\Gamma(\frac{1}{4})}{6\sqrt{2\pi}}$.
Solution: Let

$$I = \int_0^1 (1-x^4)^{1/2}\,dx.$$

Put $t = x^4 \Rightarrow x = t^{\frac{1}{4}},\ dx = \frac{1}{4}t^{-\frac{3}{4}}\,dt$. Then

$$I = \frac{1}{4}\int_0^1 t^{-\frac{3}{4}}(1-t)^{\frac{1}{2}}\,dt = \frac{1}{4}\,B\left(\frac{1}{4},\frac{3}{2}\right) = \frac{1}{4}\,\frac{\Gamma(\frac{1}{4})\Gamma(\frac{3}{2})}{\Gamma(\frac{7}{4})}.$$

Use $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$ and $\Gamma(\frac{7}{4}) = \frac{3}{4}\,\Gamma(\frac{3}{4})$ to get

$$I = \frac{1}{3}\cdot\frac{\Gamma(\frac{1}{4})\,\Gamma(\frac{3}{2})}{\Gamma(\frac{3}{4})} = \frac{\sqrt{\pi}}{6}\,\frac{\Gamma(\frac{1}{4})}{\Gamma(\frac{3}{4})}.$$

Now apply Euler's reflection formula with $z = \frac{1}{4}$:

$$\Gamma\left(\frac{1}{4}\right)\Gamma\left(\frac{3}{4}\right) = \frac{\pi}{\sin(\pi/4)} = \pi\sqrt{2} \quad\Longrightarrow\quad \frac{1}{\Gamma(\frac{3}{4})} = \frac{\Gamma(\frac{1}{4})}{\pi\sqrt{2}}.$$

Hence

$$I = \frac{\sqrt{\pi}}{6}\,\Gamma\left(\frac{1}{4}\right)\cdot\frac{\Gamma(\frac{1}{4})}{\pi\sqrt{2}} = \frac{\Gamma(\frac{1}{4})^2}{6\sqrt{2\pi}}.$$

So, the integral equals $\dfrac{\Gamma(\frac{1}{4})^2}{6\sqrt{2\pi}}$.

Example 5. Prove that $\int_0^\infty e^{-4x}x^{\frac{3}{2}}dx = \frac{3}{128}\sqrt{\pi}$.
Solution: Compute the integral by reducing it to the Gamma function. Start with the substitution $t = 4x$ (so $x = t/4,\ dx = dt/4$):

$$\int_0^\infty e^{-4x}x^{3/2}\,dx = \int_0^\infty e^{-t}\left(\frac{t}{4}\right)^{3/2}\frac{dt}{4} = \frac{1}{4^{5/2}}\int_0^\infty e^{-t}t^{5/2-1}\,dt.$$

The remaining integral is the Gamma function $\Gamma(5/2)$, so

$$\int_0^\infty e^{-4x} x^{3/2}\, dx = \frac{\Gamma(5/2)}{4^{5/2}}.$$

Evaluate $\Gamma(\frac{5}{2})$ using the recurrence $\Gamma(z+1) = z\Gamma(z)$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$:

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{2}\Gamma\left(\frac{3}{2}\right) = \frac{3}{2}\cdot\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\sqrt{\pi}.$$

Also $4^{5/2} = 32$. Therefore,

$$\int_0^\infty e^{-4x} x^{3/2}\, dx = \frac{\frac{3}{4}\sqrt{\pi}}{32} = \frac{3}{128}\sqrt{\pi},$$

as required.

Exercises:
1. Prove that $\int_0^\infty \frac{dt}{\sqrt{t}(1+t)} = B(\frac{1}{2}, \frac{1}{2}) = \pi$.
2. Show that $\int_0^\infty \sqrt{x} e^{-x^3} dx = \frac{\sqrt{\pi}}{3}$.
3. Find $\int_0^{\frac{\pi}{2}} sin^9 x dx$.
4. Prove that $\int_0^{\frac{\pi}{2}} sin^p x dx \times \int_0^{\frac{\pi}{2}} sin^{p+1} x dx = \frac{\pi}{2(p+1)}$.
5. Evaluate $\Gamma(6)$.
6. Evaluate $\Gamma\left(\frac{5}{2}\right)$.
7. Show that $\Gamma(n+1) = n!$ when n is a positive integer.
8. Evaluate $\int_0^\infty x^4 e^{-x}\, dx$.
9. Find $\int_0^\infty x^{n-1} e^{-2x}\, dx$.
10. Evaluate $\int_0^\infty \sqrt{x} e^{-x}\, dx$.
11. Prove that $\Gamma(1) = 1$.
12. Find $\int_0^\infty x^2 e^{-3x}\, dx$.
13. Show that $\int_0^\infty e^{-x^2}\, dx = \dfrac{\sqrt{\pi}}{2}$.
14. Evaluate $\int_0^\infty x^7 e^{-x}\, dx$.
15. Evaluate $B(3, 4)$.
16. Show that $B\left(\dfrac{1}{2}, \dfrac{1}{2}\right) = \pi$.
17. Evaluate $\int_0^1 x^2 (1-x)^3\, dx$.
18. Prove that $B(m, n) = \dfrac{(m-1)!(n-1)!}{(m+n-1)!}$ for integers m, n.
19. Show that $\int_0^{\pi/2} \sin^4 x \cos^2 x\, dx = \dfrac{1}{2} B\left(\dfrac{5}{2}, \dfrac{3}{2}\right)$.
20. Evaluate $\int_0^1 \sqrt{x}(1-x)^4\, dx$.
21. Show that $\int_0^1 \dfrac{x^{m-1}}{(1+x)^{m+n}} dx = B(m, n)$.

22. Find $\int_0^1 \dfrac{x^2}{\sqrt{1-x}}\,dx$.
23. Evaluate $\int_0^1 (x^3 - x^2)^2\,dx$.
24. Show that $\int_0^{\pi/2} \sin^{m-1} x \cos^{n-1} x\,dx = \dfrac{1}{2} B\left(\frac{m}{2}, \frac{n}{2}\right)$.