Semester 3:

Computational Methods → [UNIT - 1]

★ **Introduction** → There are two types of equations, i) Algebraic Equations and, ii) Transcedental Equations.

- We already know how to find roots of an algebraic equation for any polynomial degree of 'x' from quadratic to cubic and even bi-quadratic equations.

For examples,

(a) Quadratic Equation : $ax^2 + bx + c = 0$

→ can be solved by : ① Factorisation Method
② Discriminant Method
③ Completing the Squares.

(b) Cubic Equation : $ax^3 + bx^2 + cx + d = 0$.

→ can be solved by: remainder theorem.

→ **Transcedental Equations** : A transcedental equation is an equation into which transcedental functions (such as logarithm, exponentials, trigonometric and inverse trigonometric) of one of the variables have been solved for. For examples,

i) $xe^x + 2\log x = 0$, 2) $\sin(\log x) + 4\cos^{-1}x = 0$

- To solve such equations and find its roots we have several methods which includes,

1) Bisection Method
2) Newton Raphson Method
3) Secant Method

] → (In Syllabus)

4) Regula - Falsi Method
5) Iteration Method.

1) Methods to find roots of Transcedental Equations :-

1.1) BISECTION METHOD → Also called 'Bolzano Method'.

→ Let $f(x)$ be a continous function in the closed interval $[a, b]$ and let,

$f(a)$ be positive.

$f(b)$ be negative.

such that,

$$f(a) \cdot f(b) < 0$$

If it is true then there exists atleast one real root between a and b such that $f(x) = 0$.

Q.1) Solve the given equation by using the Bisection method to find its roots:

$$f(x) = x^3 - 4x - 9 = 0$$

Ans: Step 1: To find the range / interval $[a, b]$

By hit and trial for different values of x,

⇒ $f(0) = (0)^3 - 4(0) - 9 = -9$ (negative)

⇒ $f(1) = (1)^3 - 4(1) - 9 = -12$

⇒ $f(2) = (2)^3 - 4(2) - 9 = -9$

⇒ $f(3) = (3)^3 - 4(3) - 9 = +6$ (positive)

- The required interval for which $f(x)$ has two opposite sign values at a and b is, $[0, 3]$.

⇒ The root of the given equation lies in between 2 and 3, i.e.,

$$f(2) \cdot f(3) < 0 ;$$

so closest and smallest interval is $[2, 3]$

Step 2:

- The root of the equation lies in between 2 and 3, so;

$$x_1 = \frac{2+3}{2} = 2.5$$

$\Rightarrow$ $f(x_1) = f(2.5) = -3.375 < 0$

- The next shortest interval in which the root of the equation lies in is, $[x_1, 3]$.

$$x_2 = \frac{x_1+3}{2} = \frac{2.5+3}{2} = \frac{5.5}{2} = 2.75$$

$\Rightarrow$ $f(x_2) = f(2.75) = 0.7969 > 0$

- The next shortest interval in which the root of the equation lies in is, $[x_1, x_2]$.

$$x_3 = \frac{x_1+x_2}{2} = \frac{2.5+2.75}{2} = 2.625$$

$\Rightarrow$ $f(x_3) = -1.4121 < 0$

- Similarly, for interval $[x_3, x_2]$;

$$x_4 = \frac{x_2+x_3}{2} = \frac{2.75+2.625}{2} = 2.6875$$

$\Rightarrow$ $f(x_4) = -0.33911 < 0$

- Similarly, for interval $[x_2, x_4]$

$$x_5 = \frac{x_2+x_4}{2} = \frac{2.75+2.687}{2} = 2.718$$

$\Rightarrow$ $f(x_5) = 0.2072 > 0$

- Now, in interval $[x_4, x_5]$;

$\Rightarrow x_6 = \frac{x_4 + x_5}{2} = \frac{2.687 + 2.718}{2} = 2.7025$

$\therefore f(x_6) = -0.0812 < 0$.

- The next root lies in range for $[x_5, x_6]$.

$\Rightarrow x_7 = \frac{x_5 + x_6}{2} = \frac{2.718 + 2.702}{2} = 2.710$

$\therefore f(x_7) = 0.0625 > 0$.

- Similarly, further in next interval $[x_6, x_7]$ i.e.,

$\Rightarrow x_8 = \frac{x_6 + x_7}{2} = \frac{2.702 + 2.710}{2} = 2.706$.

$\therefore f(x_8) = -0.00948 < 0$.

- Further in the next interval $[x_7, x_8]$;

$\Rightarrow x_9 = \frac{x_7 + x_8}{2} = \frac{2.71 + 2.706}{2} = 2.708$.

$\therefore f(x_9) = 0.0264 > 0$.

- Finally, taking the last iteration for interval $[x_8, x_9]$;

$\Rightarrow x_{10} = \frac{x_8 + x_9}{2} = \frac{2.706 + 2.708}{2} = 2.707$.

$\therefore f(x_{10}) = f(2.707) = 0.00848 > 0$.

$\rightarrow$ Therefore, the closest root of this equation $f(x) = 0$ after 10 iterations is;

$$\boxed{X = 2.707}$$

Q.2) Solve the given below transcedental equation using the bisection method ?

$$f(x) = xe^{x} - 1$$ (Most Important)

Ans = • First, we need to find the range in which $f(x)$ lies, i.e.,

$f(0) = 0 - 1 = -1$

$f(1) = e - 1 = 1.71$ ($\because$ Value of 'e' $= 2.718$)

• The required closed interval is $[0, 1]$ where at both end-points $f(x)$ has opposite signs.

Now, finding iteration 1,

$\Rightarrow x_1 = \frac{0+1}{2} = 0.5$

$\Rightarrow f(x_1) = f(0.5) = 0.5 \times e^{0.5} - 1 = -0.175 < 0$

Finding iteration 2 in interval $[x_1, 1]$;

$\Rightarrow x_2 = \frac{x_1 + 1}{2} = \frac{0.5+1}{2} = \frac{1.5}{2} = 0.75$

$\Rightarrow f(x_2) = f(0.75) = 0.587 > 0$

Finding iteration 3 in interval $[x_1, x_2]$,

$\Rightarrow x_3 = \frac{x_1 + x_2}{2} = \frac{0.5 + 0.75}{2} = 0.625$

$\Rightarrow f(x_3) = f(0.625) = 0.1676 > 0$

For iteration 4 in $[x_1, x_3]$,

$\Rightarrow x_4 = \frac{x_1 + x_3}{2} = \frac{0.5 + 0.625}{2} = 0.562$

$\Rightarrow f(x_4) = f(0.562) = -0.0127 < 0$

For iteration 5 in the interval $[x_3, x_4]$;

$\Rightarrow x_5 = \frac{x_3 + x_4}{2} = \frac{0.625 + 0.562}{2} = 0.593$

$\Rightarrow f(x_5) = 0.0729 > 0$.

Next, for iteration 6 in the interval $[x_4, x_5]$,

$\Rightarrow x_6 = \frac{x_4 + x_5}{2} = \frac{0.562 + 0.593}{2} = 0.577$

$\Rightarrow f(x_6) = 0.0274 > 0$.

Next, for iteration 7 in the interval $[x_4, x_6]$,

$\Rightarrow x_7 = \frac{x_4 + x_6}{2} = 0.569$

$\Rightarrow f(x_7) = 0.005138 > 0$

Next, for iteration 8 in the interval $[x_4, x_7]$,

$\Rightarrow x_8 = \frac{x_4 + x_7}{2} = \frac{0.562 + 0.569}{2} = 0.565$

$\Rightarrow f(x_8) = f(0.565) = -0.00591 < 0$

Taking one last iteration 9 in the interval $[x_7, x_8]$,

$\Rightarrow x_9 = \frac{x_7 + x_8}{2} = \frac{0.569 + 0.565}{2} = 0.567$.

$\Rightarrow f(x_9) = f(0.567) = -0.0003958 < 0$

$\Rightarrow$ Hence, the closest root of this equation $f(x)$ after 9 iterations by bisection method is,

$$\boxed{x = 0.567}$$

1.2) Rate of Convergence of Bisection Method :-

(without PROOF)

→ If the bisection algorithm is applied to a continous function f on an interval $[a, b]$ where,

$$f(a) \cdot f(b) < 0$$

then,

after n steps, an approximate root will be computed with error at most,

$$= \frac{(b-a)}{2^{n+1}} \qquad \text{[Linear Convergence]}$$

- If an error tolerance (ε) has been prescibed in advance then,

$$\frac{b-a}{2^{n+1}} < \varepsilon \qquad (\varepsilon : \text{very-very small number}).$$

⇒ Taking logarithms on both sides,

$$\Rightarrow \log(b-a) - \log(2^{n+1}) < \log \varepsilon$$

$$\Rightarrow \log\left(\frac{b-a}{\varepsilon}\right) < (n+1)\log 2$$

$$\Rightarrow \frac{\log(b-a) - \log \varepsilon}{\log 2} < n+1$$

or,

$$\Rightarrow n > \frac{\log(b-a) - \log(\varepsilon)}{\log 2}$$

therefore,

$$\boxed{n > \frac{\log\left(\frac{b-a}{\varepsilon}\right)}{\log 2}}$$

→ 'n' is only dependent on the quantity ε.

(Asked in MT-2019, 3 MARKS).

1.3) NEWTON-RAPHSON METHOD → (Geometrical Interpretation)

- Applying the equation of tangent for the curve $f(x)$ at point P_0,

$\Rightarrow$ $y - y_1 = m(x - x_1)$

$\Rightarrow$ $y - y_1 = \frac{dy}{dx}(x - x_1)$

At point P_0,

$\Rightarrow$ $y - f(x_0) = \frac{dy}{dx}(x - x_0)$

$\because$ $y = f(x)$

$\Rightarrow \frac{dy}{dx} = f'(x)$

$\Rightarrow \left.\frac{dy}{dx}\right|_{(x_0)} = f'(x_0)$.

Slope at point Q;

$\Rightarrow$ $0 - f(x_0) = f'(x_0)(x_1 - x_0)$

$\Rightarrow$ $\frac{-f(x_0)}{f'(x_0)} = x_1 - x_0$

$\therefore$ $\boxed{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}}$

$\therefore$ General form can be written as,

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

where, $f'(x_n) \neq 0$.

Note: It's rate of convergence is "Quadratic", and is a much faster way to find the roots.

Q.3) Find the root of the equation by Newton-Raphson's method:

$$f(x) = x^3 - 3x - 5$$

Ans: Step 1: To find the closest value of x to near '0', i.e.,

$\Rightarrow f(1) = 1 - 3 - 5 = -7$

$\Rightarrow f(-1) = -1 + 3 - 5 = -3$

$\Rightarrow f(2) = 8 - 6 - 5 = -3$ $\rightarrow$ closest to 'zero'.

$\Rightarrow f(3) = 27 - 9 - 5 = 16$

- Clearly, for $\boxed{x_0 = 2}$, $f(x)$ is closest for zero.

Step 2: Apply the formula.

Since,

$$f(n) = x_n^3 - 3x_n - 5$$

and,

$$f'(n) = 3x_n^2 - 3$$

- From Newton-Raphson's formula,

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^3 - 3x_n - 5}{3x_n^2 - 3}$$

or,

$$\Rightarrow x_{n+1} = \frac{2x_n^3 + 5}{3x_n^2 - 3}$$

(i) For $n = 0$,

$$\Rightarrow x_1 = \frac{2x_0^3 + 5}{3x_0^2 - 3} = \frac{2(2)^3 + 5}{3(2)^2 - 3} = \frac{21}{9} = 2.333$$

$(\because x_0 = 2)$.

(ii) For $n=1$,

$$\Rightarrow x_2 = \frac{2x_1^3+5}{3x_1^2-3} = \frac{2(2.333)^3+5}{3(2.333)^2-3} = \frac{30.396}{13.328} = 2.2806$$

(iii) For $n=2$,

$$\Rightarrow x_3 = \frac{2x_2^3+5}{3x_2^2-3} = \frac{2(2.280)^3+5}{3(2.280)^2-3} = 2.2790.$$

(iv) For $n=3$,

$$\Rightarrow x_4 = \frac{2x_3^3+5}{3x_3^2-3} = \frac{2(2.279)^3+5}{3(2.279)^2-3} = 2.2790.$$

→ Clearly, for the values of $n=2$ and 3 both x_3 and x_4 are equal.
Therefore, $\boxed{x = 2.2790}$

Q.4) Solve the given function using Newton-Raphson method?

$$x = (30)^{-1/5}$$ [ET-2019, 4 MARKS]

Ans: Given, $\Rightarrow x = (30)^{-1/5}$

$$\Rightarrow x^5 = \frac{1}{30} \Rightarrow 30x^5 = 1 \Rightarrow \boxed{30x^5 - 1 = 0}$$

therefore,

$$f(x) = 30x^5 - 1$$

Considering, $x_0 = 0.5$ to get $f(x)$ closest to zero.

$\Rightarrow f(x_n) = 30x_n^5 - 1 \Rightarrow f'(x_n) = 150x_n^4$

★ From Newton's - Raphson formula,

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$\Rightarrow x_{n+1} = x_n - \frac{(30x_n^5 - 1)}{150x_n^4}$$

$$\therefore x_{n+1} = \frac{120x_n^5 + 1}{150x_n^4}$$

(i) For $n = 0$,

$$\Rightarrow x_1 = \frac{120x_0^5 + 1}{150x_0^4} = \frac{120(0.5)^5 + 1}{150(0.5)^4} = 0.5066$$

(ii) For $n = 1$,

$$\Rightarrow x_2 = \frac{120x_1^5 + 1}{150x_1^4} = \frac{120(0.506)^5 + 1}{150(0.506)^4} = \cancel{1.00098}\ 0.5064$$

(iii) For $n = 2$,

$$\Rightarrow x_3 = \frac{120x_2^5 + 1}{150x_2^4} = \frac{120(0.5064)^5 + 1}{150(0.5064)^4} = 0.50649$$

→ For the values of $n = 1$ and 2 both x_2 and x_3 are equal.

- Therefore, the root of this equation is,

$$\boxed{x = 0.5064} \quad (x \in R)$$

Q.5) Solve the given equation by Newton's Method ?

$$f(x) = \cos x - xe^{x} \quad \text{(Important)}$$

Ans: For,

$X=0$: $f(0) = \cos 0 - 0 = 1 > 0$

$X=1$: $f(1) = \cos 1 - e = -1.718 < 0.$

- Required interval is : $[0, 1]$.

→ Considering $\boxed{x_0 = 0}$ as $f(x_0)$ is closest to zero,

i.e.,

$$f(x_n) = \cos x_n - x_n e^{x_n}$$

and,

$$\Rightarrow f'(x_n) = -\sin x_n - (e^{x_n} + x_n e^{x_n})$$

$$\Rightarrow f'(x_n) = -\sin x_n - x_n e^{x_n} - e^{x_n}$$

* From Newton's - Raphson formula,

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$\Rightarrow x_{n+1} = x_n - \frac{\cos x_n - x_n e^{x_n}}{-(\sin x_n + x_n e^{x_n} + e^{x_n})}$$

further,

$$\Rightarrow x_{n+1} = \frac{x_n[-(\sin x_n + x_n e^{x_n} + e^{x_n})] - \cos x_n + x_n e^{x_n}}{-(\sin x_n + x_n e^{x_n} + e^{x_n})}$$

$$\Rightarrow x_{n+1} = \frac{\cancel{-}(x_n \sin x_n + x_n^2 e^{x_n} + \cos x_n)}{\cancel{-}(\sin x_n + x_n e^{x_n} + e^{x_n})}$$

now,

$$x_{n+1} = \frac{x_n(\sin x_n + x_n e^{x_n}) + \cos x_n}{e^{x_n}(x_n + 1) + \sin x_n}$$

(i) For $n=0$,

$$\Rightarrow x_1 = \frac{x_0(\sin x_0 + x_0 e^{x_0}) + \cos x_0}{e^{x_0}(x_0+1) + \sin x_0}$$

$$\Rightarrow x_1 = \frac{0 + \cos 0}{e^0(1) + 0} = \frac{1}{1} = \boxed{1}.$$

(ii) For $n=1$,

$$\Rightarrow x_2 = \frac{x_1(\sin x_1 + x_1 e^{x_1}) + \cos x_1}{e^{x_1}(x_1+1) + \sin x_1}$$

$$\Rightarrow x_2 = \frac{1(\sin 1^\circ + e) + \cos 1^\circ}{e(1+1) + \sin 1^\circ} = 0.684908.$$

(iii) For $n=2$,

$$\Rightarrow x_3 = \frac{x_2(\sin x_2 + x_2 e^{x_2}) + \cos x_2}{e^{x_2}(x_2+1) + \sin x_2}$$

$$\Rightarrow x_3 = \frac{0.6849[\sin(0.6849) + (0.6849)e^{0.6849}] + \cos x_2}{\exp(0.6849)[(0.6849+1)] + \sin(0.6849)}.$$

$\therefore x_3 = 0.57797$

(iv) For $n=3$,

$\Rightarrow x_4 = 0.56725$

(Substituting x_n values accordingly in the previous formula).

(v) For $n=4$,

$x_5 = 0.56712$

→ For the values of $n = 3$ and 4, x_4 and x_5 are almost equal or approximately equal.
Therefore, the real root of this equation is,

$$\boxed{x = 0.5671}$$

1.4) Rate of Convergence of Newton-Raphson's Method:-

- Rate of Convergence: It is a measure of how fast the difference between the solution point and its estimates goes to zero. By using iterative methods is approximate to close the exact root.

★ Show that the Rate of Convergence of Newton's method is always 'quadratic'? (IMPORTANT)

Ans: Let, x_k be the kth iterate root and exact root be α, then error will be ϵ_k. So,

$\Rightarrow x_k = \epsilon_k + \alpha$.

or,

$\Rightarrow x_n = \epsilon_n + \alpha$ —— ① $\Rightarrow x_{n+1} = \epsilon_{n+1} + \alpha$ —— ②

- From Newton's formula,

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

$$\Rightarrow \epsilon_{n+1} + \cancel{\alpha} = \epsilon_n + \cancel{\alpha} - \frac{f(\epsilon_n + \alpha)}{f'(\epsilon_n + \alpha)}$$ [From equations ① and ②]

$$\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{f(\epsilon_n + \alpha)}{f'(\epsilon_n + \alpha)}$$

- From taylor series expansion,

$$\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{f(\alpha) + \epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2!} f''(\alpha) + \dots \infty}{f'(\alpha) + \epsilon_n f''(\alpha) + \frac{\epsilon_n^2}{2!} f'''(\alpha) + \dots \infty}$$

Since, 'α' is an exact root, So,

$$f(\alpha) = 0$$

$$\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{0 + \epsilon_n f'(\alpha)}{f'(\alpha) + \epsilon_n f''(\alpha)}$$

(further terms are neglected as, $\epsilon_n^2 <<< 0$).

$$\Rightarrow \epsilon_{n+1} = \frac{\cancel{\epsilon_n f'(\alpha)} + \epsilon_n^2 f''(\alpha) - \cancel{\epsilon_n f'(\alpha)}}{f'(\alpha) + \epsilon_n f''(\alpha)}$$

$$\Rightarrow \epsilon_{n+1} = \frac{\epsilon_n^2 f''(\alpha)}{f'(\alpha)} \times \frac{1}{\left(1 + \epsilon_n \frac{f''(\alpha)}{f'(\alpha)}\right)}$$

$$\Rightarrow \epsilon_{n+1} = \frac{\epsilon_n^2 f''(\alpha)}{f'(\alpha)} \times \left(1 + \epsilon_n \frac{f''(\alpha)}{f'(\alpha)}\right)^{-1}$$

$$\Rightarrow \epsilon_{n+1} = \frac{\epsilon_n^2 f''(\alpha)}{f'(\alpha)} \left(1 + \epsilon_n \frac{f''(\alpha)}{f'(\alpha)}\right).$$

$\because (1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 \dots \infty$

but, $\epsilon_n^2 <<< 0 \Rightarrow$ so, the term is neglected.

$$\Rightarrow \epsilon_{n+1} = \frac{\epsilon_n^2 f''(\alpha)}{f'(\alpha)} + \epsilon_n^3 \left(\frac{f''(\alpha)}{f'(\alpha)}\right)^2$$

$\Rightarrow \epsilon_{n+1} = \epsilon_n^2 \left(\frac{f''(\alpha)}{f'(\alpha)} \right)$ ($\because \epsilon_n^3 <<< 0$, so second term is neglected)

- Taking modulus on both sides,

$\therefore$ $$\boxed{|\epsilon_{n+1}| = M |\epsilon_n|^2} \rightarrow \text{Quadratic.}$$

(Hence prooved)

where,
$$M = \frac{f''(\alpha)}{f'(\alpha)}$$
(M is a function in terms of the term = 'α').

1.5) SECANT METHOD / CHORD METHOD →

- This method is quite similar to Regula-Falsi method except for the condition,
$$f(a) \cdot f(b) < 0$$
It is not necessary for this condition to be true always.

- It can be derived from Newton-Raphson's method.

- General formula that will be used is, (no proof needed)

$$\boxed{x_{n+1} = \frac{x_{n-1} \cdot f(x_n) - x_n \cdot f(x_{n-1})}{f(x_n) - f(x_{n-1})}}$$

where,
$n = 0, 1, 2, 3, 4, \ldots\ldots$ so on

and,
→ two points are always required to find the other iteration, i.e., if;
x_0 and x_1 are known then, x_2 can be calculated.

- Method fails if $f(x_n) = f(x_{n-1})$.

Q.6) Find the real root of the given function $f(x)$ in the interval $[1, 2]$ by using Secant's method ?

$$f(x) = x^3 - 5x + 3$$

Ans: From the given interval,

$x_0 = 1$ and, $x_1 = 2$

so,

$$f(x_0) = f(1) = 1 - 5 + 3 = -1$$

and,

$$f(x_1) = f(2) = 8 - 10 + 3 = +1.$$

Now, $x_0 = 1$, $x_1 = 2$; $f(x_0) = -1$, $f(x_1) = +1$.

- Applying secant's formula,

$$\Rightarrow \quad x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}$$

(i) For $n = 1$,

$$\Rightarrow \quad x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)}$$

$$\Rightarrow \quad x_2 = \frac{1 \times 1 - 2 \times -1}{1 - (-1)} = \frac{1 + 2}{2} = 1.5$$

so,

$$f(x_2) = f(1.5) = -1.125 < 0.$$

(ii) For $n = 2$,

$$\Rightarrow \quad x_3 = \frac{x_1 f(x_2) - x_2 f(x_1)}{f(x_2) - f(x_1)}$$

$$\Rightarrow x_3 = \frac{2 \times -1.125 - 1.5 \times 1}{-1.125 - 1} = 1.76470$$

So,

$$f(x_3) = f(1.7647) = -0.32793 < 0.$$

(iii) For $n = 3$,

$$\Rightarrow x_4 = \frac{x_2 f(x_3) - x_3 f(x_2)}{f(x_3) - f(x_2)}$$

$$\Rightarrow x_4 = \frac{1.5 \times -0.3279 - 1.7647 \times -1.125}{-0.3279 - (-1.125)}$$

$$\Rightarrow x_4 = 1.87358$$

so,

$$f(x_4) = f(1.8735) = 0.20848 > 0$$

(iv) For $n = 4$,

$$\Rightarrow x_5 = \frac{x_3 f(x_4) - x_4 f(x_3)}{f(x_4) - f(x_3)}$$

$$\Rightarrow x_5 = \frac{1.7647 \times 0.2084 - 1.8735 \times -0.3279}{0.2084 - (-0.3279)}$$

$$\Rightarrow x_5 = 1.83122 \quad \text{and,} \quad f(x_5) = -0.01544 < 0$$

(v) For $n = 5$,

$$\Rightarrow x_6 = \frac{x_4 f(x_5) - x_5 f(x_4)}{f(x_5) - f(x_4)} = 1.83411$$

and, $f(x_6) = -0.000729 < 0$

(The first 'three' zeros indicate above in $f(x_6)$ after decimal that the iteration 6 is correct upto first three decimal places).

$\therefore$ The real root of $f(x)$ is : $\boxed{x = 1.834}$

Q.7) Find the real root of the given transcedental equation by using Secant method, correct upto first four decimal places ?

$$f(x) = xe^{x} - 1$$

Ans : • To find the interval at first, by hit and trial :

For,

$x = 0$, $f(0) = 0 - 1 = -1 < 0$

$x = 1$, $f(1) = e - 1 = +1.718 > 0$

• Required interval is $[0, 1]$

now,

$x_0 = 0$, $x_1 = 1$; $f(x_0) = -1$, $f(x_1) = 1.718$.

→ From Secant's formula,

$$x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}$$

(i) For $n=1$,

$$\Rightarrow x_2 = \frac{x_0 f(x_1) - x_1 f(x_0)}{f(x_1) - f(x_0)} = \frac{0+1}{1.718+1} = 0.3679$$

$$\Rightarrow f(x_2) = f(0.36791) = (0.3679)e^{(0.3679)} - 1$$

$$\rightarrow f(x_2) = -0.46849 < 0$$

(ii) For $n=2$,

$$\Rightarrow x_3 = \frac{x_1 f(x_2) - x_2 f(x_1)}{f(x_2) - f(x_1)}$$

$$\Rightarrow x_3 = \frac{1 \times -0.4684 - 0.3679 \times 1.718}{-0.4684 - 1.718}$$

$$\Rightarrow x_3 = 0.50331$$

So,

$$f(x_3) = f(0.5033) = -0.16745 < 0$$

(iii) For $n=3$,

$$\Rightarrow x_4 = \frac{x_2 f(x_3) - x_3 f(x_2)}{f(x_3) - f(x_2)} = 0.57860$$

So,

$$f(x_4) = f(0.5786) = 0.03195 > 0$$

(iv) For $n=4$,

$$\Rightarrow x_5 = \frac{x_3 f(x_4) - x_4 f(x_3)}{f(x_4) - f(x_3)} = 0.55654$$

So,

$$f(x_5) = f(0.5665) = -0.001776 < 0$$

(v) For $n=5$,

$$\Rightarrow x_6 = \frac{x_4 f(x_5) - x_5 f(x_4)}{f(x_5) - f(x_4)} = 0.56711$$

so,

$$f(x_6) = f(0.5671) = -0.000196 < 0.$$

(vi) For $n = 6$,

$$\Rightarrow x_7 = \frac{x_5 f(x_6) - x_6 f(x_5)}{f(x_6) - f(x_5)} = 0.56717$$

so,

$$f(x_7) = f(0.56717) = 0.0000780 > 0$$

$\therefore$ The real root of this given equation (correct upto first four decimal places) is,

$$\boxed{x = 0.5671}$$

1.6) Order and Rate of Convergence of Secant Method :-

★ i) Show that the rate of convergence of Secant method is Super-Linear?

(SKIP) ii) Further, proove that the order of convergence of Secant method is a golden ratio number, i.e., 1.62? (IMPORTANT)

Ans = Let x_k be the kth iterate root and the exact root be α. Then the error will be ϵ_k. So,

$$\Rightarrow x_k = \epsilon_k + \alpha$$

or,

$$\Rightarrow x_n = \epsilon_n + \alpha \quad \text{——— (1)}$$

or,

$$\Rightarrow x_{n-1} = \epsilon_{n-1} + \alpha. \quad \text{——— (2)}$$

★ From Secant formula,

$$\Rightarrow x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}.$$

- It can be also written as,

$$\Rightarrow x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}$$

→ From equations ① and ②,

$$\Rightarrow \epsilon_{n+1} + \cancel{\alpha} = \epsilon_n + \cancel{\alpha} - \frac{f(\epsilon_n + \alpha)[\epsilon_n + \cancel{\alpha} - \epsilon_{n+1} - \cancel{\alpha}]}{f(\epsilon_n + \alpha) - f(\epsilon_{n-1} + \alpha)}.$$

$$\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{f(\epsilon_n + \alpha)(\epsilon_n - \epsilon_{n-1})}{f(\epsilon_n + \alpha) - f((\epsilon_{n-1}) + \alpha)}$$

- From taylor series expansion,

$$\left(\because f(x+h) = f(x) + h f'(x) + \frac{h^2}{2!} f''(x) + \dots\right).$$

$$\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1})\left(f(\alpha) + \epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2!} f''(\alpha) \dots\right)}{\left(f(\alpha) + \epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2!} f''(\alpha)\right) - \left(f(\alpha) + \frac{\epsilon_{n-1}}{2!} f''(\alpha) + \epsilon_{n-1} f'(\alpha)\right)}$$

$\because$ 'α' is an exact root so, $f(\alpha) = 0$.
On RHS,

$$= \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1})\left(0 + \epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2} f''(\alpha) + \dots\right)}{\left(0 + \epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2} f''(\alpha) + \dots\right) - \left(0 + \epsilon_{n-1} f'(\alpha) + \frac{\epsilon_{n-1}^2}{2} f''(\alpha) + \frac{\epsilon_{n-1}^3}{6} f'''(\alpha) + \dots\right)}$$

$$= \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1})\left(\epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2} f''(\alpha) + \dots\right)}{(\epsilon_n - \epsilon_{n-1}) f'(\alpha) + (\epsilon_n^2 - \epsilon_{n-1}^2) \frac{f''(\alpha)}{2}}$$

(higher terms are ignored)

$$= \epsilon_n - \frac{\cancel{(\epsilon_n - \epsilon_{n-1})}\left(\epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2} f''(\alpha) + \dots\right)}{\cancel{(\epsilon_n - \epsilon_{n-1})}\left[f'(\alpha) + (\epsilon_n + \epsilon_{n-1}) \frac{f''(\alpha)}{2}\right]}$$

$$= \epsilon_n - \frac{\epsilon_n f'(\alpha) + \frac{\epsilon_n^2}{2} f''(\alpha)}{f'(\alpha) + (\epsilon_n + \epsilon_{n-1}) f''(\alpha)}$$

$$= \epsilon_n - \frac{\cancel{f'(\alpha)}\left[\epsilon_n + \frac{\epsilon_n^2}{2} \frac{f''(\alpha)}{f'(\alpha)}\right]}{\cancel{f'(\alpha)}\left[1 + \frac{(\epsilon_n + \epsilon_{n-1})}{2} \frac{f''(\alpha)}{f'(\alpha)}\right]}$$

$$= \epsilon_n - \left(\epsilon_n + \frac{\epsilon_n^2}{2} \frac{f''(\alpha)}{f'(\alpha)}\right)\left(1 + \frac{(\epsilon_n + \epsilon_{n-1})}{2} \frac{f''(\alpha)}{f'(\alpha)}\right)^{-1}$$

- Using expansion : $(1+x)^{-n} = 1 - nx + nx^2 \dots \infty$

$$= \epsilon_n - \left(\epsilon_n + \frac{\epsilon_n^2 f''(\alpha)}{2 f'(\alpha)}\right)\left(1 - \frac{(\epsilon_n + \epsilon_{n+1}) f''(\alpha)}{2 f'(\alpha)}\right)$$

(ignoring higher terms as, $(\epsilon_n + \epsilon_{n+1})^2 <<< 0$)

$$= \epsilon_n - \left(\epsilon_n + \epsilon_n^2 \frac{f''(\alpha)}{2 f'(\alpha)}\right)\left(1 - \frac{\epsilon_n f''(\alpha)}{2 f'(\alpha)} - \frac{\epsilon_{n-1} f''(\alpha)}{2 f'(\alpha)}\right).$$

$$= \epsilon_n - \left[\epsilon_n - \cancel{\frac{\epsilon_n^2 f''(\alpha)}{2 f'(\alpha)}} - \frac{\epsilon_n . \epsilon_{n+1} f''(\alpha)}{2 f'(\alpha)} + \cancel{\frac{\epsilon_n^2 f''(\alpha)}{2 f'(\alpha)}} \right.$$

$$\left. \underbrace{- \epsilon_n^3 \left(\frac{f''(\alpha)}{2 f'(\alpha)}\right)^2}_{\text{neglected}} \underbrace{- \epsilon_n^2 . \epsilon_{n-1} \left(\frac{f''(\alpha)}{2 f'(\alpha)}\right)^2}_{\text{neglected}} \right]$$

$$= \cancel{\epsilon_n} - \cancel{\epsilon_n} + \frac{\epsilon_n . \epsilon_{n-1} f''(\alpha)}{2 f'(\alpha)}$$

$$= \epsilon_n . \epsilon_{n-1} \frac{f''(\alpha)}{2 f'(\alpha)}.$$

We get,

$$\Rightarrow \epsilon_{n+1} = \epsilon_n . \epsilon_{n-1} . \frac{f''(\alpha)}{2 f'(\alpha)}$$

→ Taking modulus on both sides and assuming,

$$C = \frac{f''(\alpha)}{2 f'(\alpha)}.$$

$$\boxed{|\epsilon_{n+1}| = C\,|\epsilon_n|\,|\epsilon_{n-1}|} \quad \text{—— (3)}$$

→ Now, general equation of rate of convergence is,

$$|\epsilon_{n+1}| \leq C|\epsilon_n|^q \quad \text{—— (5)} \quad (q: \text{a constant}).$$

- Replacing 'n' by 'n-1' in above equation,

⇒ $\epsilon_{(n-1)+1} \leq C(\epsilon_{(n-1)})^q$

⇒ $\epsilon_n \leq C.(\epsilon_{n-1})^q$

⇒ $\frac{\epsilon_n}{C} \leq (\epsilon_{n-1})^q$

- Taking q^{th} root on both sides,

⇒ $\sqrt[q]{\frac{\epsilon_n}{C}} \leq \epsilon_{n-1}$

or,

⇒ $\epsilon_{n-1} \geq \epsilon_n^{1/q} . C^{-1/q}$ —— (4)

- Substituting the value of ϵ_{n-1} from eq-(4) in equation (3),

⇒ $\frac{\epsilon_{n+1}}{C.\epsilon_n} \geq \epsilon_n^{1/q} . C^{-1/q}$

⇒ $\epsilon_{n+1} \geq \epsilon_n^{1+1/q} . C^{1-1/q}$

- From equation-(5),

⇒ $C.\epsilon_n^q \geq \epsilon_n^{1+1/q} . C^{1-1/q}$

→ On comparing both sides we have,

$\Rightarrow 1+\frac{1}{q} = q$ or, $\Rightarrow 1+\left(-\frac{1}{q}\right) = 1$.

(this case is rejected as $q \neq 0$. If $q = 0$ then the Inequality will not hold true).

So,

$\Rightarrow q+1 = q^2$

or,

$\Rightarrow q^2 - q - 1 = 0$.

$\Rightarrow q = \frac{+1 \pm \sqrt{1-4(1)(-1)}}{2}$ (By Discriminant method)

$\Rightarrow q = \frac{1 \pm \sqrt{5}}{2}$

$\therefore q = \frac{1+\sqrt{5}}{2}$ $\left(q \neq \frac{1-\sqrt{5}}{2}\right.$ as, $\left.q > 0\right)$.

or,

$$\boxed{q = 1.62}$$

Putting the value of 'q' in equation - (5),

$$\boxed{|\epsilon_{n+1}| \leq C\,|\epsilon_n|^{1.62}} \quad \left(C = \frac{f''(\alpha)}{2f'(\alpha)}\right).$$

→ Golden ratio number

This shows that,

1) Order of Convergence is 1.62 for Secant method.

2) Rate of Convergence is "Super-Linear", i.e., it is not necessarily true that this method will always show convergence. It may fails at some times.

2) Review of Taylor's Series, Errors and Approximations :-

2.1) Review of Taylor Series :-

- Taylor series expansion for a function $f(x)$ at $x = x_0$ is,

$$\Rightarrow f(x) = f(x_0) + f'(x_0).(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \frac{f'''(x_0)}{3!}(x - x_0)^3 + \dots \infty$$

- Taylor series expansion for a function $f(x+h)$ where, $h \approx 0$;

$$\Rightarrow f(x+h) = f(x) + h.f'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \dots \infty$$

★ **Special Expansions :-**

1) $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!}$

2) $(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - x^5 + \dots \infty$

3) $(1-x)^{-1} = 1 + x + x^2 + x^3 + x^4 + x^5 + \dots \infty$

4) $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} + \dots \infty$

5) $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} + \dots \infty$

Note: Expansions ①, ② and ③ are the MOST Important.

Examples: Compute the following values by using Taylor Series?

1) $\ln(1.1) = ?$

Ans: Taylor series expansion for $\ln(1.1)$ is,

$$\Rightarrow \ln(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \frac{(x-1)^5}{5}$$

For $x = 1.1$,

$$\Rightarrow \ln(1.1) = 0.1 - \frac{0.01}{2} + \frac{0.001}{3} - \frac{0.0001}{4} + \frac{0.00001}{5}$$

On solving,

$$\boxed{\ln(1.1) \approx 0.095310333....}$$

2) $e^8 = ?$

Ans: Taylor series expansion for e^x is,

$$\Rightarrow e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + ...$$

For $x = 8$,

$$\Rightarrow e^8 = 1 + 8 + \frac{64}{2} + \frac{512}{6} + \frac{4096}{24} + \frac{32768}{120} + ...$$

On solving,

$$\boxed{e^8 = 570.066665}$$

Q.8) What is the Taylor series of the function:

$$f(x) = 3x^5 - 2x^4 + 15x^3 + 13x^2 - 12x - 5$$

at the point $x_0 = 2$?

Ans: Taylor series expansion for a function $f(x)$ at $x = x_0$ is,

$$\Rightarrow f(x) = f(x_0) + (x-x_0) f'(x_0) + \frac{(x-x_0)^2}{2!} f''(x_0) + \frac{(x-x_0)^3}{3!} f'''(x_0) + \frac{(x-x_0)^4}{4!} f''''(x_0) + \ldots\ldots$$

Since,

$\Rightarrow$ given $x_0 = 2$. So,

$\Rightarrow f(x_0) = f(2) = 207$ (Substitute $x_0 = 2$ in function $f(x)$).

Now,

$\Rightarrow f'(x_0) = 15x^4 - 8x^3 + 45x^2 + 26x - 12$, $f'(2) = 396$

$\Rightarrow f''(x_0) = 60x^3 - 24x^2 + 90x + 26$, $f''(2) = 590$

$\Rightarrow f'''(x_0) = 180x^2 - 48x + 90$, $f'''(2) = 714$

$\Rightarrow f''''(x_0) = 360x - 48$, $f''''(2) = 672$

$\Rightarrow f'''''(x_0) = 360$, $f^5(2) = 360$.

$\Rightarrow f^6(x_0) = 0$, $f^6(2) = 0$.

$$\therefore f(x) = 207 + (x-2)(396) + \frac{(x-2)^2}{2} \times 590 + \frac{(x-3)^3}{6} \times 714 + \frac{(x-4)^4}{24} \times 672 + \frac{(x-5)^5}{120} \times 360$$

- Hence, approximate function is,

$$\therefore f(x) = 207 + 396(x-2) + 295(x-2)^2 + 119(x-2)^3 + 28(x-2)^4 + 3(x-2)^5.$$

2.2) Significant Digits and Rounding off :-

i) **Significant Digits :** The significant figures of a quantity are those starting from the first non-zero digit accurately on the left to the end of last digit specified on the right.

Examples	Number of Significant Figures	
1) 58	2	
2) 603	3	
3) 1008	4	
4) 95080	5	
5) 7.0	2	
6) 0.80	2	
7) 18.00	4	
8) 0.0003	1	
9) 0.0090	2	
10) 3.20	3	
11) 0.6046	4	
12) 2.134	4	
	(As whole) ↓	(After Decimal) ↓
13) 0.003	1	3
14) 3.40	3	1
15) 0.9999	4	4
16) 2.137	4	3
17) 2.0137	5	4
18) 2.1370	5	3
19) 2.13701	6	5
20) 1000.004	7	3
	↑ In total, the number of significant figures	↑ Only after the decimal.

ii) Rounding Off Rules :-

→ To round off a number to 'n' significant digits, discard all digits to right of the nth digit, if the $(n+1)^{th}$ digit is :

Serial No.	RULES	Examples	Rounded off values to first three decimal places ($n=3$)
1)	Less than 5 : Leave the nth digit unchanged.		
		39.72431	39.724
		4.00342	4.003
2)	Greater than 5 : Increase the nth digit by 1.	79.87293	79.873
		79.87893	79.879
		3.567	3.57
3)	Exactly 5 : Increase the nth digit by 1 if it is odd, otherwise, leave it unchanged.	2.589951	2.5900
		59.4865	59.486
		47.6975	47.698
		50.0695	50.070
		6.3456	6.346
		9.864651	9.865

2.3) Errors and Approximations in Numerical Computations:-

Serial No.	Type of Error	Notation	Formula
1)	Absolute Error	E_a	$E_a = \lvert X - X' \rvert$.
2)	Relative Error	E_R	$E_R = \left\lvert \frac{X - X'}{X} \right\rvert = \frac{E_a}{\lvert X \rvert}$
3)	Percentage Error	E_P	$E_P = 100\, E_R = \frac{100\, E_a}{\lvert X \rvert}$
4)	Truncation Error	E_T	$E_T = X - X'$.

Here, $X \rightarrow$ True Value of a Quantity.
$X' \rightarrow$ Approximate Value of a Quantity.

Note: 1) Truncation Errors: Caused by using approximate results or on replacing an infinite process by a finite ones. For example,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \infty = X \text{ (say)}$$

is replaced by,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} = X' \text{ (say)}.$$

So,

$$\text{Truncation Error} = \boxed{X - X'}.$$

2) If the first significant digit of a number is K, then the relative error is less than,

$$E_R < \frac{1}{K \times 10^{n-1}}$$

where, n : any number correct upto 'n' significant figures.

Example: For the given number (N),

$$N = 342.67392$$

then the relative error will be always less than for what value of N ?

Ans : Total significant figures in 'N' $\Rightarrow n = 8$

First significant figure's value $\Rightarrow K = 3$.

So,

$$\Rightarrow E_R < \frac{1}{K \times 10^{n-1}} = \frac{1}{3 \times 10^{8-1}} = \frac{1}{3} \times 10^{-7} = 0.000005$$

therefore,

$$\boxed{E_R < 0.000005}$$

3) Error in a function $f(x)$:-

Let,

$\Rightarrow y = f(x)$

Suppose, 'x' be true value and Δx be absolute error in x

$\Rightarrow E_A = \Delta x$.

and,

$\Rightarrow E_R = \Delta x / x$

$\Rightarrow E_P = \frac{\Delta x}{x} \times 100\%$

and,

$$\boxed{\Delta y = \frac{dy}{dx} \cdot \Delta x}$$

here,

Δy = absolute error in 'y'.

Q.9) If $y(x) = 7x^7 + 3x^3$, find percentage error in y when $x = 1$ for absolute error in x given as, $\Delta x = 0.05$?

Ans: Given,

$\Rightarrow y = 7x^7 + 3x^3$. and, $y(1) = 7 + 3 = 10$.

$\Rightarrow \frac{dy}{dx} = 49x^6 + 9x^2$

$\Rightarrow \left.\frac{dy}{dx}\right|_{\text{at } x=1} = 49 + 9 = 58$

Now,

$\Rightarrow \Delta y = \frac{dy}{dx} \times \Delta x$.

$\Rightarrow \left.\Delta y\right|_{\text{at } x=1} = \left(\frac{dy}{dx}\right)_{\text{at } x=1} \times \Delta x$.

$= 58 \times 0.05$

$\therefore \Delta y = \boxed{2.9}$

- Now, percentage error in $y = \frac{\Delta y}{y} \times 100\%$

$\Rightarrow E_p = \frac{2.9}{\cancel{10}} \times 10\cancel{0}\%$

$\therefore \boxed{E_p = 29\%}$.

Note: If $u = f(x_1, x_2, x_3, \ldots, x_n)$ be a function of several variables and Δx be the error in x, where $i = 1, 2, 3, 4, \ldots, n$. Then the error (Δu) in x is given by,

$$\Delta u = \frac{\partial f}{\partial x_1}\Delta x + \frac{\partial f}{\partial x_2}\Delta x + \frac{\partial f}{\partial x_3}\Delta x + \ldots$$

Date

Q.10) If, $u(x,y,z) = xyz^2 + x^2y^2z$ and error for each parameters x, y and z is 0.001 at $x=2$, $y=4$ and $z=5$.
Compute the maximum relative error (E_R) in u ?

Ans: Given,

$\Rightarrow$ $u = xyz^2 + x^2y^2z$

also,

$\Delta x = \Delta y = \Delta z = 0.001$ for $\begin{cases} x = 2 \\ y = 4 \\ z = 5 \end{cases}$

$\Rightarrow$ $u(2,4,5) = (2)(4)(5)^2 + (2)^2(4)^2(5)$

$\Rightarrow$ $u_T = 200 + 320$

$\therefore$ $u_T = 520$. —— (1)

Now,

$\Rightarrow$ $\Delta u = yz^2\Delta x + xz^2\Delta y + 2xyz\Delta z + 2xy^2z\Delta x + 2x^2yz\Delta y + x^2y^2\Delta z$.

$\Rightarrow$ $\Delta u = (yz^2 + 2xy^2z)\Delta x + (xz^2 + 2x^2yz)\Delta y + (2xyz + x^2y^2)\Delta z$.

$\Rightarrow$ $\Delta u = (4\times25 + 2\times4\times16\times5)0.001 + (2\times25 + 2\times4\times4\times5)\times0.001 + (2\times2\times4\times5 + 4\times16)\times0.001$.

$\therefore$ $\Delta u = 0.774$ —— (2) (on solving).

Now, relative error;

$\Rightarrow$ $E_R = \dfrac{\Delta u}{u_T}\times100\% = \dfrac{0.774}{520}\times100 = \boxed{0.149\%}$

Q.11) ① Find E_p in the numerical computation of $f(z)$,
$$f(z) = Z(x, y) = x - y$$
for $x = 12.05$ and $y = 8.02$ having absolute errors $\Delta x = 0.005$ and $\Delta y = 0.006$.

Ans: Let, $Z = x - y$. So,

$\Rightarrow Z_T = 12.05 - 8.02 = 4.03$

and,

$\Rightarrow \Delta Z = \Delta x - \Delta y$

$\Rightarrow \Delta Z = 0.005 - 0.006$

$\Rightarrow |\Delta Z| = 0.001$ ($\because$ Error can't be negative).

- So, percentage error (E_p) is,

$$\Rightarrow E_p = \frac{|\Delta Z|}{|Z_T|} \times 100\%$$

$$\Rightarrow E_p = \frac{0.001}{4.03} \times 100\%$$

On solving,
$$\boxed{E_p = 0.024813\ \%}$$

② Find the relative error if the number $X = 0.004997$ is truncated to three decimal digits?

Ans: Given, $X = 0.004997$

or, $X = 4.997 \times 10^{-3} = 0.4997 \times 10^{-2}$.

After truncating to three decimal digits,
$$X' = 0.499 \times 10^{-2}$$

$$\therefore E_R = \left|\frac{X - X'}{X}\right| = \left|\frac{0.4997 \times 10^{-2} - 0.499 \times 10^{-2}}{0.4997 \times 10^{-2}}\right|$$

hence,
$$\boxed{E_R = 0.140 \times 10^{-2}}$$

3) Unconstrained One Variable Function Minimization :-

3.1) Fibonacci Search Method :- [IMPORTANT]

★ **Working Procedure** →

- Consider a function $f(x)$ defined in an interval $[a, b]$.

Step 1 : Calculate L_0.

$\Rightarrow L_0 = b - a$ for a finite interval.

Step 2 : Calculate x_1 (iteration 1) and x_2 (iteration 2) as :

$x_1 = a + L_2^* \rightarrow a + L_2^*$

$x_2 = b - L_2^*$

Step 3 : Here, $L_k^* = \frac{F_{n-k}}{F_n} \times L_0$ (n : no. of iterations specified in the question)

so, and,

$L_2^* = \frac{F_{n-2}}{F_n} \times L_0$ ($K \rightarrow$ Iteration to be calculated : $K < n$).

Note : Fibonacci Series is defined as :

$$F_n = F_{n-1} + F_{n-2} \quad (n = 2, 3, 4, 5, \ldots \infty)$$

and,

$$F_0 = F_1 = 1$$

i.e.,

Series →	0,	1,	1,	2,	3,	5,	8,	13,	21,	34, ...
Term no. →		F_0	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8

Step 4 : Now, find $F(x_1)$ and $F(x_2)$ and compare both the values :

i) If $F(x_1) > F(x_2)$:
It means function is maximized at x_1 and is minimum at x_2.

- Discard the interval $[a, x_1]$, if you want to minimize the function.
- And, discard the interval $[b, x_2]$, if you want to maximize the function.

ii) If $F(x_1) < F(x_2)$:
It means function is maximized at x_2 and is minimized at x_1.
→ Follow the above approach but opposite, i.e.,

- Discard the interval $[b, x_2]$, if you want to minimize the function.
- Discard the interval $[a, x_1]$, if you want to maximize the function.

Step 5: Calculate the next iteration, i.e., (x_3).

- First, find L_3^*, using the previously mentioned formula,

$$L_3^* = \frac{F_{n-3}}{F_n} \times L_0 \qquad \left(\begin{array}{l} L_0 = b - a \text{ (from Step 1)} \\ K = 3 \end{array} \right).$$

- Now, find x_3 [considering Step 4: Case-1].

As,
x_3 should lie in the interval $[x_1, b]$ and,

could be before the mid-point or after the mid-point (x_2)

x_3 could be located only at one position either,

$x_3 = x_1 + L_3^*$ OR $x_3 = b - L_3^*$.

i) If, $x_3 \approx x_2$ in the interval $[x_1, x_2]$, so x_3 should must be located on the opposite side.

IF EQUAL (Locate x_3 on opposite side).

ii) Vice-Versa, would be followed if, $x_3 \approx x_2$ in the interval $[x_2, b]$, so swap x_3 on the opposite side.

Step 6: Similarly, as step 4, compare $F(x_3)$ and $F(x_2)$ and decide the next interval based on either condition of maximization or minimization.

Step 7: Calculate the next iteration, i.e., (x_4) and repeat steps 5 and step 4 till $\boxed{K = n}$ th iteration.

i.e.,

$$L_K^* = \frac{F_{n-K}}{F_n} \times L_0$$

→ At $K = n$,

$$L_n^* = \frac{F_0}{F_n} \times L_0$$

- this will be the last iteration to be calculated.

Q.12) For given unconstrained one variable function $y = f(x)$ where,

$$f(x) = 2x^2 - e^x$$

Minimize the function in the interval $[0, 1]$ for number of experiments / iterations $(n) = 10$?

(7.5 marks)

Ans: Given,

$f(x) = 2x^2 - e^x$, $n = 10$, $[0, 1]$

i) $L_0 = b - a = 1 - 0 = 1$

$\Rightarrow L_2^* = \frac{F_{n-2}}{F_n} \times L_0 = \frac{F_{10-2}}{F_{10}} \times 1$ $(n = 10)$

$\Rightarrow L_2^* = \frac{F_8}{F_{10}} = \frac{34}{89} = 0.382022$

So,

$\Rightarrow x_1 = a + L_2^* = 0 + 0.382022 = 0.382022$

$\Rightarrow x_2 = b - L_2^* = 1 - 0.382022 = 0.618$

$\Rightarrow F(x_1) = F(0.382022) = -1.17336$

$\Rightarrow F(x_2) = F(0.618) = -1.09136$

(substituted values of x_1 and x_2 in function $f(x)$).

$\because F(x_2) > F(x_1)$, for minimization discarding the max^n value and the interval $[x_2, 1]$.

$\rightarrow$ Rejecting the interval $[x_2, 1]$

2) To find x_3.

$\Rightarrow L_3^* = \frac{F_{n-3}}{F_n} \times L_0$

$\Rightarrow L_3^* = \frac{F_7}{F_{10}} \times 1 = \frac{21}{89} = 0.2359$

So,

$\Rightarrow x_3 = 0 + L_3^* = 0 + 0.2359 = 0.2359$ (✓)

$\Rightarrow x_3 = x_2 - L_3^* = 0.618 - 0.2359 = 0.3821 \approx x_1$

- The value of x_3 will be the ones, that is <u>not</u> approximately equal to x_1. So,

$$x_3 = 0.2359.$$

Now,

$\Rightarrow F(x_3) = F(0.2359) = 2 \times (0.2359)^2 - e^{(0.2359)}$

i.e.,

$$F(x_3) = -1.15475$$

and, $$F(x_1) = -1.17336.$$

$\because F(x_3) > F(x_1)$, rejecting the interval $[0, x_3]$.

3) To find x_4.

$\Rightarrow L_4^* = \frac{F_{n-4}}{F_n} \times L_0$

$\Rightarrow L_4^* = \frac{F_6}{F_{10}} \times 1 = \frac{13}{89} = 0.14606$

So,

$\Rightarrow x_4 = x_3 + L_4^* = 0.2359 + 0.14606 = 0.38196$

$\Rightarrow x_4 = x_2 - L_4^* = 0.618 - 0.14606 = 0.47194$ (✓)

$\rightarrow$ Value of $x_4 = 0.47194 \neq x_1$.

$\Rightarrow$ So, $F(x_4) = -1.157646$.

now, comparing $F(x_1)$ and $F(x_4)$,

$F(x_4) > F(x_1)$. $(F(x_1) = -1.157646)$.

$\rightarrow$ i.e., $F(x)$ is maximum at x_4. So,

$\rightarrow$ rejecting the Interval $[x_4, x_2]$.

4) To find x_5.

$\Rightarrow L_5^* = \frac{F_{n-5}}{F_n} \times L_0$

$\Rightarrow L_5^* = \frac{F_5}{F_{10}} \times 1 = \frac{F_5}{F_{10}} = \frac{8}{89} = 0.089887$

So,

$\Rightarrow x_5 = x_3 + L_5^* = 0.2359 + 0.089887 = 0.32578$ (✓)

$\Rightarrow x_5 = x_4 - L_5^* = 0.47194 - 0.089887 = 0.38206$.

$\therefore F(x_5) = F(0.32578) = -1.17284$.

and,

$F(x_1) = -1.17336$.

As, $F(x_5) > F(x_1)$, reject interval $[x_3, x_5]$ for minimizing the function.

5) To find x_6.

$\Rightarrow L_6^* = \frac{F_{n-6}}{F_n} \times L_0$

$\Rightarrow L_6^* = \frac{F_4}{F_{10}} \times 1 = \frac{5}{89} \times 1 = 0.056179$

Value of x_6 will be

$\Rightarrow x_6 = x_4 - L_6^* = 0.47194 - 0.056179 = 0.41577 \neq x_1$

So,

$\Rightarrow F(x_6) = F(0.41577) = -1.169807$.

→ As, $F(x_6) > F(x_1)$, so to minimize the function rejecting the maximized interval : $[x_6, x_4]$.

6) To find x_7.

$\Rightarrow L_7^* = \frac{F_{n-7}}{F_n} \times L_0$

$\Rightarrow L_7^* = \frac{F_3}{F_{10}} \times 1 = \frac{3}{89} = 0.033707$

So,

$\Rightarrow x_7 = x_5 + L_7^* = 0.32578 + 0.033707 = 0.359487 \neq x_1$

$\Rightarrow F(x_7) = F(0.359487) = -1.1741324$

and,

$\Rightarrow F(x_1) = -1.1733627$

$\because F(x_7) < F(x_1)$, for minimizing the function rejecting the maxima at x_1, i.e., interval $[x_1, x_6]$.

7) To find x_8.

$\Rightarrow L_8^* = \frac{F_{n-8}}{F_n} \times L_0$

$\Rightarrow L_8^* = \frac{F_2}{F_{10}} \times L_0 = \frac{2}{89} = 0.022471$

So,

$\Rightarrow x_8 = x_5 + L_8^* = 0.32578 + 0.022471 = 0.34825 \neq x_7$

$\Rightarrow F(x_8) = F(0.34825) = -1.1740302$

and,

$\Rightarrow F(x_7) = -1.1741324$

• So, rejecting the interval $[x_5, x_8]$ as, $F(x_8) > F(x_7)$.

8) To find x_9.

$\Rightarrow L_9^* = \frac{F_{n-9}}{F_n} \times L_0$

$\Rightarrow L_9^* = \frac{F_1}{F_{10}} \times 1 = \frac{1}{89} = 0.011235$

So,

$\Rightarrow x_9 = x_1 - L_9^* = 0.382022 - 0.011235 = 0.37079$

i.e.,

$\Rightarrow F(x_9) = F(0.37079) = -1.1739083$

$\because F(x_7) = -1.1741324.$

$\Rightarrow F(x_9) > F(x_7)$, so rejecting the interval $[x_9, x_1]$.

9) To find x_{10}.

$\Rightarrow L_{10}^* = \frac{F_{n-10}}{F_n} \times L_0$

$\Rightarrow L_{10}^* = \frac{F_0}{F_{10}} \times 1 = \frac{1}{89} = 0.011235$

So,

$\Rightarrow x_{10} = x_8 + L_{10}^* = 0.34825 + 0.011235 = 0.359485$

and,

$\Rightarrow F(x_{10}) = F(0.359485) = -1.17413251.$

$\Rightarrow F(x_7) = F(0.35948) = -1.1741324.$

$\rightarrow$ Here, values of x_{10} and x_7 show convergence, i.e., they are both equal.

Therefore, the value of x for which the function $F(x)$ is minimum is,

$$\boxed{x = 0.359485}$$

and,

$$\boxed{F(x_{min}) = F(0.359485) = -1.17413251}$$

Q.13) For the given unconstrained one variable function $y = f(x)$ where,

$$f(x) = x^2 + 4x + 5$$

Maximize the function in the interval $[-4, 0]$ for number of iterations $(n) = 6$, using the Fibonacci Search Method ?

Ans: Given,

$$f(x) = x^2 + 4x + 5 \quad , \quad n = 6 \quad , \quad [-4, 0].$$

1) $L_0 = b - a = 0 - (-4) = 4.$

$\Rightarrow L_2^* = \dfrac{F_{n-2}}{F_n} \times L_0$

$\Rightarrow L_2^* = \dfrac{F_4}{F_6} \times 4 \qquad (\because n = 6).$

$\Rightarrow L_2^* = \dfrac{5}{13} \times 4 = 1.53846$

So,

$\Rightarrow x_1 = a + L_2^* = -4 + 1.53846 = -2.46154$

$\Rightarrow x_2 = b - L_2^* = 0 - 1.53846 = -1.53846$

and,

$\Rightarrow F(x_1) = 1.212521$ ..... [Consider the values of x_1 and

$\Rightarrow F(x_2) = 1.213444$ x_2 only upto first three decimal places.]

$\because F(x_2) > F(x_1)$, so for maximizing the function reject the minima, i.e., x_1, rejecting the interval $[-4, x_1]$.

2) $L_3^* = \dfrac{F_{n-3}}{F_n} \times L_0 = \dfrac{F_3}{F_6} \times 4 = \dfrac{3}{13} \times 4 = 0.92307$

Now,

$\Rightarrow x_3 = 0 - L_3^* = 0 - .92307 = -0.92307$

$\Rightarrow$ $F(x_3) = F(-0.92307)$

$\Rightarrow$ $F(x_3) = (-0.92307)^2 + 4 \times (-0.92307) + 5$

$\therefore$ $F(x_3) = 2.1597782$

and,

$F(x_2) = 1.213444$

$\because$ $F(x_3) > F(x_2)$, so rejecting the interval $[x_1, x_2]$.

3) $L_4^* = \frac{F_{n-4}}{F_n} \times L_0$

$\Rightarrow$ $L_4^* = \frac{F_2}{F_6} \times 4$

$\Rightarrow$ $L_4^* = \frac{2}{13} \times 4 = 0.61538$

$\Rightarrow$ $x_4 = 0 - L_4^* = 0 - 0.61538 = -0.61538$.

so,

$\Rightarrow$ $F(x_4) = F(-0.61538) = 2.9171725$.

$\because$ $F(x_4) > F(x_3)$, so rejecting the interval $[x_2, x_3]$.

4) To find x_5.

$\Rightarrow$ $L_5^* = \frac{F_{n-5}}{F_n} \times L_0$

$\Rightarrow$ $L_5^* = \frac{F_1}{F_6} \times 4$

$\Rightarrow$ $L_5^* = \frac{1}{13} \times 4 = \cancel{0.6153}\ 0.30769$.

So,

$\Rightarrow$ $x_5 = 0 - L_5^* = -0.30769$.

and,

$\Rightarrow$ $F(x_5) = F(-0.30769) = 3.863913136$.

5) To find x_6.

$\Rightarrow L_6^* = \frac{F_{n-6}}{F_n} \times L_0$

$\Rightarrow L_6^* = \frac{F_0}{F_6} \times L_0 = \frac{1}{13} \times 4 = 0.30769$.

So,

$\Rightarrow x_6 = x_4 + L_6^* = -0.61538 + 0.30769 = -0.30769$.

or,

$\Rightarrow x_6 = 0 - L_6^* = 0 - 0.30769 = -0.30769$.

and,

value of x_5 was $= -0.30769$.

i.e.,

'x_5' and 'x_6' show convergence and are clearly equal.

→ Therefore, the maxima of the function is,

$$\boxed{x_{max} = -0.30769}$$

and,

→ The maximum value of the function is

$$F(x_{max}) = F(-0.30769) = \boxed{3.863913136}$$

3.2) Golden Section Method :- [IMPORTANT].

★ Working Procedure →

- Consider a function $f(x)$ defined in an interval $[a, b]$.

→ "Same procedure as Fibonacci Search Method, only change is in 'Step-3' → formula.."

Step-3: $L_n^* = \frac{1}{\gamma^n} \times L_0$; $L_0 = b - a$
$\gamma = 1.618$ (a constant).

Note: ① 'γ' is called Golden Ratio Number. It is a constant and its value is,

$$\gamma = \frac{1+\sqrt{5}}{2} = 1.618.$$

② If the value of 'n' is not given in the question, then calculate the value of x, till the number of iterations reach an error (ϵ) of 0.15 (approx.). That is,

$$\boxed{x \approx \epsilon = 0.15 \ldots\ldots \text{ (approximately)}}.$$

Q.14) Using the Golden Section Method, minimize the following given function $f(x)$,

$$f(x) = 4x^3 + x^2 - 7x + 14$$

in the interval $[0, 1]$ for $n = 6$?

Ans: Given,

$$f(x) = 4x^3 + x^2 - 7x + 14, \quad n = 6, \quad [0, 1].$$

1) $L_0 = b - a = 1 - 0 = 1$

$\Rightarrow L_2^* = \frac{1}{\gamma^2} \times L_0 = \frac{1}{(1.618)^2} \times 1 = 0.3819$

$\Rightarrow x_1 = a + L_2^* = 0 + 0.3819 = 0.3819$

$\Rightarrow x_2 = b - L_2^* = 1 - 0.3819 = 0.6181$

so,

$\Rightarrow F(x_1) = F(0.3819) = 11.6953$

$\Rightarrow F(x_2) = F(0.6181) = 10.9999$

As, $F(x_1) > F(x_2)$, so to minimize the function,

rejecting the maxima (x_1) and interval $[0, x_1]$.

2) $L_3^* = \frac{1}{\gamma^3} \times L_0$

$\Rightarrow L_3^* = \frac{1}{(1.618)^3} \times 1 = 0.2361$.

$\Rightarrow x_3 = 1 - L_3^* = 1 - 0.2361 = 0.7639$

so,

$\Rightarrow F(x_3) = 11.019317$ and, $F(x_2) = 10.9999$

$\because F(x_3) > F(x_2)$, so rejecting the interval $[x_3, 1]$.

3) $L_4^* = \frac{1}{\gamma^4} \times L_0$

$\Rightarrow L_4^* = \frac{1}{(1.618)^4} = 0.1459102$

so,

$\Rightarrow x_4 = x_1 + L_4^* = 0.52781$ (✓).

$\Rightarrow x_4 = x_3 - L_4^* = 0.61799$

so,

$\Rightarrow F(x_4) = 11.1720698$ and, $F(x_2) = 10.9999$.

$\because F(x_4) > F(x_2)$, so rejecting the interval $[x_1, x_4]$.

4) $L_5^* = \frac{1}{\gamma^5} \times L_0$

$\Rightarrow L_5^* = \frac{1}{(1.618)^5} = 0.0901794$

so,

$\Rightarrow x_5 = x_3 - L_5^* = 0.7639 - 0.090179 = 0.673721$

i.e.,

$\Rightarrow F(x_5) = 10.9610608$, and, $F(x_2) = 10.9999$

$\because F(x_2) > F(x_5)$, so rejecting the interval $[x_4, x_2]$.

5) For $n=6$,

$$L_6^* = \cancel{F_{n=6}}\ \frac{1}{\gamma^6} \times L_0 = \frac{1}{(1.618)^6} \times 1 = 0.055735$$

$\Rightarrow x_6 = x_3 - L_6^*$

$\Rightarrow x_6 = 0.7639 - 0.055735$

$\Rightarrow x_6 = 0.708165. \neq x_5.$

So,

$\Rightarrow F(x_6) = F(0.708165) = 10.964915$

$\Rightarrow F(x_5) = 10.9610608.$

$\because F(x_6) > F(x_5)$, so rejecting the interval $[x_6, x_3]$.

$\rightarrow$ Hence, the interval remains $[x_2, x_6]$. So, optimised value of x_{min} will be;

$\Rightarrow x_{min} = \dfrac{x_2 + x_6}{2}$ $\left(\because x_{opt} = \dfrac{a+b}{2} \text{ for } [a,b]\right)$.

$\Rightarrow x_{min} = \dfrac{0.6181 + 0.708165}{2} = \boxed{0.663325}$

So,

$\Rightarrow \boxed{F(x_{min}) = 10.964253 \sim F(x_5)}$

<u>Note</u>: If, number of iterations (n) and interval (L_0) is not given
(Important) for either Fibonaci Search and Golden Section method, but error (ϵ) is given then, assume, $L_0 = [0, 1]$ for both the methods and 'n' shall be calculated as;

1). <u>Fibonaci Search Method</u>: Assume, $L_0 = [0, 1]$

$\because$ 'ϵ' is given, but 'n' is not, so

★ Using the formula,

$$\boxed{\epsilon = \frac{L_n}{L_0}}$$

where,

① 'ϵ' is called as measure of efficiency.

② L_n = Length of interval of uncertainity after 'n' iterations

L_0 = Length of initial interval of uncertainity.

③ Ratio $\left(\frac{L_n}{L_0}\right)$ is called as "Reduction Ratio."

So, assume value of $L_n = \frac{1}{F_n} L_0$.

$\therefore \epsilon = \left(\frac{\cancel{L_0}}{F_n}\right) \times \frac{1}{\cancel{L_0}} \Rightarrow \boxed{F_n \geq \frac{1}{\epsilon}}$ {'ϵ' will be given. For fibonacci method}

→ Greater than or equal to.

→ When calculated 'F_n', compare this value to the fibonacci series to find the nearest value of 'n'. For example, if F_n comes out to be 377 then, $\boxed{n = 14}$.

Sequence (F_n)	1	1	2	3	5	8	13	21	34	55	89	144	233	377 ✓
No. of term (n)	1	2	3	4	5	6	7	8	9	10	11	12	13	$\boxed{14}$ ↑

2) <u>Golden Section Method</u>: Assume $L_0 = [0, 1]$.

Applying above formula,

$$\epsilon = \frac{L_n}{L_0}$$

So, in this case, assume $L_n = \frac{1}{\gamma^{n-1}} L_0$.

$\therefore \epsilon = \left(\frac{\cancel{L_0}}{\gamma^{n-1}}\right) \times \frac{1}{\cancel{L_0}} \Rightarrow \boxed{\epsilon = (0.618)^{n-1} = (1.618)^{1-n}}$

→ Use this formula, to solve and find 'n'.

4) Multivariate Two Variables Function Minimization :-

4.1) Steepest Descent Method :-

★ **Working Procedure :-**

Step 1 : Start with an arbitary initial point X_1, considering it as iteration 1 where,

$$X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{(for two variables)}$$

Step 2 : Find the Search Direction (S_n), where n: number of iteration,

$$S_n = -\nabla f_n = -\begin{bmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \end{bmatrix} X_n$$

- Example : If $n=1$,

$$S_1 = -\underset{\text{(Gradient Symbol)}}{\nabla} f_1 = -\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} X_1 \quad \text{or} \quad -\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \text{at} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Step 3 : Determine the optimal step Length (L_1^*). For minimizing the function,

① Find $f(X_1 + L_1^* S_1) = ?$ (for $n=1$)

② Now, differentiate this value w.r.t L_1^*, i.e., find

$$\Rightarrow \frac{d}{dL_1^*}\left[f(X_1 + L_1^* S_1)\right]$$

③ Put the above value equal to zero and find L_1^*.

$$\Rightarrow \frac{d}{dL_1^*}\left[f(X_1 + L_1^* S_1)\right] = 0 \quad \Rightarrow \boxed{L_1^* = ?}$$

Step 4: Now, determine the next arbitary point X_2 using the formula,

$$\boxed{X_2 = X_1 + L_1^* S_1}$$

Step 5: Find the next Search Direction (S_2) at point X_2.

$$S_2 = -\nabla f_2 = -\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_2} \text{ or } -\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} \text{ at } \begin{bmatrix} x_K \\ x_{K+1} \end{bmatrix}$$

Step 6: Check whether, $|\nabla f_2| \cong X_1$. [$\cong$: Approximately Equal].
If NOT,
then perform the next iteration, i.e., find X_3 then S_3 and compare it to X_1. If not approximately equal then repeat the procedure till it shows convergence.

Note: Break the iteration, if it does not show convergence or gets maximized instead of minimizing.

Q.15) Minimize the given function $f(x_1, x_2)$ where,

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$$

Starting from the point $X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ where, $x_1 = x_2 = 0$ using the Steepest Descent Method? (HARD)

Ans: Given, initial arbitary point $(X_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Now,

Iteration 1:

$$\Rightarrow \nabla f_1 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_1} = \begin{bmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{bmatrix} \text{ at } \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \nabla f_1 = \begin{bmatrix} 1+4(0)+2(0) \\ -1+2(0)+2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

hence,

$$\Rightarrow S_1 = -\nabla f_1 = -\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$\Rightarrow$ To find λ_1^* or L_1^*, minimizing the function,

$$\Rightarrow f(X_1 + L_1^* S_1) = f\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix} + L_1^* \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right]$$

$$\Rightarrow f\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -L_1^* \\ L_1^* \end{bmatrix}\right] = f\begin{bmatrix} -L_1^* \\ +L_1^* \end{bmatrix} = f(\underset{x_1}{\underset{\downarrow}{-L_1^*}}, \underset{x_2}{\underset{\downarrow}{L_1^*}})$$

$$\Rightarrow -L_1^* - L_1^* + 2(L_1^*)^2 + 2L_1^*(-L_1^*) + (L_1^*)^2$$

$$\Rightarrow -2L_1^* + \cancel{2L_1^{*2}} - \cancel{2L_1^{*2}} + L_1^{*2}$$

$$\Rightarrow L_1^{*2} - 2L_1^*$$

$\rightarrow$ Differentiating this value w.r.t L_1^* and equating to zero,

$$\Rightarrow \frac{d}{dL_1^*}\left[f(-L_1^*, L_1^*)\right] = \frac{d}{dL_1^*}(L_1^{*2} - 2L_1^*)$$

$\Rightarrow 2L_1^* - 2 = 0 \qquad \Rightarrow 2L_1^* = 2$

$\Rightarrow \boxed{L_1^* = 1}$.

So,

$\Rightarrow X_2 = X_1 + L_1^* S_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Iteration 2: Now,

$\Rightarrow \nabla f_2 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_2} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$ at $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

$\Rightarrow \nabla f_2 = \begin{bmatrix} 1 + 4(-1) + 2(1) \\ -1 + 2(-1) + 2(1) \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \neq X_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\because \nabla f_2 \neq X_1$, we have to continue with next iteration.

hence,

$S_2 = -\nabla f_2 = -\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$\rightarrow$ To find L_2^*, finding the function : $f(X_2 + L_2^* S_2)$.

$= f\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} + L_2^* \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) = f\left(\begin{bmatrix} -1 + L_2^* \\ 1 + L_2^* \end{bmatrix} \right) \begin{matrix} \rightarrow x_1 \\ \rightarrow x_2 \end{matrix}$

$\Rightarrow f\left([-1 + L_2^*], [1 + L_2^*] \right)$

$\Rightarrow (-1 + L_2^*) - (1 + L_2^*) + 2(-1 + L_2^*)^2 + 2(-1 + L_2^*)(1 + L_2^*)$

$+ (1 + L_2^*)^2$

$\Rightarrow -1 + \cancel{L_2^*} - 1 - \cancel{L_2^*} + 2(L_2^{*2} + 1 - 2L_2^*) + 2(L_2^{*2} - 1)$

$+ (1 + L_2^{*2} + 2L_2^*)$

$\Rightarrow -2 + 5L_2^{*2} - 2L_2^* + 1$

$\Rightarrow 5L_2^{*2} - 2L_2^* - 1$

$\rightarrow$ On differentiating w.r.t L_2^*;

$\Rightarrow \frac{d}{dL_2^*}(5L_2^{*2} - 2L_2^* - 1) = 10L_2^* - 2$

$\Rightarrow 10L_2^* - 2 = 0$ (for minimization)

$\Rightarrow L_2^* = \frac{2}{10}$

$\therefore \boxed{L_2^* = \frac{1}{5}}$

So,

$\Rightarrow X_3 = X_2 + L_2^* S_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{1}{5}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 + 0.2 \\ 1 + 0.2 \end{bmatrix}$

$\Rightarrow X_3 = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix}$

Now,

$\Rightarrow \nabla f_3 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{\text{at } X_3} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix}$ at $\begin{bmatrix} -0.8 \\ +1.2 \end{bmatrix}$.

$\Rightarrow \nabla f_3 = \begin{bmatrix} 1 + 4(-0.8) + 2(1.2) \\ -1 + 2(-0.8) + 2(1.2) \end{bmatrix}$

$\Rightarrow \nabla f_3 = \begin{bmatrix} 0.2 \\ -0.2 \end{bmatrix} \neq X_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

$\because$ $\nabla f_3 \neq X_1$, we have to continue with the next iteration.

Iteration 3: So,

$S_3 = -\nabla f_3 = \begin{bmatrix} -0.2 \\ +0.2 \end{bmatrix}$.

$\rightarrow$ Now, L_3^* shall be calculated as : $f(X_3 + L_3^* S_3) =$

$\Rightarrow f\left(\begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} + L_3^* \begin{bmatrix} -0.2 \\ 0.2 \end{bmatrix} \right)$

$\Rightarrow f\left(\left[-0.8 + (-0.2) L_3^* \right], \left[1.2 + (0.2) L_3^* \right] \right)$.

$\Rightarrow (-0.8 - 0.2L_3^*) - (1.2 + 0.2L_3^*) + 2(-0.8 - 0.2L_3^*)^2$

$+ 2(-0.8 - 0.2L_3^*)(1.2 + 0.2L_3^*) + (1.2 + 0.2L_3^*)^2$

$\Rightarrow -2 - 0.4L_3^* + 2(0.64 + 0.04L_3^{*2} + 0.32L_3^*)$.

$+ 2(-0.96 - 0.16L_3^* - 0.24L_3^* - 0.04L_3^{*2}) +$

$(1.44 + 0.04L_3^{*2} + 0.48L_3^*)$.

$\Rightarrow 0.04L_3^{*2} - 0.08L_3^* - 1.2$ (on solving..)

$\rightarrow$ Now, differentiating this value, and equating it to zero.

$\Rightarrow \frac{d}{dL_3^*}(0.04L_3^{*2} - 0.08L_3^* - 1.2) = 0$.

$\Rightarrow \quad 0.08 L_3^* - 0.08 = 0$

so, $\boxed{L_3^* = 1}$

$\Rightarrow \quad X_4 = X_3 + L_3^* S_3$

$$\Rightarrow \quad X_4 = \begin{bmatrix} -0.8 \\ 1.2 \end{bmatrix} + 1 \begin{bmatrix} -0.2 \\ 0.2 \end{bmatrix}$$

$$\Rightarrow \quad X_4 = \begin{bmatrix} -0.8 - 0.2 \\ 1.2 + 0.2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1.4 \end{bmatrix}$$

Now,

$$\Rightarrow \quad \nabla f_4 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_4} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \text{ at } \begin{bmatrix} -1 \\ 1.4 \end{bmatrix}$$

$$\Rightarrow \quad \nabla f_4 = \begin{bmatrix} 1 + 4(-1) + 2(1.4) \\ -1 + 2(-1) + 2(1.4) \end{bmatrix} = \begin{bmatrix} -0.2 \\ -0.2 \end{bmatrix} \neq X_1.$$

$\because \quad \nabla f_4 \neq X_1$, we have to continue with the next iteration.

<u>Iteration 4</u>: So,

$$S_4 = -\nabla f_4 = -\begin{bmatrix} -0.2 \\ -0.2 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}.$$

Now,

$$\Rightarrow \quad X_5 = X_4 + L_4^* S_4 = \begin{bmatrix} -1 \\ 1.4 \end{bmatrix} + L_4^* \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix}.$$

$$\Rightarrow \quad X_5 = \begin{bmatrix} -1 + 0.2 L_4^* \\ 1.4 + 0.2 L_4^* \end{bmatrix}.$$

→ Now, to find the value of L_4^*, first finding $f: (X_4 + L_4^* S_4)$

$$\Rightarrow f\left(\left[\underbrace{-1+0.2L_4^*}_{x_1}, \underbrace{1.4+0.2L_4^*}_{x_2}\right]\right).$$

$$\Rightarrow (-1+\cancel{0.2L_4^*}) - (1.4+\cancel{0.2L_4^*}) + 2(-1+0.2L_4^*)^2$$
$$+ 2(-1+0.2L_4^*)(1.4+0.2L_4^*) + (1.4+0.2L_4^*)^2.$$

$$\Rightarrow -2.4 + 2(0.04L_4^{*2} + 1 - 0.4L_4^*) + 2(-1.4 - 0.2L_4^*$$
$$+ 0.28L_4^* + 0.04L_4^{*2}) + (1.96 + 0.04L_4^{*2} + 0.56L_4^*)$$

$$\Rightarrow -2.4 + 0.08L_4^{*2} + 2 - 0.8L_4^* - 2.8 - 0.4L_4^* + 0.56L_4^*$$
$$+ 0.08L_4^{*2}) + (1.96 + 0.04L_4^{*2} + 0.56L_4^*).$$

$$\Rightarrow 0.2L_4^{*2} - 0.08L_4^* - 1.24.$$

→ On differentiating, w.r.t. L_4^*,

$$\Rightarrow \frac{d}{dL_4^*}(0.2L_4^{*2} - 0.08L_4^* - 1.24) = 0.$$

$$\Rightarrow 0.4L_4^* - 0.08 = 0$$

hence,

$$\boxed{L_4^* = 0.2}$$

so,

$$X_5 = \begin{bmatrix} -1+0.2L_4^* \\ 1.4+0.2L_4^* \end{bmatrix} = \begin{bmatrix} -1+0.04 \\ 1.4+0.04 \end{bmatrix} = \begin{bmatrix} -0.96 \\ 1.44 \end{bmatrix}$$

→ Now, evaluating ∇f_5:

$$\Rightarrow \nabla f_5 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_5} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \text{ at } \begin{bmatrix} -0.96 \\ 1.44 \end{bmatrix}$$

$$\Rightarrow \nabla f_5 = \begin{bmatrix} 1 + 4(-0.96) + 2(1.44) \\ -1 + 2(-0.96) + 2(1.44) \end{bmatrix}$$

$$\therefore \nabla f_5 = \begin{bmatrix} 0.04 \\ -0.04 \end{bmatrix} \approx X_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Here,
∇f_5 is approximately close to X_1 (initial point).
So,
stopping the iteration here, we have;
$x_1 = 0.04$ and $x_2 = -0.04$.
i.e.,
the minimised value of the function $f(x_1, x_2)$ at ∇f_5 is,

$$f(\nabla f_5) = f(0.04, -0.04)$$

$$\boxed{f_{min} = 0.0816}$$

4.2) Newton's Method :-

Q.16) Minimize the given function $f(x_1, x_2)$, starting from the point $X_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, given $x_1 = x_2 = 0$ using Newton's method if :

$$f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2.$$

Ans : i) Step 1: Find the Hessian's matrix $[H_1]$.

Here,

$$[H_1] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial^2 x_1 x_2} \\ \frac{\partial^2 f}{\partial^2 x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}$$

where,

$\Rightarrow \frac{\partial f}{\partial x_1} = 1 + 4x_1 + 2x_2$ $\qquad \Rightarrow \frac{\partial^2 f}{\partial x_1^2} = 4.$

$\Rightarrow \frac{\partial f}{\partial x_2} = -1 + 2x_1 + 2x_2$ $\qquad \Rightarrow \frac{\partial^2 f}{\partial x_2^2} = 2.$

$\Rightarrow \frac{\partial^2 f}{\partial^2 x_1 x_2} = \frac{\partial}{\partial x_1}\left(\frac{\partial f}{\partial x_2}\right) = 2$

$\Rightarrow \frac{\partial^2 f}{\partial^2 x_2 x_1} = \frac{\partial}{\partial x_2}\left(\frac{\partial f}{\partial x_1}\right) = 2.$

$\therefore$ Value of matrix $[H_1]$ is,

$$H_1 = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}.$$

ii) Step 2: Find inverse of the matrix H.

$$\Rightarrow [H_1]^{-1} = \frac{1}{4\times2 - 2\times2}\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix}$$

$$\Rightarrow [H_1]^{-1} = \frac{1}{4}\begin{bmatrix} 2 & -2 \\ -2 & 4 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1 \end{bmatrix}.$$

Note: (INVERSE OF A MATRIX): For a given matrix A, if,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{where, } ad - bc \neq 0.$$

Then, $[A^{-1}]$ is,

$$\Rightarrow [A^{-1}] = \frac{1}{(ad-bc)} \times \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

iii) Step 3: Find ∇f_1 at the given initial point X_1 using the formula,

$$\Rightarrow \nabla f_1 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_1} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \text{ at } \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

iv) Step 4: Find the next iteration X_2 using the formula,

$$\Rightarrow X_2 = X_1 - ([H_1]^{-1}).\nabla f_1$$

$$\Rightarrow X_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

$$\Rightarrow X_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ -3/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$$

v) Step 5: Find ∇f_2 at the point X_2.

$$\Rightarrow \nabla f_2 = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}_{X_2} = \begin{bmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{bmatrix} \text{ at } \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$$

on solving,

$$\Rightarrow \nabla f_2 = \begin{bmatrix} 1 - 4 + 3 \\ -1 - 2 + 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = X_1$$

vi) Step 6: (If $\nabla f_2 \neq X_1$ or $\nabla f_2 \neq X_1$ then..)

Find the next iteration X_3. i.e., also ~~ca~~ find matrix $[H_2]$

$$X_3 = X_2 - [H_2]^{-1} \cdot \nabla f_2$$

* [(General Formula): To find $(n+1)^{th}$ iteration we have,

$$X_{n+1} = X_n - [H_n]^{-1} \cdot \nabla f_n$$]

vii) Step 7: (Continuation of Step-6;)

→ Find ∇f_3 at the point X_3. Then compare ∇f_3 with X_1, if approximately equal then stop at this iteration, and if not equal then repeat steps - 6 and 7 again..

∵ In this question, at step-5; $\nabla f_2 = X_1$, i.e., the minimum value of the function is X_2: $(-1, 1.5)$ and $\boxed{F_{min} = \cancel{1.75}}$ → $\boxed{-1.25}$

4.3) Nelder-Mead Downhill Simplex Method :- (HARD)

(Nelder-Mead's Algorithm for Two Dimensions and Three Sides → A Triangle...)

- Nelder-Mead Method is a simplex method for finding a local minimum of a function of several variables.

→ Simplex : Elementary Geometrical figure formed in N Dimensions and has (N+1) sides.
For example,
1) A 2-simplex is a triangle (2 dimensions and 3 sides).
2) A 3-simplex is a tetrahedron (3 dimensions and 4 sides).

⊙ We will study this method only for a 2-simplex figure.

Goal: To find the local minima (minimum point) by,

1) Moving the Simplex towards the (local) minimum. (R/E).
2) Surrounding the minimum, then,
3) Contracting the Simplex around the minimum. (C/S)

- This procedure shall be repeated till an acceptable error (ϵ) has been reached.

- Four Searching Movements are involved, i.e.,

1) Reflection (R)
2) Expansion (E)
3) Contraction (C)
4) Shrinking (S).

* Notations to be involved :-

1)	B	Best Vertex	→ Vertices of the Triangle.
2)	G	Good Vertex	
3)	W	Worst Vertex	
4)	R	Reflection Point	→ Searching Movements.
5)	E	Expanding Point	
6)	C	Contraction Point	
7)	S	Shrinking Point	
8)	M	Mid-Point of $\overline{BG}$.	
9)	C_1	Mid-Point of Contracting Point (C) of $\overline{WM}$.	
10)	C_2	Mid-Point of Contracting Point (C) of $\overline{MR}$.	

* Working Procedure :-

Step 1: Consider initial triangle BGW and choose the three vertices : (x_K, y_K) where, K = 1, 2 and 3.

① $B = (x_1, y_1)$ (the best vertex)
② $G = (x_2, y_2)$ (the good vertex)
③ $W = (x_3, y_3)$ (the worst vertex).

⇒ To select which will be B or G or W vertex, use the relation,

$$\boxed{f(B) \leq f(G) \leq f(W)} \quad \text{—— ①}$$

Example: Given, $f(x, y) = x + y$. And vertices of the triangle are $(0,0)$, $(2,3)$ and $(4,-1)$. Determine B, G & W?

Ans:
$f(0,0) = 0+0 = 0 \longrightarrow B$
$f(2,3) = 2+3 = 5 \longrightarrow W$
$f(4,-1) = 4-1 = 3 \longrightarrow G$
} Selected on the basis of relation - ①.

$\therefore$ $B = (0,0)$, $G = (4,-1)$, $W = (2,3)$.

Step 2: Find the Mid-Point of the Good Side $\overline{BG}$.

$$\Rightarrow \boxed{M = \frac{B+G}{2}} = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$$

Step 3: Determine the Reflection Point (R) or Expansion Point (E) based on the conditions. (discussed later)

$$\Rightarrow \boxed{R = 2M - W}$$

(REFLECTION)

$$\Rightarrow \boxed{E = 2R - M}$$

(EXPANSION)

Step 4: Determine the Contraction Point (C).
There will be two contracting points, i.e.,

1). Contraction Point (C_1) lying between $\overline{WM}$.

$$\boxed{C_1 = \frac{W+M}{2}}$$

2) Contraction Point (C_2) lying between $\overline{MR}$.

$$\boxed{C_2 = \frac{M+R}{2}}$$

(CONTRACTION)

Step 5: Determine the Shrinking Point (S).

⇒ $\boxed{S = \frac{B+W}{2}}$ (shrink towards B)

(SHRINKING).

★ BUT,
on what conditions will steps - 2, 3, 4 and 5 be applied and which point will we determined among R, E, C or S?

→ **Order of Procedure :-**

1) If, $f(R) < f(G)$, then the searching movement will be either Reflection (R) or Expansion / Extension (E), and CASE - 1 shall be performed.

2) If, $f(R) \nless f(G)$, then the searching movement will be either Contraction (C) or Shrinking (S), and CASE - 2 shall be performed.

Step Number	CASE - 1	CASE - 2
(1) Ⓐ	If, $f(B) < f(R)$ then, Replace W with R.*	If, $f(R) < f(W)$, then Replace W with R. AND, Compute C and $f(C)$, i.e., $C_1 = \frac{W+M}{2}$, $C_2 = \frac{M+R}{2}$ and, $f(C_1) = f(C_2) = f(C)$. Note: Use the new point W to find C_1 and further on..
Ⓑ	If, $f(B) \nless f(R)$ then, Compute $E = 2R - M$ and, evaluate $f(E)$.	If $f(R) \geq f(W)$ then. Compute C and $f(C)$ directly. $C = C_1 = \frac{W+M}{2} = C_2 = \frac{M+R}{2}$

(2) Ⓐ	If, $f(E) < f(B)$ then, Replace W with E.*	If, $f(C) < f(W)$ then, Replace W with C.*
Ⓑ	If, $f(E) \nless f(B)$ then, Replace W with R.*	CONTINUED..
	– CASE (1) ENDS –	
(3)	–	If, $f(C) \nless f(W)$ then, find S, i.e., $S = \frac{B+W}{2}$ and $f(S)$. Replace W with S.* Replace G with M.*
		– CASE (2) ENDS –

Q.17) Using Nelder-Mead Downhill Simplex method, find the minimum value of the function: (upto 6 iterations).

$$f(x,y) = x^2 - 4x + y^2 - y - xy$$

Given vertices for the 2-D simplex as:
$V_1 = (0,0)$, $V_2 = (1.2, 0)$, $V_3 = (0, 0.8)$.

Ans: ① Iteration 1: First find BGW triangle,

$\Rightarrow f(V_1) = f(0,0) = 0 \rightarrow W$
$\Rightarrow f(V_2) = f(1.2, 0) = -3.36 \rightarrow B$
$\Rightarrow f(V_3) = f(0, 0.8) = -0.16 \rightarrow G.$

$\because$ The relation $f(B) < f(G) < f(W)$ should be satisfied.
So,
$B = (1.2, 0)$, $G = (0, 0.8)$, $W = (0, 0)$.

$\rightarrow$ Now, finding mid-point (M) of $\overline{BG}$.

$\Rightarrow M = \frac{B+G}{2} = \left(\cancel{1+2}\ \frac{1.2+0}{2}, \frac{0+0.8}{2}\right) = (0.6, 0.4)$.

$\rightarrow$ Now, $R = 2M - W = 2(0.6, 0.4) - (0, 0)$.
$\Rightarrow R = (1.2, 0.8) - (0, 0)$.
so,
$$R = (1.2, 0.8)$$
$\Rightarrow f(R) = (1.2)^2 - 4(1.2) + (0.8)^2 - (0.8) - (1.2)(0.8)$
$\Rightarrow f(R) = -4.48$.

$\rightarrow$ Since, $f(R) = -4.48$ and, $f(G) = -0.16$
As, $\boxed{f(R) < f(G)}$, so <u>CASE-1</u> will be followed.

$\rightarrow$ Now, $f(B) = -3.36 \nless f(R) = -4.48$. So, computing value of E,
$\Rightarrow E = 2R - M = 2(1.2, 0.8) - (0.6, 0.4)$
$\Rightarrow E = (2.4, 1.6) - (0.6, 0.4)$
$\Rightarrow E = (1.8, 1.2)$.
and
$f(E) = (1.8)^2 - 4(1.8) + (1.2)^2 - (1.2) - (1.8)(1.2)$
$f(E) = -4.68 - 1.2 = -5.88$.

$\rightarrow$ Now, $f(E) = \overset{-5.88}{\cancel{-4.68}} < f(B) = -3.36$, so replacing W with E, i.e.,
$$W = (1.8, 1.2).$$

② **Iteration 2:** Find BGW triangle again,

$\Rightarrow f(1.2, 0) = -3.36 \rightarrow G$

$\Rightarrow f(0, 0.8) = -0.16 \rightarrow W$

$\Rightarrow f(1.8, 1.2) = -5.88 \rightarrow B$

$\because$ The relation $f(B) \leq f(G) \leq f(W)$ should be satisfied, so
$B = (1.8, 1.2)$ $G = (+3.2, 0)$ $W = (0, 0.8)$.

$\rightarrow$ Now, mid-point (M) of line segment $\overline{BG}$.

$$\Rightarrow M = \frac{B+G}{2} = \left(\frac{1.8+1.2}{2}, \frac{1.2+0}{2}\right) = (1.5, 0.6)$$

$\rightarrow$ Value of Reflextion Point (R),

$\Rightarrow R = 2M - W = 2(1.5, 0.6) - (0, 0.8) = (3, 0.4)$

and

$\Rightarrow f(R) = (3)^2 - 4(3) + (0.4)^2 - (0.4) - (3)(0.4)$

$\therefore f(R) = -4.44$.

$\because \boxed{f(R) < f(G)}$, so CASE-1 will be followed.

i) Now, because $f(B) < f(R)$, so replacing W with R, i.e.,
$W = (3, 0.4)$.

ii) Since, $f(B) < f(R)$, so there was no expansion point (E) so, case-1 will end here and we move to the next iteration. (No Step-2 will be performed).

③ **Iteration 3:** Find BGW triangle again,

$\Rightarrow f(1.2, 0) = -3.36 \rightarrow W$.

$\Rightarrow$ $f(3, 0.4) = -4.44 \longrightarrow G$
$\Rightarrow$ $f(1.8, 1.2) = -5.88 \longrightarrow B$

$\because$ The relation $f(B) \leq f(G) \leq f(W)$ should be satisfied, so
$B = (1.8, 1.2)$, $G = (3, 0.4)$, $W = (1.2, 0)$.

$\rightarrow$ Now, mid-point (M) of line segment $\overline{BG}$.

$\Rightarrow$ $M = \dfrac{B+G}{2} = \left(\dfrac{1.8+3}{2}, \dfrac{1.2+0.4}{2}\right) = (2.4, 0.8)$

$\rightarrow$ Value of Reflextion Point (R),
$\Rightarrow$ $R = 2M - W = 2(2.4, 0.8) - \cancel{(0, 0.8}\ (1.2, 0)$
$\Rightarrow$ $R = (4.8, 1.6) - (1.2, 0)$
$\therefore$ $R = (3.6, 1.6)$
and,
$\Rightarrow$ $f(R) = (3.6)^2 - 4(3.6) + (1.6)^2 - (1.6) - (3.6)(1.6)$
$\therefore$ $f(R) = -6.24$.

$\because$ $\boxed{f(R) < f(G)}$, so <u>CASE - 1</u> begins.

i) Now, because $f(B) \nless f(R)$ then, finding the expanding point (E) and $f(E)$.
$\Rightarrow$ $E = 2R - M = 2(3.6, 1.6) - (2.4, 0.8)$
$\Rightarrow$ $E = (4.8, 2.4)$
and,
$\Rightarrow$ $f(E) = -4.32$

ii) Now, because $f(E) \nless f(B)$ so, replacing W with R.
That is,
$W = (3.6, 1.6)$

④ Iteration 4: Find BGW triangle again.

$\Rightarrow f(3.6, 1.6) = -6.24 \rightarrow B$
$\Rightarrow f(3, 0.4) = -4.44 \rightarrow W$
$\Rightarrow f(1.8, 1.2) = -5.88 \rightarrow G$

$\because$ The relation $f(B) \leq f(G) \leq f(W)$ should be satisfied so,
$B = (3.6, 1.6)$, $G = (1.8, 1.2)$, $W = (3, 0.4)$.

$\rightarrow$ Mid-Point of line segment of $\overline{BG}$.
$\Rightarrow M = \frac{B+G}{2} = \left(\frac{3.6+1.8}{2}, \frac{1.6+1.2}{2}\right) = (2.7, 1.4)$

$\rightarrow$ Value of Reflextion Point (R),
$\Rightarrow R = 2M - W = 2(2.7, 1.4) - (3, 0.4) = (2.4, 2.4)$
and,
$\Rightarrow f(R) = -(2.4)^2 - 4(2.4) + (2.4)^2 - (2.4) - (2.4)(2.4)$
$\therefore f(R) = -6.24$.

$\because$ $\boxed{f(R) < f(G)}$, so CASE-1 begins.

i) $\because$ Here, $f(B) = f(R) = -6.24$, but the condition was strictly for $f(B) < f(R)$ so, we can say;
$f(B) \nless f(R)$.
We will find E and f(E).
$\Rightarrow E = 2R - M = 2(2.4, 2.4) - (2.7, 1.4)$
$\therefore E = (2.1, 3.4)$
and,
$\Rightarrow f(E) = -2.97$.

ii) $\because f(E) \nless f(B)$ $[-2.97 \nless -6.24]$, so replacing W with R.

$W = (2.4, 2.4)$.

(5) Iteration 5: Find BGW triangle again with new vertices,

$\Rightarrow f(3.6, 1.6) = -6.24 \longrightarrow B$
$\Rightarrow f(2.4, 2.4) = -6.24 \longrightarrow G$
$\Rightarrow f(1.8, 1.2) = -5.88 \longrightarrow W$

$\because$ The relation $f(B) \leq f(G) \leq f(W)$ should be satisfied,
$B = (3.6, 1.6)$ $G = (2.4, 2.4)$ $W = (1.8, 1.2)$.

$\rightarrow$ Mid-Point (M) of $\overline{BG}$.

$\Rightarrow M = \frac{B+G}{2} = \left(\frac{3.6+2.4}{2}, \frac{1.6+2.4}{2}\right) = (3, 2)$

$\rightarrow$ Values of (R) and f(R).

$\Rightarrow R = 2M - W = 2(3, 2) - (1.8, 1.2) = (4.2, 2.8)$

and,

$\Rightarrow f(R) = -5.88$

$\because$ $\boxed{f(R) \nleq f(G)}$, so CASE-2 begins.

i) Since, $f(R) \leq f(W)$, i.e., $-5.88 = -5.88 = f(R) = f(W)$
Now, computing the Contraction Point (C), i.e.,

$\Rightarrow C_1 = \frac{W+M}{2} = \left(\frac{1.8+3}{2}, \frac{1.2+2}{2}\right) = (2.4, 1.6)$

and,

$\Rightarrow C_2 = \frac{M+R}{2} = \left(\frac{3+4.2}{2}, \frac{2+2.8}{2}\right) = (3.6, 2.4)$

and,

$\rightarrow f(C_1) = f(C_2) = -6.72$ (On solving..)

Both have the same value. So,

$$C = (2.4, 1.6)$$
$$f(C) = -6.72$$

ii) Now, because $f(C) < f(W)$, so replace W with C.

$$W = (2.4, 1.6).$$

(6) Iteration 6: Find BGW triangle again,

$f(3.6, 1.6) = -6.24 \longrightarrow G$
$f(2.4, 2.4) = -6.24 \longrightarrow W$
$f(2.4, 1.6) = -6.72 \longrightarrow B$

$\because$ The relation $f(B) \leq f(G) \leq f(W)$ should be satisfied so,
$B = (2.4, 1.6)$, $G = (3.6, 1.6)$, $W = (2.4, 2.4)$

→ Mid-point (M) of $\overline{BG}$,

$$\Rightarrow M = \frac{B+G}{2} = \left(\frac{2.4+3.6}{2}, \frac{1.6+1.6}{2}\right) = (3, 1.6)$$

→ Value of Reflection Point (R) and f(R),

$$\Rightarrow R = 2M - W = 2(3, 1.6) - (2.4, 2.4) = (3.6, 0.8)$$

and,

$$\Rightarrow f(R) = -4.48$$

$\because$ $f(R) \nless f(G)$, so CASE-2 begins.

i) Since, $f(R) \nless f(W)$, i.e., $-4.48 \nless -6.24$.
so, computing value of $f(C)$, i.e.,

$$\Rightarrow C_2 = C = \frac{M+R}{2} = \left(\frac{3+3.6}{2}, \frac{1.6+0.8}{2}\right)$$

$$\therefore C = (3.3, 1.2).$$

Note: • In the question it will be specified as which value of C we have to assume either C_1 or C_2 throughout the whole process (procedure).
• If NOT mentioned, then assume C to be either C_1 or C_2 and solve the question throughout based on this assumption and then apply the further conditions.

⇒ Now, $f(C) = f(C_2) = -6.03$
and,
$f(W) = -6.24$.

ii) Since, $f(C) \neq f(W)$. So, we have to find the Shrinking Point (S) and f(S).

⇒ $S = \frac{B+W}{2} = \left(\frac{2.4+2.4}{2}, \frac{1.6+2.4}{2}\right) = (2.4, 2)$

and,
⇒ $f(S) = (2.4)^2 - 4(2.4) + (2)^2 - (2) - (2.4)(2)$
⇒ $f(S) = -6.64$.

→ Further, Replacing W with S and G with M, i.e.,

updated values → $B = (2.4, 1.6)$, $G = (3, 1.6)$, $\cancel{M}\ W = (2.4, 2)$.

Now, $f(2.4, 1.6) = -6.72 \longrightarrow G$
$f(3, 1.6) = -6.84 \longrightarrow B$
$f(2.4, 2) = -6.64 \longrightarrow W$.

ANSWER:
∴ Stopping the iteration here, so new vertices are:
$B = (3, 1.6)$, $G = (2.4, 1.6)$, $W = (2.4, 2)$.

→ The value of F_{min} upto 6 iterations is: $\boxed{-6.84}$ at B.

5) Rolle's Theorem and Mean Value Theorem :-

5.1) Rolle's Theorem :-

→ Let $f(x)$ be a function defined on $[a, b]$ such that,

① $f(x)$ is a continous function on $[a, b]$
② $f(x)$ is a differentiable function on (a, b).
③ $f(a) = f(b)$.

Then for a real number, $c \in (a, b)$ we have,

$$\boxed{f'(c) = 0}$$

5.2) Mean Value Theorem :-

Let $f(x)$ be a function defined on $[a, b]$ such that,

① $f(x)$ is a continous function on $[a, b]$.
② $f(x)$ is a differentiable function on (a, b).

Then, there exists some $c \in (a, b)$ such that,

$$\boxed{f'(c) = \frac{f(b) - f(a)}{b - a}}$$

Example: Verify Rolle's Theorem in $[-1, 1]$ for the function :
$f(x) = x^2$.

Ans: ⇒ $f(a) = f(b)$ $\quad (\exists\ a = -1\ \forall\ b = 1)$
⇒ $f(-1) = f(1)$
⇒ $(-1)^2 = (1)^2$
⇒ $1 = 1$ ⇒ LHS = RHS

also, given $f(x)$ is both continous and differentiable on $[-1, 1]$

→ For a real number, $c \in (-1, 1)$ we have,

$\Rightarrow f'(c) = 0$

$\Rightarrow \frac{d}{dc}[f(c)] = 0 \quad \Rightarrow \frac{d}{dc}[c^2] = 0$

$\Rightarrow 2c = 0$

$\therefore \boxed{c = 0} \quad \Rightarrow \boxed{c \in (-1, 1)}$

$\therefore$ Rolle's Theorem is verified for the given function.

Q.18) Verify Rolle's Theorem and Mean Value Theorem for the given function in the interval $[-3, 0]$.

$$f(x) = x(x+3)e^{\frac{-x}{2}}$$ ~~(IF POSSIBLE)~~.

Ans: ① Rolle's Theorem:

i) We have, $a = -3$ and $b = 0$ for interval $[-3, 0]$. So,

$\Rightarrow f(a) = f(b)$.

$\Rightarrow f(-3) = f(0)$.

$\Rightarrow (-3)(-3+3) \times \exp\left(\frac{+3}{2}\right) = (0)(0+3) \times \exp\left(\frac{-0}{2}\right)$.

$\Rightarrow \boxed{0 = 0} \quad \therefore$ LHS = RHS.

Now,

ii) For a real number, $c \in (a, b)$ we have,

$\Rightarrow f'(c) = 0$ or, $\frac{d}{dc}[f(c)] = 0$

$\Rightarrow \frac{d}{dc}\left(c(c+3)\exp\left(\frac{-c}{2}\right)\right) = 0$.

$\Rightarrow \frac{d}{dc}\left[(c^2+3c) \cdot e^{\frac{-c}{2}}\right] = 0$

$$\Rightarrow (2c+3)e^{\frac{-c}{2}} + (c^2+3c)\left(\frac{-1}{2}\right)e^{\frac{-c}{2}} = 0$$

$$\Rightarrow e^{\frac{-c}{2}}(4c+6-c^2-3c) = 0$$

$$\Rightarrow e^{\frac{-c}{2}}(c^2-c-6) = 0$$

$\because$ e^x is never equal to zero and graphically also, it always tend to zero. So,

$$e^{\frac{-c}{2}} \neq 0 \quad \text{(Case Rejected)}$$

But,

$$\Rightarrow c^2-c-6 = 0$$

$$\Rightarrow \boxed{c=-2}$$ ($c=3$ rejected since it lies outside the interval $[-3, 0]$).

$\therefore$ For a real number, $\boxed{c=-2}$, $c \in (-3, 0)$ we have

$$\boxed{f'(-2) = 0}$$

hence,

Rolle's Theorem is verified.

(2) <u>Mean Value Theorem</u> :

$\rightarrow$ Since, $f(x)$ is already continous and differentiable for the interval $(-3, 0)$. So,

there exists some $c \in (a, b)$ such that,

$$f'(c) = \frac{f(b)-f(a)}{b-a}$$

So,

$a = -3$, $f(a) = 0$.

$b = 0$, $f(b) = 0$

i.e.,

$$\Rightarrow \quad f'(c) = \frac{0-0}{0-(-3)} = 0.$$

$$\therefore \boxed{f'(c) = 0}$$

Same Condition as Rolle's theorem.

Hence, the value of c will be same again, i.e.,

$$\boxed{c = -2}$$

$\therefore$ Mean Value Theorem is also verified.

Note: Logically, if Rolle's Theorem is verified then MVT is automatically also verified or vice-versa.

—X—

Question) Using Newton - Raphson Method, find the root of the equation $f(x) = x\sin x + \cos x = 0$ which is near $x = \pi$ correct to three decimal places?
(DIY) (ET 2018, 6.5 marks)

Answer:

VANSH CSE
05115602721

Unit II

Interpolation

Introduction -

Suppose we are given the following values of $y=f(x)$ for a set of values of x:

$$x : x_0 \quad x_1 \quad x_2 \ldots x_n$$
$$y : y_0 \quad y_1 \quad y_2 \ldots y_n$$

The process of finding the values of y corresponding to any value of $x = x_i$ b/w x_0 & x_n is called interpolation.

Assumptions of Interpolation -

1) There are no sudden jumps or falls in the values during the period under consideration.
2) The rise and fall in the values should be uniform.
3) When we apply calculus of finite differences, we assume that the given set of observations are capable of being expressed in a polynomial form.

Errors in polynomial Interpolation -

If the fn $f(x)$ is known explicitly, the value of y corresponding to any value of x can easily be found. If the fn $f(x)$ is not known, it is required to find a simpler fn. say $\varphi(x)$ such that $f(x)$ & $\varphi(x)$ agree at the set of tabulated points. Such a process is called interpolation. If $\varphi(x)$ is a polynomial, then the process is called polynomial interpolation & $\varphi(x)$ is called interpolating polynomial.

Let the fn $y(x)$ defined by $(n+1)$ points (x_i, y_i) $i=0,1\ldots n$ be continuous & differentiable $(n+1)$ times & let $y(x)$ be approximated by a polynomial $\varphi_n(x)$ of degree not

exceeding n such that

$$\varphi_n(x_i) = y_i ; i = 0, 1, 2 \ldots n$$

Now the problem lies in finding the accuracy of this approximation, if we use $\varphi_n(x)$ to obtain approximate value of $y(x)$ at some points other than those given.

Since the expression $y(x) - \varphi_n(x)$ vanishes for $x = x_0, x_1, \ldots x_n$ we put

$$y(x) - \varphi_n(x) = L(x-x_0)(x-x_1)\ldots(x-x_n) \quad \text{—①}$$

where

$$= L\Pi_{n+1}(x)$$

L is to be determined such that ① holds for any intermediate value of x say x' where $x_0 < x' < x_n$

Clearly $$L = \frac{y(x') - \varphi_n(x')}{\Pi_{n+1}(x')} \quad \text{—②}$$

Construct a fn. $$F(x) = y(x) - \varphi_n(x) - L\Pi_{n+1}(x) \quad \text{—③}$$

where L is given by ②

It is clear that

$$F(x_0) = F(x_1) = \ldots = F(x_n) = F(x') = 0$$

ie $F(x)$ vanishes $(n+2)$ times in interval $[x_0, x_n]$ Consequently, by repeated application of Rolle's theorem, $F'(x)$ must vanish $(n+1)$ times, $F''(x)$ must vanish n times in the interval $[x_0, x_n]$.

Particularly $F^{n+1}(x)$ must vanish once in $[x_0, x_n]$

Let this point be

$$x = \xi \qquad x_0 < \xi < x_n$$

Differentiating ③ $(n+1)$ time w.r.t x & put $x = \xi$, we get

Finite differences.

The Calculus of finite differences deals with the changes that take place in the value of the function (dependent variable), due to finite changes in the independent variable. Through this we also study the relations that exist b/w the values assumed by the function, whenever the independent variable changes by finite jumps whether equal or unequal. On the other hand, in infinit-esimal calculus, we study those changes of the function which occur when the independent variable changes continously in a given interval. Now we will study the variations in the function when the independent variable changes by equal intervals.

Finite Differences— Suppose that the function $y = f(x)$ is tabulated for the equally spaced values $x = x_0, x_0+h, x_0+2h, \ldots\ldots x_0+nh$ giving $y = y_0, y_1, \ldots\ldots y_n$. To determine the values of $f(x)$ or $f'(x)$ for some intermediate values of x, the following three type of differences are found useful:

① <u>Forward differences</u>— The differences $y_1-y_0;\ y_2-y_1;\ \ldots\ y_n-y_{n-1}$ when denoted by $\Delta y_0, \Delta y_1, \ldots\ldots \Delta y_{n-1}$, respectively are called the first forward differences where Δ is the forward difference operator. similarly, the second forward differences are defined by

$$\Delta^2 y_r = \Delta y_{r+1} - \Delta y_r \qquad \left[\text{1st forward difference is } \Delta y_r = y_{r+1} - y_r\right]$$

In general, $\Delta^p y_r = \Delta^{p-1} y_{r+1} - \Delta^{p-1} y_r$

defines the pth forward differences.

These differences are systematically set out as follows in what is called a forward difference Table.

value of x	value of y	1st differ.	2nd differ.	3rd differ.	4th differ.	5th differ
x_0	y_0					
		Δy_0				
x_0+h	y_1		$\Delta^2 y_0$			
		Δy_1		$\Delta^3 y_0$		
x_0+2h	y_2		$\Delta^2 y_1$		$\Delta^4 y_0$	
		Δy_2		$\Delta^3 y_1$		$\Delta^5 y_0$
x_0+3h	y_3		$\Delta^2 y_2$		$\Delta^4 y_1$	
		Δy_3		$\Delta^3 y_2$		
x_0+4h	y_4		$\Delta^2 y_3$			
		Δy_4				
x_0+5h	y_5					

x is called the argument & y the function or the entry. y_0, the first entry is called the leading term & $\Delta y_0, \Delta^2 y_0, \Delta^3 y_0$ etc. are called the leading differences.

Note ①: Any higher order forward difference can be expressed in terms of the entries. we have

$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0$$
$$= (\Delta y_2 - \Delta y_1) - (\Delta y_1 - \Delta y_0)$$
$$= \Delta y_2 - 2\Delta y_1 + \Delta y_0$$
$$= (y_3 - y_2) - 2(y_2 - y_1) + (y_1 - y_0)$$
$$= y_3 - 3y_2 + 3y_1 - y_0$$

$y_2 - 2y_1 + y_0.$

The coefficients occuring on the right hand side being the binomial coefficients, we have in general,

$$\Delta^n y_0 = y_n - {}^nC_1 y_{n-1} + {}^nC_2 y_{n-2} - \cdots + (-1)^n y_0.$$

Note ②: The operator Δ obeys the distributive, commutative & index laws. i.e

(i) $\Delta[f(x) \pm g(x)] = \Delta f(x) \pm \Delta g(x)$

(ii) $\Delta[cf(x)] = c\Delta f(x)$; c is a constant.

(iii) $\Delta^m \Delta^n f(x) = \Delta^{m+n} f(x)$, m, n are +ve constant integers.

(iv) $\Delta[f(x).g(x)] \neq \Delta f(x).\Delta g(x)$.

② <u>Backward differences</u> – The differences $y_1 - y_0, y_2 - y_1, \ldots y_n - y_{n-1}$ when denoted by $\nabla y_1, \nabla y_2, \ldots \nabla y_n$, respectively, are called the first backward differences where ∇ is the backward difference operator. Similarly we define higher order backward differences. Thus, we have

$$\nabla y_r = y_r - y_{r-1}$$

$$\nabla^2 y_r = \nabla y_r - \nabla y_{r-1}$$

$$\nabla^3 y_r = \nabla^2 y_r - \nabla^2 y_{r-1} \quad \text{etc.}$$

The backward difference table is.

Value of x	Value of y	1st differ.	2nd differ.	3rd differ.	4th differ.	5th differ.
x_0	y_0					
		∇y_1				
x_0+h	y_1		$\nabla^2 y_2$			
		∇y_2		$\nabla^3 y_3$		
x_0+2h	y_2		$\nabla^2 y_3$		$\nabla^4 y_4$	
		∇y_3		$\nabla^3 y_4$		$\nabla^5 y_5$
x_0+3h	y_3		$\nabla^2 y_4$		$\nabla^4 y_5$	
		∇y_4		$\nabla^3 y_5$		
x_0+4h	y_4		$\nabla^2 y_5$			
		∇y_5				
x_0+5h	y_5					

③ <u>Central differences</u> — The central differences operator δ is defined by the relation

$$y_1 - y_0 = \delta y_{1/2} \; ; \quad y_2 - y_1 = \delta y_{3/2} \; ; \; \ldots\ldots \; ; \; y_n - y_{n-1} = \delta y_{n-\frac{1}{2}}$$

similarly, higher order central differences are defined as

$$\delta y_{3/2} - \delta y_{1/2} = \delta^2 y_1$$

$$\delta y_{5/2} - \delta y_{3/2} = \delta^2 y_2 \ldots\ldots$$

$$\delta^2 y_2 - \delta^2 y_1 = \delta^3 y_{3/2} \text{ \& so on.}$$

These differences are shown in the following:

Value of x	Value of y	1st differ.	2nd differ.	3rd differ.	4th differ	5th differ.
x_0	y_0					
		$\delta y_{1/2}$				
x_0+h	y_1		$\delta^2 y_1$			
		$\delta y_{3/2}$		$\delta^3 y_{3/2}$		
x_0+2h	y_2		$\delta^2 y_2$		$\delta^4 y_2$	
		$\delta y_{5/2}$		$\delta^3 y_{5/2}$		$\delta^5 y_{5/2}$.
x_0+3h	y_3		$\delta^2 y_3$		$\delta^4 y_3$	
		$\delta y_{7/2}$		$\delta^3 y_{7/2}$		
x_0+4h	y_4		$\delta^2 y_4$			
		$\delta y_{9/2}$				
x_0+5h	y_5					

We see from this table that the central differences on the same horizontal line have the same suffix. Also the differences of odd order are known only for half values of the suffix & those of even order for only integral values of the suffix. To find the mean of adjacent values in the same column of differences. We denote this mean by μ. Thus

$$\mu\delta y_1 = \frac{1}{2}(\delta y_{1/2} + \delta y_{3/2})$$

$$\mu\delta^2 y_{3/2} = \frac{1}{2}(\delta^2 y_1 + \delta^2 y_2)$$

Note: It is only the notation which changes & not the differences. e.g:

$$y_1 - y_0 = \nabla y_0 = \Delta y_1 = \delta y_{1/2}.$$

eg: Evaluate (i) $\Delta \tan^{-1} x$ (ii) $\Delta(e^x \log 2x)$

Solu: (i) $\Delta \tan^{-1}x = \tan^{-1}(x+h) - \tan^{-1}x$

$$= \tan^{-1}\left[\frac{x+h-x}{1+(x+h)x}\right] = \tan^{-1}\left[\frac{h}{1+hx+x^2}\right].$$

(ii) $\Delta(e^x \log 2x) = e^{x+h}\log 2(x+h) - e^x \log 2x$

$$= e^{x+h}\log 2(x+h) - e^{x+h}\log 2x + e^{x+h}\log 2x - e^x \log 2x$$

$$= e^{x+h}\log\left(\frac{x+h}{x}\right) + (e^{x+h} - e^x)\log 2x$$

$$= e^x\left[e^h \log\left(\frac{x+h}{x}\right) + (e^h - 1)\log 2x\right].$$

$$= e^x\left[e^h \log\left(1+\frac{h}{x}\right) + (e^h - 1)\log 2x\right].$$

eg: Form a table of differences for the fn $f(x) = x^3 + 5x - 7$; for $x = -1, 0, 1, 2, 3, 4, 5$.

continue the table to obtain $f(6)$.

Solu: $f(-1) = -13$ $\quad f(3) = 35$

$f(0) = -7$ $\quad f(4) = 77$

$f(1) = -1$ $\quad f(5) = 143.$

$f(2) = 11$

x	y	1st diff	2nd diff	3rd diff	4th diff	5th diff.
-1	-13					
		6				
0	-7		0			
		6		6		
1	-1		6		0	
		12		6		0
2	11		12		0	
		24		6		0
3	35		18		0	
		42		6		
4	77		24			
		66				
5	143					

① Newton's forward Interpolation formula –

Let the function $y = f(x)$ take the values $y_0, y_1 \ldots y_n$ corresponding to the values $x_0, x_1 \ldots x_n$ of x. Let these values of x be equispaced such that $x_i = x_0 + ih$; $i = 0, 1, 2 \ldots$ Assuming $y(x)$ to be a polynomial of nth degree in x such that $y(x_0) = y_0$; $y(x_1) = y_1$; $y(x_2) = y_2$; $\ldots$ $y(x_n) = y_n$.

We can write.

$$y(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2)$$

$$+ \ldots\ldots + a_n(x - x_0)(x - x_1)\ldots(x - x_{n-1}) \quad \text{—①}$$

Putting $x = x_0, x_1, x_2 \ldots x_n$ successively in ①

$$y_0 = a_0 \quad \text{—①}$$

$$y_1 = a_0 + a_1(x_1 - x_0) \quad \text{—②}$$

$$y_2 = a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1) \quad \text{—③ \& so on}$$

from these, we find that

$$a_0 = y_0$$

$$\Delta y_0 = y_1 - y_0 = a_1(x_1 - x_0) = a_1 h$$

$$\therefore a_1 = \frac{\Delta y_0}{h}$$

from ③

$$y_2 - a_0 - a_1(x_2 - x_0) = a_2 \times 2h \times h$$

$$y_2 - y_0 - \frac{\Delta y_0}{\cancel{h}} \times 2\cancel{h} = 2h^2 a_2$$

$$(y_2 - y_1) + (y_1 - y_0) - 2\Delta y_0 = 2h^2 a_2$$

$$\Delta y_1 + \Delta y_0 - 2\Delta y_0 = 2h^2 a_2$$

$$a_2 = \frac{\Delta y_1 - \Delta y_0}{2h^2} = \frac{\Delta^2 y_0}{2h^2}$$

Also $\Delta y_1 = y_2 - y_1 = a_2(x_2 - x_0)(x_2 - x_1) + a_1(x_2 - x_0) - a_1(x_1 - x_0)$

$$= a_2(x_2 - x_0)(x_2 - x_1) + a_1(x_2 - x_1)$$

$$= a_1 h + a_2(2h)(h) = a_1 h + 2h^2 a_2.$$

Putting value of a_1 in this eq.

$$\Delta y_1 = a_1 h + 2h^2 a_2 = \frac{\Delta y_0}{\cancel{h}} \cancel{h} + 2h^2 a_2$$

$$\therefore 2h^2 a_2 = \Delta y_1 - \Delta y_0$$

$$a_2 = \frac{1}{2h^2}(\Delta y_1 - \Delta y_0) = \frac{1}{2!\,h^2} \Delta^2 y_0$$

$$a_3 = \frac{1}{3!\,h^3}\Delta^3 y_0 \quad \& \text{ so on.}$$

Substituting these values of a_i's in ①, we obtain

$$y(x) = y_0 + \frac{\Delta y_0}{h}(x-x_0) + \frac{\Delta^2 y_0}{2!\,h^2}(x-x_0)(x-x_1) + \dots\dots + \dots\dots \quad \text{—②}$$

Now if it is required to evaluate y for $x = x_0 + ph$, then

$$x - x_0 = ph,$$
$$x - x_1 = x - x_0 - (x_1 - x_0) = ph - h = (p-1)h$$
$$x - x_2 = (x - x_1) - (x_2 - x_1) = (p-1)h - h$$
$$= (p-2)h \text{ etc.}$$

Hence, writing $y(x) = y(x_0 + ph) = y_p$, ② becomes

$$y(x) = y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \dots\dots +$$
$$+ \frac{p(p-1)\dots(p-(n-1))}{n!}\Delta^n y_0 \quad \text{—③}$$

It is called Newton's forward interpolation formula as ③ contains y_0 & the forward differences of y_0.

e.g: Using Newton's forward formula, find the value of $f(1.6)$ if

x	:	1	1.4	1.8	2.2
$f(x)$	:	3.49	4.82	5.96	6.5

Solu: The difference table is as under.

x	$f(x)$	Δ	Δ^2	Δ^3	Δ^4
1	3.49				
		1.33			
1.4	4.82		−0.19		
		1.14		−0.41	
1.8	5.96		−0.6		
		0.54			
2.2	6.5				

$$\therefore p = \frac{x - x_0}{h} = \frac{1.6 - 1.4}{0.4} = 0.5$$

$$y(1.6) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0.$$

$$= 4.82 + 0.5 \times 1.14 + \frac{0.5(0.5-1)}{2!} \times (-0.6)$$

$$= 4.82 + 0.57 + \frac{(-0.25)}{2!}(-0.6)$$

$$= 4.82 + 0.57 + \frac{0.15}{2!} = 4.82 + 0.57 + 0.075$$

$$= 5.465.$$

e.g: Estimate the values of $f(22)$ & $f(42)$ from the following data

x:	20	25	30	35	40	45
$f(x)$:	354	332	291	260	231	204

Solu: The difference table is

x	$f(x)$	Δ	Δ^2	Δ^3	Δ^4	Δ^5
20	354					
		-22				
25	332		-19			
		-41		29		
30	291		+10		-37	
		-31		-8		45
35	260		2		8	
		-29		0		
40	231		2			
		-27				
45	204					

(i) To find $f(22)$.

$x_0 = 20$ $\quad h = 5$ $\quad \Delta = 22$ $\quad \Delta^3 = 29$ $\quad \Delta^5 = 45$.

$x = 22$ $\quad y_0 = 354$ $\quad \Delta^2 = -19$ $\quad \Delta^4 = -37$ $\quad p = \frac{22-20}{5} = 0.4$

$$f(22) = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \frac{p(p-1)(p-2)(p-3)}{4!}\Delta^4 y_0 + \frac{p(p-1)(p-2)(p-3)(p-4)}{5!}\Delta^5 y_0$$

$$= 354 + 0.4 \times -22 + \frac{0.4(0.4-1)(-19)}{2!} + \frac{0.4(0.4-1)(0.4-2)\times 29}{3!}$$

$$+ \frac{0.4(0.4-1)(0.4-2)(0.4-3)}{4!} \times (-37) + \frac{(0.4)(0.4-1)(0.4-2)(0.4-3)(0.4-4)}{5!} \times 45$$

-0.6×-1.6
$\times -2.6$
-3.6

$$= 354 - 8.8 + 2.28 + 1.856 + 1.5392 + 1.34784$$

$$= 352.22 \approx 352.$$

(ii) To find $f(42)$.

$x_0 = 40$ $\quad h = 5$

$y_0 = 231$ $\quad p = \frac{42-40}{5} = 0.4$

$\Delta = -27$.

$$f(42) = y_0 + p\Delta y_0$$

$$= 231 + 0.4 \times (-27)$$

$$= 231 - 10.8 = 220.2.$$

we can also use Newton's backward Interpolation formula to solve part (2) i.e to calculate $f(42)$.

② <u>Newton's backward Interpolation formula</u> — Let the function $y = f(x)$ take the values $y_0, y_1, y_2 \ldots$ corresponding to the values $x_0, x_0+h, x_0+2h, \ldots$ of x. Suppose it is required to evaluate $f(x)$ for $x = x_n + ph$, where p is any real number. Then we have

$$y_p = f(x_n + ph) = E^p f(x_n) = (1-\nabla)^{-p} y_n \qquad (\because E^{-1} = (1-\nabla))$$

$$= \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 + \ldots\right] y_n$$

$$= y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \ldots \quad \text{—①}$$

It is called Newton's backward interpolation formula as ① contains y_n & backward differences of y_n.

Note: Newton's forward Inter. formula is used for interpolating the values of y near the beginning of a set of tabulated values & Newton's backward Inter. formula is used for interpolating the values of y near the end of a set of tabulated values.

e.g: The area A of a circle of diameter d is given for the following values:

d :	80	85	90	95	100
A :	5026	5674	6362	7088	7854

Calculate area of a circle of diameter 105.

Solu: The difference table is given below

$x_n = 100$

$x = 105$

$$p = \frac{x - x_n}{h} = \frac{105-100}{5} = 1.$$

$\nabla y_n = 766$, $\nabla^2 y_n = 40$, $\nabla^3 y_n = 2$, $\nabla^4 y_n = 4$

d	A	Δ	Δ^2	Δ^3	Δ^4
80	5026				
		648	40		
85	5674			-2	
		688			4.
90	6362		38		
		726		2	
95	7088		40		
		766			
100	7854.				

$$\therefore\ y(105) = y(100) + p\nabla y(100) + \frac{p(p+1)}{2!}\nabla^2 y(100) + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n$$

$$+ \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y(100)$$

$$= 7854 + 1\times 766 + \frac{1\times 2}{2!}\times 40 + \frac{1\times 2\times 3}{3!}\times 2 + \frac{1.2.3.4}{4!}\times 4$$

$$= 7854 + 766 + 40 + 2 + 4 = 8666.$$

e.g : construct Newton's backward Interpolation ~~[illegible]~~ polynomial for the data

x :	1	2	3	4	5
y :	1	-1	1	-1	1

The Newton's backward Interpolation poly can be obtained using difference table.

x	y	Δ	Δ^2	Δ^3	Δ^4
1	1				
		-2			
2	-1		4		
		2		-8	
3	1		-4		16.
		-2		8	
4	-1		4		
		2			
5	1				

The Newton's backward formula is

$$y(p) = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \dots$$

Here $y_n = 1.$

$x_n = 5$

$p = \frac{x-5}{1} = (x-5).$

$$y(x) = 1 + (x-5).2 + \frac{(x-5)(x-4)}{2!}4 + \frac{(x-5)(x-4)(x-3)}{3!}\times 8$$

$$+ \frac{(x-5)(x-4)(x-3)(x-2)}{4!}\times 16$$

$$= 1 + 2x - 10 + 2(x^2 - 9x + 20) + \frac{4}{3}(x^2 - 9x + 20)(x-3)$$

$$+ \frac{2}{3}(x^2 - 9x + 20)(x^2 - 5x + 6)$$

$$= 1 + 2x - 10 + 2x^2 - 18x + 40 + \frac{4}{3}(x^3 - 3x^2 - 9x^2 + 27 + 20x - 60)$$

$$+ \frac{2}{3}(x^4 - 5x^3 + 6x^2 - 9x^3 + 45x^2 - 54 + 20x^2 - 100x + 120)$$

$$= 1 + 2x - 10 + 2x^2 - 18x + 40 + \frac{4}{3}x^3 - 4x^2 - 12x^2 + 36$$

$$+ \frac{80}{3}x - 80 + \frac{2}{3}x^4 - \frac{10}{3}x^3 + 4x^2 - 6x^3 + 30x^2 - 36 + \frac{40}{3}x^2 - \frac{200}{3}x + 80$$

$y = \frac{2}{3}x^4 - 8x^3 + \frac{100}{3}x^2 - 56x + 31.$ is the req. polynomial

3 Newton's Divided difference interpolation formula –

Let $y_0, y_1, \ldots y_n$ be the values of $y = f(x)$ corresponding to the arguments $x_0, x_1, \ldots x_n$. Then from the definition of the divided differences, we have

$$[x, x_0] = \frac{y - y_0}{x - x_0}.$$

$$\therefore \quad y = y_0 + (x - x_0)[x, x_0] \qquad \text{— (1)}$$

Again $$[x, x_0, x_1] = \frac{[x, x_0] - [x_0, x_1]}{x - x_1}.$$

which gives $$[x, x_0] = [x_0, x_1] + (x - x_1)[x, x_0, x_1]$$

Substituting this value of $[x, x_0]$ in (1), we get

$$y = y_0 + (x - x_0)\left[[x_0, x_1] + (x - x_1)[x, x_0, x_1]\right]$$

$$= y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x, x_0, x_1]. \qquad \text{— (2)}$$

Also

$$[x, x_0, x_1, x_2] = \frac{[x, x_0, x_1] - [x_0, x_1, x_2]}{x - x_2}$$

which gives

$$[x, x_0, x_1] = (x - x_2)[x, x_0, x_1, x_2] + [x_0, x_1, x_2]$$

Substituting in (2)

$$y = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)\{(x - x_2)[x, x_0, x_1, x_2] + [x_0, x_1, x_2]\}.$$

$$= y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)(x - x_2)[x, x_0, x_1, x_2] + (x - x_0)(x - x_1)[x_0, x_1, x_2].$$

Proceeding in this manner, we get

$$y = f(x) = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3] + \ldots + (x - x_0)(x - x_1)\ldots(x - x_n)[x, x_0, x_1, \ldots, x_n] \qquad \text{— (3)}$$

which is called Newton's general interpolation formula with divided difference

e.g: Use Newton's divided difference formula to find $f(x)$ from the following data.

x :	0	1	2	4	5	6
$f(x)$:	1	14	15	5	6	19.

Solu: The divided difference table is

x	$f(x)$	1st div.differ.	2nd div.differ.	3rd div.diff.	4th div.differ
x_0 0	y_0 1				
		13			
x_1 1	y_1 14		-6		
		1		1	
x_2 2	y_2 15		-2		0
		-5		1	
x_3 4	y_3 5		2		0
		1		1.	
x_4 5	y_4 6		6.		
x_5 6	y_5 19.	13.			

$[x_0 x_1] = 13$
$[x_1 x_2] = 1$
$[x_0 x_1 x_2] = -6$
$[x_1 x_2 x_3] = -2$

$$f(x) = f(x_0) + (x-x_0)[x_0, x_1] + (x-x_0)(x-x_1)[x_0, x_1, x_2] + (x-x_0)(x-x_1)(x-x_2)[x_0, x_1, x_2, x_3] + (x-x_0)(x-x_1)(x-x_2)(x-x_3)[x; x_0, x_1, x_2, x_3].$$

$$= 1 + (x-0)\times 13 + (x-0)(x-1)\times -6 + (x-0)(x-1)(x-2)\times 1$$

$$= 1 + (x)\times 13 + (x^2-x)\times -6 + x(x^2-3x+2)$$

$$= 1 + 13x - 6x^2 + 6x + x^3 - 3x^2 + 2x$$

$$= x^3 - 9x^2 + 21x + 1.$$

e.g: Using Newton's divided difference formula, evaluate $f(8)$ for the data given as

x :	4	5	7	10	11	13
$f(x)$:	48	100	294	900	1210	2028

The difference table is

x	$f(x)$	1st	IInd	3rd	4th
4	48				
		52			
5	100		15		
		97		1	
7	294		21		0
		202		1	
10	900		27		0
		310		1	
11	1210		33		
		409			
13	2028				

$294 + (x-7) \times 202 + (x-7)(x-10) \times 27 + (x-7)(x-10)(x-11) \times 1$

$f(8) = 294 + 202 + 27 \times (-2) + (-2)(-3) \times 1$

$= 294 + 202 - 54 + 6$

$= 502 - 54 = 448.$

$\therefore f(x) = f(x_0) + [x_0, x_1](x - x_0) + (x - x_0)(x - x_1)[x_0, x_1, x_2] + \dots$

$= 48 + (x-4)[52] + (x-4)(x-5)(15) + (x-4)(x-5)(x-7) \times 1$

$= 48 + (x-4) \times 52 + (x^2 - 9x + 20) \times 15 + (x^3 - 9x^2 + 20x - 7x^2 + 63x - 140)$

$= 48 + 52x - 208 + 15x^2 - 135x + 300 + x^3 - 9x^2 + 20x - 7x^2 + 63x - 140$

$f(x) = x^3 - x^2$ is the polynomial

$f(8) = (8)^3 - (8)^2 = 448.$

(4) Lagrange's Interpolation formula for unequal differences –

If $y = f(x)$ be a function which takes the values (x_0, y_0) $(x_1, y_1), \dots (x_n, y_n)$. Since there are $(n+1)$ pairs of values of x & y $\therefore$ we can represent $f(x)$ by a polynomial in x of degree n. Let this polynomial be of the form.

$$y = f(x) = a_0 (x - x_1)(x - x_2) \dots (x - x_n) + a_1 (x - x_0)(x - x_2) \dots (x - x_n) + a_2 (x - x_0)(x - x_1)(x - x_3) \dots (x - x_n) + \dots + a_n (x - x_0)(x - x_1) \dots (x - x_{n-1}). \quad \text{—(1)}$$

Putting $x = x_0$, $y = y_0$ in this eq, we get

$$y_0 = a_0 (x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)$$

$$\therefore \; a_0 = \frac{y_0}{(x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)}$$

Similarly putting $x = x_1$, $y = y_1$ we get

$$a_1 = \frac{y_1}{(x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n)}$$

Proceeding the same way, we get $a_2, a_3 \ldots a_n$.

Putting these values of a_i's in the eq ①, we get

$$f(x) = \frac{(x - x_1)(x - x_2) \ldots (x - x_n)}{(x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)} y_0 + \frac{(x - x_0)(x - x_2) \ldots (x - x_n)}{(x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n)} y_1$$

$$+ \ldots + \frac{(x - x_0)(x - x_1) \ldots (x - x_{n-1})}{(x_n - x_0)(x_n - x_1) \ldots (x_n - x_{n-1})} y_n$$

This is known as Lagrange's Interpolation formula.

Note ①: It is easy to remember, but quite cubersome to apply.

②: This formula can also be used to split the given function into partial fraction. If we divide both sides of ① by $(x - x_0)(x - x_1) \ldots (x - x_n)$, we get

$$\frac{f(x)}{(x - x_0)(x - x_1) \ldots (x - x_n)} = \frac{y_0}{(x_0 - x_1)(x_0 - x_2) \ldots (x_0 - x_n)} \cdot \frac{1}{x - x_0} +$$

$$\frac{y_1}{(x_1 - x_0)(x_1 - x_2) \ldots (x_1 - x_n)} \cdot \frac{1}{x - x_1} +$$

$$\ldots + \frac{y_n}{(x_n - x_0)(x_n - x_1) \ldots (x_n - x_{n-1})} \cdot \frac{1}{x - x_n}$$

①

Given data find $y(10)$, using Lagrange's Interpolation formula.

x :	$\overset{x_0}{5}$	$\overset{x_1}{6}$	$\overset{x_2}{9}$	$\overset{x_3}{11}$
y :	$\overset{y_0}{12}$	$\overset{y_1}{13}$	$\overset{y_2}{14}$	$\overset{y_3}{16}$

Solu :

$$y(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}y_1$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}y_3$$

$$= \frac{\overset{(x^2-15x+54)(x-11)}{(x-6)(x-9)(x-11)}}{(-1)(-4)(-6)} \times 12 + \frac{\overset{(x^2-14x+45)(x-11)}{(x-5)(x-9)(x-11)}}{1 \times -3 \times -5} \times 13$$

$$+ \frac{\overset{(x^2-11x+30)(x-11)}{(x-5)(x-6)(x-11)}}{4 \times 3 \times -2} \times 14 + \frac{\overset{(x^2-11x+30)(x-9)}{(x-5)(x-6)(x-9)}}{6 \times 5 \times 2} \times 16$$

$$= \frac{\overset{x^3-26x^2+219x-594}{(x-6)(x-9)(x-11)}}{-2} + \frac{\overset{x^3-25x^2+189x-495}{(x-5)(x-9)(x-11)}}{15} \times 13 +$$

$$-\frac{7}{12}\overset{x^3-22x^2+151x-330}{(x-5)(x-6)(x-11)} + \frac{4}{15}\overset{x^3-20x^2+129x-270}{(x-5)(x-6)(x-9)}$$

This is the polynomial. Now we will find $y(10)$

$$y(10) = \frac{(10-6)(10-9)(10-11)}{-2} + \frac{(10-5)(10-9)(10-11)}{15} \times 13 -$$

$$\frac{7}{12}(10-5)(10-6)(10-11) + \frac{4}{15}(10-5)(10-6)(10-$$

$$= \frac{\cancel{4}^{2} \times 1 \times (-1)}{\cancel{-2}} + \frac{\cancel{5} \times 1 \times (-1) \times 13}{\cancel{15}_{3}} - \frac{7}{\cancel{12}_{3}} \times 5 \times 4 \times (-1) +$$

$$\frac{4}{\cancel{15}_{3}}(5) \times 4 \times 1.$$

$$= 2 - \frac{13}{3} + \frac{35}{3} + \frac{16}{3} = \frac{6-13+35+16}{3} = \frac{44}{3} = 14.6$$

Numerical Integration – The process of evaluating a definite Integral from a set of tabulated values of the integrand $f(x)$ is called numerical Integration. This process when applied to a function of a single variable, is known as quadrature.

The problem of numerical Integration, like that of numerical differentiation, is solved by representing $f(x)$ by an interpolation formula and then integrating it between the given limits. First we study the Newton Cotes quadrature formula and then use it to to study ① Trapezoidal rule ② Simpson's one third rule (iii) Simpson's three eighth rule.

Newton Cotes Quadrature formula –

$$\text{Let } I = \int_a^b f(x)dx$$

where $f(x)$ takes the values $y_0, y_1, y_2 \ldots y_n$ for $x = x_0, x_1, \ldots x_n$.

Let us divide the interval (a,b) into n subintervals of width h so that $x_0 = a$, $x_1 = a+h$, $x_2 = x_0 + 2h, \ldots\ldots x_0 + nh = x_n$

Then

$$I = \int_{x_0}^{x_0+nh} f(x)dx$$

Putting $x = x_0 + rh$ & $dx = hdr$.

$$= h\int_0^n f(x_0+rh)\,dr$$

$$= h\int_0^n \Big[y_0 + r\Delta y_0 + \frac{\overset{(r^2-r)}{r(r-1)}}{2!}\Delta^2 y_0 + \frac{\overset{r(r^2-3r+2)=r^3-3r^2+2r}{r(r-1)(r-2)}}{3!}\Delta^3 y_0 +$$

$$\frac{r(r-1)(r-2)(r-3)}{4!}\Delta^4 y_0 + \frac{r(r-1)(r-2)(r-3)(r-4)}{5!}\Delta^5 y_0$$

$$+ \frac{r(r-1)(r-2)(r-3)(r-4)(r-5)}{6!}\Delta^6 y_0 + \dots\Big]\,dr$$

(by Newton's forward Interpolation).

Integrating term by term, we obtain

$$\int_{x_0}^{x_0+nh} f(x)\,dx = nh\Big[y_0 + \frac{n}{2}\Delta y_0 + \frac{n(2n-3)}{12}\Delta^2 y_0 + \frac{n(n-2)^2}{24}\Delta^3 y_0 +$$

$$\left(\frac{n^4}{5} - \frac{3n^3}{2} + \frac{11n^2}{3} - 3n\right)\frac{\Delta^4 y_0}{4!} +$$

$$\left(\frac{n^5}{6} - 2n^4 + \frac{35n^3}{4} - \frac{50n^2}{3} + 12n\right)\frac{\Delta^5 y_0}{5!} +$$

$$\left(\frac{n^6}{7} - \frac{15n^5}{6} + 17n^4 - \frac{225n^3}{4} + \frac{274n^2}{3} - 60n\right)\frac{\Delta^6 y_0}{6!}$$

$$+\dots\Big]. \quad \text{— (1)}$$

This is known as Newton Cotes quadrature formula. From this general formula we deduce the following quadrature rules by takes different values of n.

(1) Trapezoidal Rule –

Putting $n=1$ in (1) & taking the curve through (x_0, y_0) and (x_1, y_1) as a straight line i.e a polynomial of first order so that the differences of order higher than first become zero, we get

$$\int_{x_0}^{x_0+h} f(x)\,dx = h\left(y_0 + \frac{1}{2}\Delta y_0\right) = \frac{h}{2}(y_0+y_1). \qquad \Delta y_0 = y_1 - y_0$$

Similarly

$$\int_{x_0+h}^{x_0+2h} f(x)dx = h\left(y_1 + \frac{1}{2}\Delta y_1\right) = \frac{h}{2}(y_1+y_2)$$

$$\vdots$$

$$\int_{x_0+(n-1)h}^{x_0+nh} f(x)dx = \frac{h}{2}(y_{n-1}+y_n).$$

Adding these n integrals, we obtain

$$\int_{x_0}^{x_0+nh} f(x)\,dx = \frac{h}{2}\left[(y_0+y_n) + 2(y_1+y_2+\dots+y_{n-1})\right] \quad \text{— (2)}$$

This is known as Trapezoidal Rule.

(2) Simpson's One third Rule –

Putting $n=2$ in (1) above & taking the curve through (x_0, y_0), (x_1, y_1) & (x_2, y_2) as a parabola ie "a polynomial of second order so that differences of order higher than second vanish", we get

$$\int_{x_0}^{x_0+2h} f(x)dx = 2h\left(y_0 + \Delta y_0 + \frac{1}{6}\Delta^2 y_0\right) = \frac{h}{3}(y_0 + 4y_1 + y_2).$$

$$= 2h\left(y_0 + y_1 - y_0 + \frac{1}{6}(\Delta y_1 - \Delta y_0)\right)$$

$$= 2h\left(y_0 + y_1 - y_0 + \frac{1}{6}(y_2 - y_1 - y_1 + y_0)\right) = 2h\left[\frac{y_0}{6} + \frac{2}{3}y_1 + \frac{y_2}{6}\right]$$

Similarly

$$\int_{x_0+2h}^{x_0+4h} f(x)\,dx = \frac{h}{3}(y_2 + 4y_3 + y_4)$$

$$\vdots$$

$$\int_{x_0+(n-2)h}^{x_0+nh} f(x)dx = \frac{h}{3}(y_{n-2} + 4y_{n-1} + y_n) \; ; \; n \text{ being even.}$$

Adding all these integrals, we have when n is even

$$\int_{x_0}^{x_0+nh} f(x)\,dx = \frac{h}{3}\left[(y_0+y_n) + 4(y_1+y_3+\dots+y_{n-1}) + 2(y_2+y_4+\dots+y_{n-2})\right] \quad \text{— (3)}$$

This is known as simpson's One-third Rule or simply simpson's rule & is most commonly used. while applying this rule the given interval must be divided into even number of equal sub-intervals, since we find the area of two stripes at a time.

③ Simpson's three Eighth rule –

Putting $n=3$ in ① & taking the curve through (x_0, y_0) (x_1, y_1), (x_2, y_2) & (x_3, y_3) as a polynomial of third order so that differences above the third order vanish, we get

$$\int_{x_0}^{x_0+3h} f(x)dx = 3h\left(y_0 + \frac{3}{2}\Delta y_0 + \frac{3}{2}\Delta^2 y_0 + \frac{1}{8}\Delta^3 y_0\right)$$

$$= \frac{3}{8}h\left[y_0 + 3y_1 + 3y_2 + y_3\right].$$

Similarly

$$\int_{x_0+3h}^{x_0+6h} f(x)dx = \frac{3h}{8}\left[y_3 + 3y_4 + 3y_5 + y_6\right] \text{ \& so on.}$$

Adding all such expressions from x_0 to x_0+nh, where n is a multiple of 3, we obtain.

$$\int_{x_0}^{x_0+nh} f(x)dx = \frac{3h}{8}\left[(y_0+y_n) + 3(y_1+y_2+y_4+y_5+\dots+y_{n-1}) + 2(y_3+y_6+\dots+y_{n-3})\right]. \quad \text{— ④}$$

which is known as Simpson's three Eighth rule. while applying ④ the number of sub-intervals should be taken as a multiple of 3.

e.g: Evaluate $\int_0^1 \frac{dx}{1+x^2}$ using ① Trapezoidal rule taking $h=1/4$.

② Simpson's One third rule taking $h=1/4$.

③ Simpson's three eighth rule taking $h=1/6$

Solu: ① $h = 1/4$ $\therefore$ $x_0 = 0$ $x_1 = 0.25$ $x_2 = 0.50$ $x_3 = 0.75$ $x_4 = 1$.

x	0	0.25	0.50	0.75	1
$f(x)$	1	0.941	0.8	0.64	0.5
	y_0	y_1	y_2	y_3	y_4.

By Trapezoidal Rule

$$\int_0^1 \frac{dx}{1+x^2} = \frac{h}{2}\left[(y_0+y_4) + 2(y_1+y_2+y_3)\right]$$

$$= \frac{1}{4} \times \frac{1}{2}\left[(1+0.5) + 2(0.941+0.8+0.64)\right]$$

$$= \frac{1}{8}\left[6.262\right] = 0.7827.$$

② $h = 1/4$ & by simpson's 1/3 rule

$$\int_0^1 \frac{dx}{1+x^2} = \frac{h}{3}\left[(y_0+y_4) + 4(y_1+y_3) + 2y_2\right]$$

$$= \frac{1}{4} \times \frac{1}{3}\left[(1+0.5) + 4(0.941+0.64) + 2\times 0.8\right].$$

$$= \frac{1}{12}\left[9.424\right] = 0.7853.$$

③ By simpson's 3/8 rule

$$\int_0^1 \frac{dx}{1+x^2} = \frac{3h}{8}\left[(y_0+y_4) + 3(y_1+y_2) + 2(y_3)\right].$$

$$= \frac{3}{8} \times \frac{1}{4}\left[(1+0.5) + 3(0.941+0.8) + 2\times 0.64\right].$$

$$= 0.7502$$

Errors in quadrature formula –

If y_p is the polynomial representing the function $y=f(x)$ in the interval $[a,b]$ then error in the quadrature formula is given by

$$E = \int_a^b f(x)\,dx - \int_a^b y_p\,dx$$

① Error in Trapezoidal rule –

Expanding $y=f(x)$ in the neighbourhood of $x=x_0$ by Taylor's series, we get

$$y = y_0 + (x-x_0)y_0' + \frac{(x-x_0)^2}{2!}y_0'' + \dots \quad (1)$$

$$\therefore \int_{x_0}^{x_1} y\,dx = \int_{x_0}^{x_0+h}\left[y_0 + (x-x_0)y_0' + \frac{(x-x_0)^2}{2!}y_0'' - \dots\right]dx$$

$$= hy_0 + \frac{h^2}{2!}y_0' + \frac{h^3}{3!}y_0'' + \dots \quad (2)$$

Also A_1 = area of first trepezium in the interval $[x_0, x_1]$

$$= \frac{1}{2}h(y_0+y_1) \quad \text{(by Trapezoidal rule)} \quad (3)$$

Putting $x=x_0+h$ & $y=y_1$ in (1) we get

$$y_1 = y_0 + hy_0' + \frac{h^2}{2!}y_0'' + \dots$$

$$\therefore A_1 = \frac{1}{2}h\left[y_0 + y_0 + hy_0' + \frac{h^2}{2!}y_0'' \dots\right] \quad \text{by putting in (3)}$$

$$= hy_0 + \frac{h^2}{2}y_0' + \frac{h^3}{2.2!}y_0'' + \dots \quad (4)$$

$\therefore$ Error in the interval $[x_0, x_1]$

$$= \int_{x_0}^{x_1} y\,dx - A_1 \quad [\text{by } (2)-(4)]$$

$$= \left(hy_0 + \frac{h^2}{2!}y_0' + \dots\right) - \left(hy_0 + \frac{h^2}{2!}y_0'' \dots\right)]$$

$$= \left(\frac{1}{3!} - \frac{1}{2.2!}\right) h^3 y_0'' + \dots = -\frac{h^3}{12} y_0'' + \dots$$

i.e Principal part of error in $[x_0, x_1] = -\frac{h^3}{12} y_0''$

Similarly " " " " " $[x_1\ x_2] = -\frac{h^3}{12} y_1''$ & so on

Hence total error is

$$E = -\frac{h^3}{12}\left[y_0'' + y_1'' + \dots + y_{n-1}''\right]$$

Assuming $y''(x)$ is the largest of n quantities $[y_0'', y_1'' \dots y_{n-1}'']$ we obtain

$$E < -\frac{nh^3}{12} y''(x) = -\frac{(b-a)h^2}{12} y''(x) \quad \text{—(5)}$$

$\because (b-a = nh)$

$\therefore$ error in trapezoidal rule is of order h^2.

(2) Error in Simpson 1/3 rule —

Expanding $y = f(x)$ around $x = x_0$ by Taylor's series we get (1)

& over the first double strip, we get

$$\int_{x_0}^{x_2} y\,dx = \int_{x_0}^{x_0+2h} \left[y_0 + (x-x_0)y_0' + \frac{(x-x_0)^2}{2!} y_0'' + \dots\right] dx$$

$$= 2hy_0 + \frac{4h^2}{2!} y_0' + \frac{8h^3}{3!} y_0'' + \frac{16h^4}{4!} y_0''' + \dots \quad \text{—(6)}$$

Also area of first double strip by simpson's 1/3 rule

$$A_1 = \frac{1}{3} h\left[y_0 + 4y_1 + y_2\right] \quad \text{—(7)}$$

Putting $x = x_0 + h$ & $y = y_1$ in (1), we get

$$y_1 = y_0 + hy_0' + \frac{h^2}{2!} y_0'' + \frac{h^3}{3!} y_0''' + \dots$$

Putting $x = x_0 + 2h$ & $y = y_2$ in ①, we get

$$y_2 = y_0 + 2hy_0' + \frac{4h^2}{2!}y_0'' + \frac{8h^3}{3!}y_0''' + \dots$$

Substituting these values of y_1 & y_2 in ⑦ we get

$$A_1 = \frac{1}{3}h\left[y_0 + 4\left(y_0 + hy_0' + \frac{h^2}{2!}y_0'' + \frac{h^3}{3!}y_0''' + \dots\right) + \left(y_0 + 2hy_0' + \frac{4h^2}{2!}y_0'' + \frac{8h^3}{3!}y_0''' - \dots\right)\right]$$

$$= 2hy_0 + 2h^2y_0' + \frac{4h^3}{3}y_0'' + \frac{2h^2}{3}y_0''' + \frac{5h^5}{18}y_0^{iv} \dots \quad ⑧$$

∴ Error in interval $[x_0, x_2]$

$$= \int_{x_0}^{x_2} y\,dx - A_1 = \left[2hy_0 + \frac{4h^2}{2!}y_0' + \frac{8h^3}{3!}y_0'' + \frac{16h^4}{4!}y_0''' \dots\right] - \left[2hy_0 + 2h^2y_0' + \frac{4h^3}{3}y_0'' \dots\right]$$

(by ⑥ - ⑧)

$$= \left(\frac{4}{15} - \frac{5}{18}\right)h^5 y_0^{iv} + \dots$$

∴ Principal part of error in $[x_0, x_2] = -\frac{h^5}{90}y_0^{iv}$

Similarly " " " " " $[x_2, x_4] = -\frac{h^5}{90}y_2^{iv}$ & so on

Hence total error

$$E = -\frac{h^5}{90}\left[y_0^{iv} + y_2^{iv} + \dots + y_{2(n-1)}^{iv}\right]$$

Assuming $y^{iv}(x)$ as the largest of $y_0^{iv}, y_2^{iv}, y_4^{iv} \dots$, we get

$$E < -\frac{nh^5}{90}y^{iv}(x) = -\frac{(b-a)h^4}{180}y^{iv}(x)$$

∴ error in simpson's ⅓ rule is of order h^4.

③ Error in Simpson's $\frac{3}{8}$th rule – The principal part of error in the interval $[x_0, x_3] = -\frac{3h^5}{80} y_0^{iv}$

Gaussian quadrature formula –

In Newton cotes quadrature formula, for the evaluation of $\int_0^b f(x)dx$, we require the values of the function at equally spaced points of the interval. Gauss derived a formula which uses the same no. of functional values but with different spacing & yields better accuracy.

Gauss formula is expressed as

$$\int_{-1}^{1} f(x)dx = w_1 f(x_1) + w_2 f(x_2) + \dots + w_n f(x_n) \quad (1)$$
$$= \sum_{i=1}^{n} w_i f(x_i)$$

where w_i & x_i are called the weights & abscissae respectively. The abscissae and weights are symmetrical with respect to the middle point of the interval. There being $2n$ unknowns in (1), $2n$ relations between them are necessary so that the formula is exact for all polynomials of degree not exceeding $(2n-1)$. Thus, we consider

$$f(x) = C_0 + C_1 x + C_2 x^2 + \dots + C_{2n-1} x^{2n-1} \quad (2)$$

Then (1) gives

$$\int_{-1}^{1} f(x)dx = \int_{-1}^{1} (C_0 + C_1 x + C_2 x^2 + \dots + C_{2n-1} x^{2n-1}) dx$$

$$= 2C_0 + \frac{2C_2}{3} + \frac{2C_4}{5} + \dots \quad (3)$$

Putting $x = x_i$ in (2), we get

$$f(x_i) = C_0 + C_1 x_i + C_2 x_i^2 + C_3 x_i^3 + \cdots + C_{2n-1} x_i^{2n-1}$$

Substituting these values on RHS of (1) we get

$$\int_{-1}^{1} f(x)dx = w_1 [C_0 + C_1 x_1 + C_2 x_1^2 + \cdots + C_{2n-1} x_1^{2n-1}] +$$
$$w_2 [C_0 + C_1 x_2 + C_2 x_2^2 + \cdots + C_{2n-1} x_2^{2n-1}] +$$
$$\vdots$$
$$+ w_n [C_0 + C_1 x_n + C_2 x_n^2 + \cdots + C_{2n-1} x_n^{2n-1}]$$

$$= C_0 (w_1 + w_2 + \cdots + w_n) + C_1 (w_1 x_1 + w_2 x_2 \cdots + w_n x_n) +$$
$$C_2 (w_1 x_1^2 + w_2 x_2^2 + \cdots + w_n x_n^2) + \cdots + C_n (w_1 x_1^{2n-1} +$$
$$w_2 x_2^{2n-1} \cdots + w_n x_n^{2n-1}) \quad \text{— (4)}$$

Since (3) & (4) are same for all values of C. $\therefore$ comparing the coefficients of C_i, we obtain $2n$ eqns in $2n$ unknowns w_i & x_i

$$\left.\begin{aligned}
&w_1 + w_2 + \cdots + w_n = 2 \\
&w_1 x_1 + w_2 x_2 + \cdots + w_n x_n = 0 \\
&w_1 x_1^2 + w_2 x_2^2 + \cdots + w_n x_n^2 = \frac{2}{3} \\
&\vdots \\
&w_1 x_1^{2n-1} + w_2 x_2^{2n-1} + \cdots + w_n x_n^{2n-1} = 0
\end{aligned}\right] \quad (5)$$

Solution of these eqns is out of scope. However, the values x_i comes out to be zeros of $(n+1)$th Legendre polynomial.

Gauss formula for $n=2$ is

$$\int_{-1}^{1} f(x)dx = w_1 f(x_1) + w_2 f(x_2)$$

Then (5) becomes

$$w_1 + w_2 = 2$$
$$w_1 x_1 + w_2 x_2 = 0$$
$$w_1 x_1^2 + w_2 x_2^2 = \frac{2}{3}$$
$$w_1 x_1^3 + w_2 x_2^3 = 0$$

Solving, we get

$$w_1 = w_2 = 1, \quad x_1 = \frac{-1}{\sqrt{3}}, \quad x_2 = \frac{1}{\sqrt{3}}$$

$\therefore$ Thus the gauss formula for $n=2$ is

$$\int_{-1}^{1} f(x)\,dx = f\left(-\frac{1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right)$$

which gives the correct value of Integral of $f(x)$ in $(-1,1)$.

Note: The abscissae x_i & weights w_i in (1) are tabulated for different values of n.

n	x_i	w_i
2	-0.57735	1
	0.57735	1
3	-0.77460	0.5555
	0.0000	0.8888
	0.77460	0.5555
4	-0.86114	0.34785
	-0.33998	0.65214
	0.33998	0.65214
	0.86114	0.34785

& ctd.

Note: Gauss formula imposes a restriction on the limits of Integration to be from -1 to 1. In general, the limits of integration $\int_a^b f(x)dx$ are changed to -1 to 1 by means of transformation

$$x = \frac{1}{2}(b-a)u + \frac{1}{2}(b+a)$$

Evaluate $\int_{-1}^{1} \frac{dx}{1+x^2}$ using gauss formula for $n=2$ & $n=3$

Solu: Gauss formula for $n=2$ is

$$I = \int_{-1}^{1} \frac{dx}{1+x^2} = \omega_1 f(x_1) + \omega_2 f(x_2)$$

where $f(x_1) = f(-1/\sqrt{3})$; $\omega_1 = 1$

$f(x_2) = f(1/\sqrt{3})$ $\omega_2 = 1$

$$I = 1.f\left(\frac{-1}{\sqrt{3}}\right) + 1.f\left(\frac{1}{\sqrt{3}}\right)$$

$$= \frac{1}{1+(-1/\sqrt{3})^2} + \frac{1}{1+(1/\sqrt{3})^2} = \frac{3}{4} + \frac{3}{4} = 1.5$$

Gauss formula for $n=3$

$$I = \omega_1 f(x_1) + \omega_2 f(x_2) + \omega_3 f(x_3)$$

$\omega_1 = 0.55555$; $\omega_2 = 0.88889$

$\omega_3 = 0.55555$

$f(x_1) = \frac{1}{1+(-0.77460)^2}$ $f(x_2) = \frac{1}{1+(0.0)^2}$

$f(x_3) = \frac{1}{1+(0.77460)^2}$

$$I = 0.55555 \times \frac{1}{1+(-0.77460)^2} + 0.88889 \times \frac{1}{1+(0.0)^2} + 0.55555 \times \frac{1}{1+(0.7746)^2}$$

$$= 1.5833$$

Evaluate $\int_0^1 \frac{dx}{1+x}$ using $n=3$

Solu: We first change limits $(0,1)$ to $(-1,1)$

$$x = \frac{1}{2}(b-a)u + \frac{1}{2}(b+a)$$
$$= \frac{1}{2}(1-0)u + \frac{1}{2}(1+0) = \frac{u}{2} + \frac{1}{2} = \frac{u+1}{2}$$

$$\therefore I = \int_0^1 \frac{dx}{1+x} = \int_{-1}^{1} \frac{\cancel{2}}{(u+3)} \frac{du}{\cancel{2}} = \int_{-1}^{1} \frac{du}{u+3}$$

$$= \omega_1 f(u_1) + \omega_2 f(u_2) + \omega_3 f(u_3)$$

$\omega_1 = 0.55555$ $\quad f(-0.77460) = \frac{1}{(-0.77460)+3}$

$\omega_2 = 0.88889$ $\quad f(0) = \frac{1}{0+3}$

$\omega_3 = 0.55555$ $\quad f(0.77460) = \frac{1}{0.77460+3}$

$I = 0.6931$

Romberg's Algorithm

Romberg's Algorithm provides a simple modification to the quadrature formulae for finding their better approximation. Let us here improve upon the value of integral

$$I = \int_a^b f(x)\,dx$$

by Trapezoidal rule. If I_1 & I_2 be the values of I with subintervals of width h_1, h_2 & E_1, E_2 be their corresponding errors resp., then

$$E_1 = -\frac{(b-a)h_1^2}{12} y''(x)$$

$$E_2 = -\frac{(b-a)^2 h_2^2}{12} y''(\bar{x})$$

Since $y''(\bar{x})$ is also the largest value of $y''(x)$, we can reasonably assume that $y''(x)$ & $y''(\bar{x})$ are very nearly equal.

$$\frac{E_1}{E_2} = \frac{h_1^2}{h_2^2} \quad \text{or} \quad \frac{E_1}{E_2 - E_1} = \frac{h_1^2}{h_2^2 - h_1^2} \qquad -(1)$$

Now since $I = I_1 + E_1 = I_2 + E_2$

$$\therefore \quad E_2 - E_1 = I_1 - I_2 \qquad -(2)$$

from (1) & (2), we have

$$E_1 = \frac{h_1^2}{h_2^2 - h_1^2}(I_1 - I_2)$$

Hence $I = I_1 + E_1 = I_1 + \frac{h_1^2}{h_2^2 - h_1^2}(I_1 - I_2)$

$$\therefore I = \frac{I_1 h_2^2 - I_2 h_1^2}{h_2^2 - h_1^2} \qquad -(3)$$

which is better approximation of I

To evaluate I systematically, we take $h_1 = h$ & $h_2 = \frac{1}{2}h$

so (3) gives

$$I = \frac{I_1\left(\frac{h}{2}\right)^2 - I_2(h)^2}{\left(\frac{h}{2}\right)^2 - h^2} = \frac{4I_2 - I_1}{3}$$

$$I(h, h/2) = \frac{1}{3}\left[4I\left(\frac{h}{2}\right) - I(h)\right] \qquad -(4)$$

Now we use the trapezoidal rule several times successively halving h & apply (4) to each pair of values as per the following scheme

$I(h)$

$I(h, h/2)$

$I(h/2)$

$I(h, \frac{h}{2}, \frac{h}{4})$

$I(h/2, h/4)$

$I(h, \frac{h}{2}, \frac{h}{4}, \frac{h}{8})$

$I(h/4)$

$I(\frac{h}{2}, \frac{h}{4}, \frac{h}{8})$

$I(h/4, h/8)$

$I(h/8)$

The computation is continued till successive values are close to each other. An obvious advantage of this method is that the accuracy of the computed value is known at each step.

Unit III

System of Linear Algebraic Eq^n.

Let us consider the following system of eq^{n}s.

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

This is called a system of n algebraic eq^{n}s which is linear in $x_1, x_2 \dots x_n$. Here a_{ij} & b_i, $i = 1 \dots n$, $j = 1 \dots n$ are constants & x_i's are unknowns. This system can be written in matrix form as

$$AX = B \quad \text{where}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{n1} & \dots & \dots & a_{nn} \end{bmatrix} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

A is called coefficient matrix.

Existence of the solution —

The matrix $[A|B] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} & b_n \end{array}\right]$

is called Augmented matrix.

1) If rank(A) = rank(aug A) = no. of unknowns, then solution is unique
2) If rank(A) = rank(Aug A) < no of unknowns, infinitely many solutions exists
3) If rank(A) $\neq$ rank(Aug A), no solution exists for the system.

Def :

① Symmetric matrix — A matrix A is symmetric if

$$A^T = A \quad \text{or} \quad [a_{ij}] = [a_{ji}]$$

If $A = -A^T$, then matrix is called skew symmetric.

② Lower triangular / upper triangular matrix —
A matrix in which all elements above its main diagonal are zero, its called lower triangular matrix. On the other hand, if all elements below the main diagonal are zero, the matrix is called upper triangular matrix.

If the diagonal elements are '1's, the matrices are called unit lower triangular matrix & unit upper triangular matrix.

3) Orthogonal matrix- A matrix is called orthogonal if

$$A^T = A^{-1} \quad \text{or} \quad AA^T = A^TA = I$$

e.g $\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$.

4) Unitary matrix — A matrix A is unitary if $A^* = A^{-1}$
where $A^* = (\bar{A})^T$

5) Positive definite matrix — A matrix A is said to be positive definite if for any non zero column vector x; $x^TAx > 0$.

If all the leading principal minors are +ve, the matrix will be positive definite.

Gauss Elimination method —

We consider the eq^{ns}

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned} \quad \text{— ①}$$

Step 1: To eliminate x, from second & third equations

Assuming $a_1 \neq 0$, we eliminate x from the second eq. by subtracting (a_2/a_1) times the first equation from second equation. Similarly we eliminate x from the third equation by eliminating (a_3/a_1) times the first eq. from third equation.

We thus get the new system

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ b_2'y + c_2'z &= d_2' \\ b_3'y + c_3'z &= d_3' \end{aligned} \quad \text{—(2)}$$

(2) $- \frac{a_2}{a_2} \times$ (1)

(3) $- \frac{a_3}{a_1} \times$ (1)

The first eq is called the pivotal eq. & a_1 is called the first pivot.

Step 2: To eliminate y from third equation in (2).:

Assuming $b_2' \neq 0$, we eliminate y from the third eq of (2) by subtracting (b_3'/b_2') times the second eq from third eq.

We thus get the new system

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ b_2'y + c_2'z &= d_2 \\ c_3''z &= d_3'' \end{aligned} \quad \text{—(3)}$$

(3') $-$ (2') $\times \frac{b_3'}{b_2'}$

Here the second eq is the pivotal eq & b_2' is the new pivot.

Step 3: To evaluate the unknowns:

The values of x, y, z are found from the reduced system (3) by back substitution.

Observations —

1. The method will fail if any one of the pivots a_1, b_2', c_3'' becomes zero. In such cases, we rewrite the equations in a different order so that the pivots are non zero.
2. The eq with largest coefficient of x should be taken as first eq. Now to choose second eq, we see the largest coefficient of y from remaining eqs. This eq is second equation. Then the coefficient of z will be seen and so on. This procedure

is called partial pivoting.

eg: Apply Gauss Elimination method to solve the eqs.

$$x+4y-z=-5$$
$$x+y-6z=-12$$
$$3x-y-z=4$$

Using partial pivoting, we get.

$$3x-y-z=4$$
$$x+4y-z=-5$$
$$x+y-6z=-12$$

Step 1: Multiply first eq by $\frac{1}{3}$ & subtract from second equation and third eq.

$$3x-y-z=4$$
$$\frac{13}{3}y-\frac{2}{3}z=-\frac{19}{3}$$
$$\frac{4}{3}y-\frac{17}{3}z=-\frac{40}{3}.$$

Step 2: Multiply second eq. by $\frac{4}{13}$ & subtract from third eq.

$$3x-y-z=4$$
$$\frac{13}{3}y-\frac{2}{3}z=-\frac{19}{3}$$
$$-\frac{213}{39}z=-\frac{444}{39}.$$
$$z=\frac{444}{213}=\frac{148}{71}.$$

$$\begin{array}{r} 81 \\ -\cancel{243} \\ \cancel{237} \\ \cancel{213} \\ 71 \end{array}$$

$$z=\frac{148}{71}$$

Step 3: By Back substitution.

$$\frac{13}{3}y=-\frac{19}{3}+\frac{2}{3}z=-\frac{19}{3}+\frac{2}{3}\times\frac{148}{71}=-\frac{19}{3}+\frac{296}{213}$$
$$=\frac{-1349+296}{213}$$

$$y=-\frac{\overset{81}{\cancel{1053}}}{\underset{71}{\cancel{213}}}\times\frac{\cancel{3}}{\cancel{13}}=-\frac{81}{71}.$$

$$3x = 4 + z + y$$
$$= 4 + \frac{148}{71} - \frac{81}{71} = \frac{284 + 148 - 81}{71}$$

$$3x = \frac{351}{71}$$

$$x = \frac{\overset{117}{\cancel{351}}}{71} \times \frac{1}{\cancel{3}} = \frac{117}{71}.$$

$$-5 + \frac{148}{71} + 4 \times \frac{81}{71}$$
$$\frac{-5 \times 71 + 148 + 324}{71}$$
$$= -355 + 148 + 324$$

$$\begin{array}{r} 472 \\ 355 \\ \hline 117 \end{array}$$

eg: $2x + y + z = 10$
$3x + 2y + 3z = 18$
$x + 4y + 9z = 16.$

Using partial pivoting

$$3x + 2y + 3z = 18$$
$$x + 4y + 9z = 16$$
$$2x + y + z = 10$$

Step 1: Elimination of x.

Multiply 1st eq by $\frac{1}{3}$ subtract from second ←
" " " by $\frac{2}{3}$ " " third.

$$3x + 2y + 3z = 18$$
$$\frac{10}{3}y + 8z = 10$$
$$-\frac{1}{3}y - z = -2.$$

Step 2: $\left[(iii) + \frac{1}{10}(ii)\right]$, we get

$$3x + 2y + 3z = 18$$
$$\frac{10}{3}y + 8z = 10$$
$$-\frac{1}{5}z = -1$$

Step 3: Using back substitution

$$z = 5$$

$$\frac{10}{3}y = 10 - 8z = 10 - 8(5) = 10 - 40 = -30$$

$$y = -9$$

$$3x = 18 - 3z - 2y$$

$$= 18 - 3(5) - 2(-9) = 21$$

$$x = 7.$$

Concept of pivoting —

In Gauss elimination method suppose we have to solve

$$2x_2 + 3x_3 = 8$$
$$4x_1 + 6x_2 + 7x_3 = -3$$
$$2x_1 + x_2 + 6x_3 = 5$$

Here $a_{11} = 0$, which creates a problem of division by zero. Same problem arise when a coefficient is very close to zero. The technique of pivoting has been developed to partially avoid these problems.

To overcome this, we determine the largest available coefficient in the column below the pivot element. The row can then be switched so that the largest element is the pivot element. This is called partial pivoting.

It is used to avoid roundoff errors that could be caused when dividing every entry of a row by a pivot value that is relatively small in comparison to its remaining row entries.

Gauss Jordan Method – This is modification of Gauss elimination method. In this method, elimination of unknowns is performed not in the equations below but in the equations above also, ultimately reducing the system to a diagonal matrix form. from this the unknowns x, y, z can be obtained easily.

Let the system be

$$\begin{aligned} a_1 x + b_1 y + c_1 z &= d_1 \\ a_2 x + b_2 y + c_2 z &= d_2 \\ a_3 x + b_3 y + c_3 z &= d_3 \end{aligned} \Rightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix}$$

Steps: 1) Make $a_{11} = 1$ using either R_2 or R_3

2) Make a_{21} & a_{31} zero using a_{11}

3) Make $a_{22} = 1$ by division or using R_3

4) Make a_{12} & a_{32} zero using a_{22}.

5) Make $a_{33} = 1$

6) Make a_{13} & a_{23} zero using a_{33}.

Example: Solve the system of Eq. using Gauss Jordan method

$$\begin{aligned} 2x_1 + 4x_2 - 6x_3 &= -8 \\ x_1 + 3x_2 + x_3 &= 10 \\ 2x_1 - 4x_2 - 2x_3 &= -12 \end{aligned}$$

Solu:

$$\left[\begin{array}{ccc|c} 2 & 4 & -6 & -8 \\ 1 & 3 & 1 & 10 \\ 2 & -4 & -2 & -12 \end{array}\right]$$

$R_1 \to R_1/2$ we get

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & -4 \\ 1 & 3 & 1 & 10 \\ 2 & -4 & -2 & -12 \end{array}\right]$$

$R_2 \to R_2 - R_1$

$R_3 \to R_3 - 2R_1$

$$\left[\begin{array}{ccc|c} 1 & 2 & -3 & -4 \\ 0 & 1 & 4 & 14 \\ 0 & -8 & 4 & -4 \end{array}\right]$$

$a_{22} = 1$ already $\therefore$

$R_1 \to R_1 - 2R_2$

$R_3 \to R_3 + 8R_2$

$$\left[\begin{array}{ccc|c} 1 & 0 & -11 & -32 \\ 0 & 1 & 4 & 14 \\ 0 & 0 & 36 & 108 \end{array}\right]$$

$R_3 \to R_3/36$

$$\left[\begin{array}{ccc|c} 1 & 0 & -11 & -32 \\ 0 & 1 & 4 & 14 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$R_1 \to R_1 + 11R_3$

$R_2 \to R_2 - 4R_3$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

$\therefore \quad x_1 = 1 \quad x_2 = 2 \quad x_3 = 3.$

Factorization methods – A square matrix A can be decomposed / factorised into a product of two matrices L & U ie $A = LU$, where L & U are lower & upper triangular matrices respectively.

Doolittle's method –

Consider the eqns.

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \qquad (1)$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3.$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = LU = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Then we have $LUX = B$

Taking $UX = V$ we get

$$LV = B$$

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$v_1 = b_1$$
$$l_{21}v_1 + v_2 = b_2$$
$$l_{31}v_1 + l_{32}v_2 + v_3 = b_3$$

Solving this for v_1, v_2, v_3 we get $UX = V$

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

from this, we obtain x_1, x_2, x_3.

Now to compute L & U, we have $A = LU$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

$\therefore$ $u_{11} = a_{11}$ $\quad u_{12} = a_{12}$ $\quad u_{13} = a_{13}$

$l_{21} u_{11} = a_{21}$ $\quad l_{21} u_{12} + u_{22} = a_{22}$ $\quad l_{21} u_{13} + u_{23} = a_{23}$

$l_{21} = \dfrac{a_{21}}{a_{11}}$ $\quad u_{22} = a_{22} - l_{21} u_{12}$ $\quad u_{23} = a_{23} - l_{21} u_{13}$.

Now IIIrd row

$l_{31} u_{11} = a_{31}$ $\quad l_{31} u_{12} + l_{32} u_{22} = a_{32}$ $\quad l_{31} u_{13} + l_{32} u_{23} + u_{33} = a_{33}$

$l_{31} = \dfrac{a_{31}}{u_{11}}$ $\quad l_{32} = \dfrac{a_{32} - l_{31} u_{12}}{u_{22}}$ $\quad u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23}$

Thus we compute the elements of L & U in the following order -

1) First row of U
2) First Column of L
3) Second row of U
4) Second column of L
5) Third row of U.

eg: Solve the eqns

$$3x + 2y + 7z = 4$$
$$2x + 3y + z = 5$$
$$3x + 4y + z = 7$$

Solu: Let $A = LU$

$$\begin{bmatrix} 3 & 2 & 7 \\ 2 & 3 & 1 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

R_1 of U : $u_{11} = 3$ $\quad u_{12} = 2$ $\quad u_{13} = 7$

C_1 of L : $l_{21} u_{11} = 2$ $\quad l_{31} u_{11} = a_{31} = 3$

$l_{21} = 2/3$ $\quad l_{31} = 3/3 = 1$

R_2 of U : $l_{21} u_{12} + u_{22} = a_{22} \Rightarrow u_{22} = 3 - \frac{2}{3} \times 2 = \frac{5}{3}$

$l_{21} u_{13} + u_{23} = a_{23} \Rightarrow u_{23} = 1 - \frac{2}{3} \times 7 = -11/3$

C_2 of L : $l_{32} = \dfrac{a_{32} - l_{31} u_{12}}{u_{22}} \Rightarrow l_{32} = \dfrac{4 - 1 \times 2}{5/3} = \dfrac{6}{5}$

R_3 of U : $u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23}$

$= 1 - 1 \times 7 - \frac{6}{5} \times \left(-\frac{11}{3}\right) = 1 - 7 + \frac{22}{5} = -\frac{8}{5}$.

$\therefore$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1 & 6/5 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 7 \\ 0 & 5/3 & -11/3 \\ 0 & 0 & -8/5 \end{bmatrix}$$

writing $UX = V$, we get

$$\begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1 & 6/5 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 7 \end{bmatrix}$$

solving $v_1 = 4$ $\quad \frac{2}{3} v_1 + v_2 = 5$ or $v_2 = 7/3$

$v_1 + \frac{6}{5} v_2 + v_3 = 7$ or $v_3 = 1/5$

Hence original system becomes.

$$\begin{bmatrix} 3 & 2 & 7 \\ 0 & 5/3 & -11/3 \\ 0 & 0 & -8/5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 7/3 \\ 1/5 \end{bmatrix}.$$

$$3x + 2y + 7z = 4$$
$$\frac{5}{3}y - \frac{11}{3}z = \frac{7}{3}$$
$$-\frac{8}{5}z = \frac{1}{5}$$

By back substitution, we get

$z = -1/8 \quad y = 9/8 \quad \& \quad x = 7/8$

Crout's Algorithm – In crout's Algorithm the matrix A can be expressed as the product of lower & upper triangular matrices as $A = LU$, where

$$L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \quad \& \quad U = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

The system of eqs $AX = B$ can be now written as

We consider $LUX = B$

$UX = Y$, where Y is an unknown matrix.

Then $LY = B$

After finding Y, we write $UX = Y$ & solve for X.

eg:
$$5x_1 + 4x_2 + x_3 = 3.4$$
$$10x_1 + 9x_2 + 4x_3 = 8.8$$
$$10x_1 + 13x_2 + 15x_3 = 19.2$$

Solve the system of eqns using crout's method.

Soln: we write
$$\begin{bmatrix} 5 & 4 & 1 \\ 10 & 9 & 4 \\ 10 & 13 & 15 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

We have $l_{11} = 5$ $\quad l_{21} = 10$ $\quad l_{31} = 10$

$l_{11} u_{12} = 4$

$\therefore u_{12} = \frac{4}{5}$

$l_{21} u_{12} + l_{22} = 9$

$l_{22} = 9 - l_{21} u_{12}$

$= 9 - 10 \times \frac{4}{5}$

$= 1$

$l_{31} u_{12} + l_{32} = 13$

$l_{32} = 13 - l_{31} u_{12}$

$= 13 - 10 \times \frac{4}{5}$

$= 5$

$l_{11} u_{13} + 0 = 1$

$u_{13} = \frac{1}{l_{11}} = \frac{1}{5}$

$l_{21} u_{13} + l_{22} u_{23} = 4$

$l_{22} u_{23} = 4 - l_{21} u_{13}$

$= 4 - 10 \times \frac{1}{5}$

$l_{22} u_{23} = 2$

$u_{23} = \frac{2}{l_{22}} = 2$

$l_{31} u_{13} + l_{32} u_{23} + l_{33} = 15$

$l_{33} = 15 - l_{31} u_{13} - l_{32} u_{23}$

$= 15 - 10 \times \frac{1}{5} - 5 \times 2$

$= 3$

$$\therefore \quad A = \begin{bmatrix} 5 & 0 & 0 \\ 10 & 1 & 0 \\ 10 & 5 & 3 \end{bmatrix} \begin{bmatrix} 1 & 4/5 & 1/5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

Now $LY = B$

$$\begin{bmatrix} 5 & 0 & 0 \\ 10 & 1 & 0 \\ 10 & 5 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 3.4 \\ 8.8 \\ 19.2 \end{bmatrix}$$

$5y_1 = 3.4 \qquad y_1 = 0.68$

$10y_1 + y_2 = 8.8 \Rightarrow y_2 = 2$

$10y_1 + 5y_2 + 3y_3 = 19.2 \Rightarrow y_3 = 0.80$

$\therefore$ The given system can be written as $UX = Y$

$$\begin{bmatrix} 1 & 4/5 & 1/5 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.68 \\ 2 \\ 0.80 \end{bmatrix}$$

$x_3 = 0.80 \qquad x_2 + 2x_3 = 2 \Rightarrow x_2 = 0.40$

$x_1 + \frac{4}{5} x_2 + \frac{1}{5} x_3 = 0.68 \Rightarrow x_1 = 0.20.$

Cholesky's method—

This method deals with a special case when the given matrix is symmetric & positive definite. If A is symmetric matrix, it can be expressed as product of two matrices L & L^T where L is a lower triangular matrix i.e

$$A = LL^T \text{ or } (U^TU). \quad \text{i.e}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

For the system $AX = B$, we get

$$LL^TX = B$$

We take $L^TX = Y$ so that

$$LY = B.$$

Solving for $LY = B$, we get the unknown vector Y

Using Y, we get $L^TX = Y$ & solve for X.

eg: Solve

$$x_1 - x_2 + 2x_3 = 17$$
$$-x_1 + 5x_2 - 4x_3 = 31$$
$$2x_1 - 4x_2 + 6x_3 = -5$$

Solu: Here $A = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 5 & -4 \\ 2 & -4 & 6 \end{bmatrix}$. We take $A = LL^T$

$$\begin{bmatrix} 1 & -1 & 2 \\ -1 & 5 & -4 \\ 2 & -4 & 6 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

$l_{11}^2 = 1$ $\quad$ $l_{11}l_{21} = -1$ $\quad$ $l_{11}l_{31} = 2$

$\therefore l_{11} = 1$ $\quad$ $l_{21} = -\frac{1}{1} = -1$ $\quad$ $l_{31} = \frac{2}{1} = 2$

$l_{21} \cdot l_{11} = -1$

$l_{21} = -1$

$l_{21}^2 + l_{22}^2 = 5$

$l_{22}^2 = 5 - (-1)^2$

$= 4$

$l_{22} = 2$

$l_{21} l_{31} + l_{22} l_{32} = -4$

$l_{32} = \dfrac{-4 - l_{21}}{l_{22}}$

$= \dfrac{-4 + 1 \times 2}{2}$

$= \dfrac{-4+2}{2} = -1$

$l_{31}^2 + l_{32}^2 + l_{33}^2 = 6$

$\therefore l_{33} = 1$

$\therefore$

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & -1 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix}}_{L^T}$$

Now

$$LL^T X = B$$

Let $L^T X = Y$ then $LY = B$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 17 \\ 31 \\ -5 \end{bmatrix}$$

$y_1 = 17$

$-y_1 + 2y_2 = 31 \Rightarrow y_2 = 24$

$2y_1 - y_2 + y_3 = -5 \Rightarrow y_3 = -15$

Now $L^T X = Y$

$$\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 17 \\ 24 \\ -15 \end{bmatrix}$$

$x_1 - x_2 + 2x_3 = 17 \Rightarrow x_1 = 51.5$

$2x_2 - x_3 = 24 \Rightarrow x_2 = 4.5$

$x_3 = -15$

By back substitution.

Eigen Value problem –

Power method – This method computes the numerically largest eigen value & the corresponding eigen vector of a matrix.

Let $X_1, X_2 \ldots X_n$ be eigen vectors of a matrix A, corresp-onding to eigen values $\lambda_1 \lambda_2 \ldots \lambda_n$; then any arbitrary column vector can be written as

$$X = K_1 X_1 + K_2 X_2 + \ldots + K_n X_n$$

$$AX = K_1 AX_1 + K_2 AX_2 + \ldots + K_n AX_n$$

$$= K_1 \lambda_1 X_1 + K_2 \lambda_2 X_2 + \ldots + K_n \lambda_n X_n$$

Then $A^r X = K_1 \lambda_1^r X_1 + K_2 \lambda_2^r X_2 + \ldots + K_n \lambda_n^r X_n$

(If $\lambda_1, \lambda_2 \ldots \lambda_n$ are eigen values of A, then $\lambda_1^p, \lambda_2^p \ldots \lambda_n^p$ are eigen values of A^p).

As r increases, & if $|\lambda_1| > |\lambda_2| > \ldots |\lambda_n|$, then the contribution of term $K_1 \lambda_1^r X_1$ to the sum on the right increases. So every time we increase r, the vector X becomes nearer to X_1 (if X_1 is an eigen vector then $(K_1 X_1)$ is also eigen vector). Finally we divide by K_1. The resulting vector is largest eigen vector.

To start the process, we take up a column vector X which is as near the solution as possible & evaluate AX which is written as $\lambda^{(1)} X^{(1)}$ after normalization. This gives the first approximation $\lambda^{(1)}$ to the eigen value & $X^{(1)}$ to the eigen ~~value~~ vector. Similarly, we evaluate $AX^{(1)} = \lambda^{(2)} X^{(2)}$ which gives IInd approximation. We repeat this process till $[X^r - X^{r-1}]$ becomes negligible. Then $\lambda^{(r)}$ will be the largest eigen value & $X^{(r)}$, the

corresponding eigen vector.

Note: Since we have

$$A^{-1}X = \frac{1}{\lambda}X$$

This equation, if we use, will yields the smallest eigen value.

eg: Determine the largest eigen value & the corresponding eigen vector of the matrix

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

Solu: Let the initial vector be $[1, 0, 0]'$. Then

$$AX = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix}$$

So the first approximation to the eigen value is 2 & eigen vector is $X^{(1)} = [1, -0.5, 0]'$

Now $$AX^{(1)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 2.5 \\ -2 \\ 0.5 \end{bmatrix} = 2.5\begin{bmatrix} 1 \\ -0.8 \\ 0.2 \end{bmatrix}$$

Repeating the process taking $X^{(2)} = \begin{bmatrix} 1 \\ -0.8 \\ 0.2 \end{bmatrix}$

$$AX^{(2)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -0.8 \\ 0.2 \end{bmatrix}$$

$$= \begin{bmatrix} 2.8 \\ -2.8 \\ 1.2 \end{bmatrix} = 2.8\begin{bmatrix} 1 \\ -1 \\ 0.428 \end{bmatrix}$$

$$X^{(3)} = \begin{bmatrix} 1 \\ -1 \\ 0.428 \end{bmatrix}.$$

$$AX^{(3)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0.428 \end{bmatrix}$$

$$= \begin{bmatrix} 3 \\ -3.428 \\ 1.856 \end{bmatrix} = 3.428 \begin{bmatrix} 0.875 \\ -1 \\ 0.541 \end{bmatrix}$$

$$X^{(4)} = \begin{bmatrix} 0.875 \\ -1 \\ 0.541 \end{bmatrix}$$

$$AX^{(4)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0.875 \\ -1 \\ 0.541 \end{bmatrix} = \begin{bmatrix} 2.75 \\ -3.416 \\ 2.082 \end{bmatrix} = 3.416 \begin{bmatrix} 0.80 \\ -1 \\ 0.609 \end{bmatrix}$$

Illy $X^{(5)} = \begin{bmatrix} 0.80 \\ -1 \\ 0.609 \end{bmatrix}$ $X^{(6)} = \begin{bmatrix} 0.76 \\ -1 \\ 0.65 \end{bmatrix}$

$$AX^{(5)} = \underbrace{\begin{bmatrix} 0.76 \\ -1 \\ 0.65 \end{bmatrix}}_{X^{(6)}} \underbrace{0.341}_{\lambda^{(6)}} \qquad AX^{(6)} = \underbrace{3.41}_{\lambda^{(7)}} \underbrace{\begin{bmatrix} 0.74 \\ -1 \\ 0.67 \end{bmatrix}}_{X^{(7)}}$$

$\lambda^{(6)}$ & $\lambda^{(7)}$ are same & $X^{(6)}$ & $X^{(7)}$ are approximately same so we take largest eigen value as 3.41 & corresponding eigen vector as $\begin{bmatrix} 0.74 \\ -1 \\ 0.67 \end{bmatrix}$.

Approximation by Spline function –

A spline function is a function that consists of polynomial pieces joined together with certain smoothness conditions. Let us assume, without loss of generality, that $x_0 < x_1 < \ldots < x_n$. Instead of a single polynomial for the entire domain (x_0, x_n) we can approximate the fn by several polynomials defined over subdomains of (x_0, x_n).

A simple example is a polygonal fn. (also called spline fn of degree 1) whose pieces are linear polynomials joined together to acheive continuity, as in fig.

The points $x_0, x_1 \ldots x_4$ at which the fn. changes its character are termed as Knots. We can define

$$S(x) = \begin{cases} S_0(x) & x_0 < x < x_1 \text{ or } x \in [x_0, x_1] \\ S_1(x) & x \in [x_1, x_2] \\ \vdots & \vdots \\ S_4(x) & x \in [x_3 \in x_4] \end{cases} \quad \text{—(1)}$$

where $S_i(x) = a_i x + b_i \quad \forall\ i = 0, 1, 2, 3, 4.$

because each piece of $S(x)$ is a linear polynomial. Such a fn. is piecewise linear.

If the fn S defined by (1) is <u>continuous</u>, it is called first degree spline.

Outside the interval $[x_0, x_n]$, $S(x)$ is usually defined to be the same fn on the left of x_0 as it is on the leftmost

subinterval $[x_0, x_1]$ & the same on right of x_n as it is on the rightmost subinterval $[x_{n-1}, x_n]$.

Continuity of a fn at a point of spline is defined in usual manner.

e.g: $s(x) = \begin{cases} x & x \in [-1, 0] \\ 1-x & x \in (0,1) \\ 2x-2 & x \in [1,2] \end{cases}$

Can not be a first degree spline, since it is not continuous at $x=0$.

$$\lim_{x \to 0^-} s(x) = 0 \qquad \lim_{x \to 0^+} s(x) = 1$$

The spline fn. of degree 1 can be used for interpolation.

Suppose we are given

x	x_0	x_1	$\dots$	x_n
y	y_0	y_1	$\dots$	y_n

Let $x_0 < x_1 < x_2 \dots < x_n$. The pts $(x_0, y_0) \dots (x_n\, y_n)$ can be represented on a plane. These points can be joined by polygonal lines. The eqn of line joining (x_i, y_i) & (x_{i+1}, y_{i+1}) is

$$y - y_i = \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i) \Rightarrow y = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i)$$

OR

$$s_i(x) = y_i + \frac{y_{i+1} - y_i}{x_{i+1} - x_i}(x - x_i)$$

where $(y_{i+1} - y_i)/(x_{i+1} - x_i)$ represents the slope of line joining $(x_i\, y_i)$ & (x_{i+1}, y_{i+1})

Second degree splines –

Let Q be a second degree spline then

$$Q(x) = \begin{cases} Q_1(x) & x \in [x_0, x_1] \\ \vdots & \\ Q_n(x) & x \in [x_{n-1}, x_n] \end{cases}$$

where all Q_i's are atmost quadratic fn. Also Q as well as $Q'(x)$ are continuous. i.e Quadratic spline is continuously differentiable

eg:

$$Q(x) = \begin{cases} x^2 & -10 \le x \le 0 \\ -x^2 & 0 \le x \le 1 \\ 1-2x & 1 \le x \le 20 \end{cases}$$

we have all $Q_i(x)$ are quadratic or linear. Also Q_i's are continuous at $x=0$ & $x=1$ (ie at Knots).

Now to check continuity of $Q'(x)$

$$\lim_{x \to 0^-} Q'(x) = \lim_{x \to 0^-} 2x = 0$$

$$\lim_{x \to 0^+} Q'(x) = \lim_{x \to 0^+} (-2x) = 0$$

$\therefore$ Continuous at $x=0$

$$\lim_{x \to 1^-} Q'(x) = \lim_{x \to 1^-} (-2x) = -2$$

$$\lim_{x \to 1^+} Q'(x) = \lim_{x \to 1^+} (-2) = -2$$

$\therefore$ Continuous at $x=1$

$\therefore$ the fn $Q(x)$ is a spline of degree 2.

Interpolating quadratic splines –

Suppose the table of given values is

x	x_0	x_1	...	x_n
y	y_0	y_1	....	y_n

A quadratic spline consists of n quadratic fns

$$x \rightarrow a_i x^2 + b_i x + c_i$$

one for each subinterval $[x_i, x_{i+1}]$

Also on each subinterval, the spline fn Q_i must satify

$$Q_i(x_i) = y_i$$

$$Q_i(x_{i+1}) = y_{i+1}$$

Since $Q(x)$ is continuously differentiable on $[x_0, x_n]$

we can put $z_i = Q_i'(x_i)$

We take up the following formula for Q_i

$$Q_i(x) = \frac{z_{i+1} - z_i}{2(x_{i+1} - x_i)} (x - x_i)^2 + z_i (x - x_i) + y_i$$

To check, if this is correct, we verify that

$$Q_i(x_i) = y_i$$

$$Q_i'(x_i) = z_i \quad \& \quad Q_i'(x_{i+1}) = z_{i+1}.$$

Hence $Q(x)$ is continuous as well as $Q_i'(x)$ is continuous. On simplification, we get

$$z_{i+1} = -z_i + 2\left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i}\right) \qquad 0 \le i \le n-1$$

eg: find a quadratic spline interpolant for the data

x	-1	0	$\frac{1}{2}$	1	2	$\frac{5}{2}$
y	2	1	0	1	2	3

assuming $z_0 = 0$

Solu:

$$Q(x) = \begin{cases} Q_0(x) & x \in [-1, 0] \\ Q_1(x) & x \in [0, \frac{1}{2}] \\ Q_2(x) & x \in [\frac{1}{2}, 1] \\ Q_3(x) & x \in [1, 2] \\ Q_4(x) & x \in [2, \frac{5}{2}] \end{cases}$$

where $Q_i(x) = \dfrac{z_{i+1} - z_i}{2(x_{i+1} - x_i)} (x - x_i)^2 + z_i(x - x_i) + y_i$

& $z_{i+1} = -z_i + 2\left(\dfrac{y_{i+1} - y_i}{x_{i+1} - x_i}\right)$.

Since $z_0 = 0$

$$z_1 = -z_0 + 2\left(\frac{y_1 - y_0}{x_1 - x_0}\right) = -0 + 2\left(\frac{1-2}{0-(-1)}\right) = 2\left(\frac{-1}{1}\right) = -2$$

$$z_2 = -z_1 + 2\left(\frac{y_2 - y_1}{x_2 - x_1}\right) = -(-2) + 2\left(\frac{0-1}{\frac{1}{2}-0}\right)$$
$$= 2 + 2\left(\frac{-1}{1/2}\right) = 2 - 4 = -2$$

$$z_3 = -z_2 + 2\left(\frac{y_3 - y_2}{x_3 - x_2}\right) = -(-2) + 2\left(\frac{1-0}{1-1/2}\right)$$
$$= 2 + 2\left(\frac{1}{1/2}\right) = 2 + 4 = 6$$

$$z_4 = -z_3 + 2\left(\frac{y_4 - y_3}{x_4 - x_3}\right) = -6 + 2\left(\frac{2-1}{2-1}\right) = -6 + 2 = -4$$

$$z_5 = -z_4 + 2\left(\frac{y_5 - y_4}{x_5 - x_4}\right) = -(-4) + 2\left(\frac{3-2}{\frac{5}{2}-2}\right) = 4 + 2\left(\frac{1}{1/2}\right)$$
$$= 4 + 4 = 8$$

$$Q_0(x) = \frac{z_1 - z_0}{2(x_1 - x_0)}(x - x_0)^2 + z_0(x - x_0) + y_0$$

$$= \frac{-2-0}{2(0+1)}(x+1)^2 + 0 + 2$$

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

$$Q_1(x) = \frac{z_2 - z_1}{2(x_2 - x_1)}(x - x_1)^2 + z_1(x - x_1) + y_1$$

$$= \frac{-2-(-2)}{2(\frac{1}{2}-0)}(x-0)^2 + (-2)(x-0) + 1$$

$$= -2x + 1$$

$$Q_2(x) = \frac{z_3 - z_2}{2(x_3 - x_2)}(x - x_2)^2 + z_2(x - x_2) + y_2$$

$$= \frac{6-(-2)}{2(1-\frac{1}{2})}(x-\tfrac{1}{2})^2 + (-2)(x-\tfrac{1}{2}) + 0$$

$$= \frac{8}{1}(x-\tfrac{1}{2})^2 + (-2)(x-\tfrac{1}{2}) = 8(x-\tfrac{1}{2})^2 - 2(x-\tfrac{1}{2})$$

$$Q_3(x) = \frac{z_4 - z_3}{2(x_4 - x_3)}(x - x_3)^2 + z_3(x - x_3) + y_3$$

$$= \frac{-4-6}{2(2-1)}(x-1)^2 + 6(x-1) + 1$$

$$= \frac{-10}{2}(x-1)^2 + 6(x-1) + 1 = -5(x-1)^2 + 6(x-1) + 1$$

$$Q_4(x) = \frac{z_5 - z_4}{2(x_5 - x_4)}(x - x_4)^2 + z_4(x - x_4) + y_4$$

$$= \frac{8-(-4)}{2(\frac{5}{2}-2)}(x-2)^2 + (-4)(x-2) + 2$$

$$= \frac{12}{1}(x-2)^2 - 4(x-2) + 2 = 12(x-2)^2 - 4(x-2) + 2$$

Natural cubic spline—

The 1st & 2nd degree splines have obvious imperfections due to the discontinuity of 1st & 2nd order derivatives respectively. In 1st order spline the slope of curve abruptly change at knots, whereas in 2nd degree spline the discontinuity is in second order derivative, because of which the curvature changes abruptly. So, whenever more smoothness is needed in approximation, higher degree splines are used. The choice of degree most frequently made for a spline fn is 3. The resulting splines are termed cubic spline, where we join cubic polynomials together in such a way that the resulting fn. has two continuous derivatives everywhere, so, at each knot three continuity conditions are imposed. Since S, S', S'' are continuous, the graph of the fn will appear smooth to the eye. Discontinuities, may occur in the third derivative, but cannot be easily detected visually. Moreover the splines of degree greater than 3 seldom yields any advantage. In general, odd degree splines behave better than even degree splines.

Let us consider the data

x	x_0	x_1	x_2	$\dots$	x_n
y	y_0	y_1	y_2	$\dots$	y_n

We define
$$S(x) = \begin{cases} S_0(x) & x \in [x_0, x_1] \\ S_1(x) & x \in [x_1, x_2] \\ \vdots & \\ S_{n-1}(x) & x \in [x_{n-1}, x_n] \end{cases}$$

where $S_i(x)$ denotes the cubic polynomial. We have

$$S(x_i) = y_i \quad \forall \; i = 0, 1 \dots n$$

The continuity conditions are S' & S'' are continuous

at Knots, i'e $\underset{x \to x_i^-}{Lt} S(x_i) = \underset{x \to x_i^+}{Lt} S(x_i)$

$$\underset{x \to x_i^-}{Lt} S'(x_i) = \underset{x \to x_i^+}{Lt} S'(x_i)$$

$$\& \ \underset{x \to x_i^-}{Lt} S''(x_i) = \underset{x \to x_i^+}{Lt} S''(x_i)$$

In addition, we have to consider

$$S''(x_0) = S''(x_n) = 0$$

The resulting spline function is termed as natural cubic spline.

Using the conditions of continuity, we get

$$S_i(x) = \frac{z_{i+1}}{6h_i}(x - x_i)^3 + \frac{z_i}{6h_i}(x_{i+1} - x)^3 + \left(\frac{y_{i+1}}{h_i} - \frac{h_i}{6} z_{i+1}\right)(x - x_i) +$$

$$+ \left(\frac{y_i}{h_i} - \frac{h_i}{6} z_i\right)(x_{i+1} - x)$$

where $x_{i+1} - x_i = h_i \ \forall \ 0 \le i \le n-1$.

To obtain the values of z_i's, the eqs are

$$h_{i-1} z_{i-1} + 2(h_{i-1} + h_i) z_i + h_i z_{i+1} = 6(b_i - b_{i-1})$$

where $b_i = \frac{1}{h_i}(y_{i+1} - y_i)$ for $0 \le i \le n-1$

If we take $2(h_{i-1} + h_i) = u_i$

$6(b_i - b_{i-1}) = v_i$; we obtain a tridiagonal system of linear Eqns.

$$z_0 = 0$$

$$h_{i-1} z_{i-1} + u_i z_i + h_i z_{i+1} = v_i \quad 1 \le i \le n-1$$

$$z_n = 0$$

Solving this tridiagonal system of linear Eqns, we obtain the values of Z_i's. $Z_0 = 0$ & $Z_n = 0$ are obtained due to conditions $S''(x_0) = S''(x_n) = 0$

eg: Derive the eqns of natural cubic spline for the following table:

x	-1 (x_0)	0 (x_1)	1 (x_2)
y	1 (y_0)	2 (y_1)	-1 (y_2)

Solu: we first obtain $h_0 = x_1 - x_0 = 1$

$h_1 = x_2 - x_1 = 1$

Next $b_1 = \frac{1}{h_1}(y_2 - y_1) = \frac{1}{1}(-1-2) = -3$

$b_0 = \frac{1}{h_0}(y_1 - y_0) = \frac{1}{1}(2-1) = 1$

$u_1 = 2(h_0 + h_1) = 2(1+1) = 4$

$v_1 = 6(b_1 - b_0) = 6(-3-0) = -24$

$\therefore$ The TD system becomes

$$\begin{cases} Z_0 = 0 \\ h_0 Z_0 + u_1 Z_1 + h_1 Z_2 = v_1 \\ Z_2 = 0 \end{cases}$$

$$\therefore \begin{cases} Z_0 = 0 \\ Z_0 + 4Z_1 + Z_2 = -24 \\ Z_2 = 0 \end{cases}$$

Solving we obtain $Z_0 = 0$, $Z_1 = -6$, $Z_2 = 0$

$\therefore$

$$S_0(x) = \frac{z_1}{6h_0}(x-x_0)^3 + \frac{z_0}{6h_0}(x_1-x)^3 + \left(\frac{y_1}{h_0} - \frac{h_0}{6}z_1\right)(x-x_0) +$$

$$\left(\frac{y_0}{h_0} - \frac{h_0}{6}z_0\right)(x_1-x)$$

$$= \frac{-\cancel{6}}{\cancel{6}\times 1}(x-(-1))^3 + 0 + \left(\frac{2}{1} - \frac{1.(-\cancel{6})}{\cancel{6}}\right)(x-(-1))$$

$$+ \left(\frac{1}{1} - \frac{1}{6}\times 0\right)(0-x)$$

$$= -(x+1)^3 + 3(x+1) - x \quad ; \quad x \in [-1,0]$$

$$S_1(x) = \frac{z_2}{6h_1}(x-x_1)^3 + \frac{z_1}{6h_1}(x_2-x)^3 + \left(\frac{y_2}{h_1} - \frac{h_1 z_2}{6}\right)(x-x_1)$$

$$+ \left(\frac{y_1}{h_1} - \frac{h_1 z_1}{6}\right)(x_2-x)$$

$$= 0 + \frac{(-\cancel{6})}{\cancel{6}\cdot 1}(1-x)^3 + \left(\frac{-1}{1} - \frac{1.0}{6}\right)(x-0)$$

$$+ \left(\frac{2}{1} - \frac{1.(-6)}{\cancel{6}}\right)(1-x)$$

$$= -(1-x)^3 - x + 3(1-x)$$

$$= -(1-x)^3 - x + 3 - 3x \quad ; \quad x \in [0,1].$$

An alternate form to write Natural cubic spline is

$$S_i(x) = y_i + (x-x_i)\left[B_i + (x-x_i)\left(\frac{z_i}{2} + \frac{1}{6h_i}(x-x_i)\cdot(z_{i+1}-z_i)\right)\right]$$

where

$$B_i = -\frac{h_i}{6} z_{i+1} - \frac{h_i z_i}{3} + \frac{1}{h_i}(y_{i+1}-y_i).$$

B spline –

Let us consider an infinite set of Knots t_i defined as

$$\ldots < t_{-2} < t_{-1} < t_0 < t_1 < t_2 < \ldots$$

$$\lim_{i\to\infty} t_i = \infty = -\lim_{i\to\infty} t_{-i}$$

We define B_i^0 spline (i.e B spline of degree 0) as

$$B_i^0(x) = \begin{cases} 1 & t_i \le x < t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

The graph of B_i^0 is as follows

Obviously B_i^0 is discontinuous. However, it is continuous from the right at all points even where the jump occurs.

Thus

$$\lim_{x\to t_i^+} B_i^0(x) = 1 = B_i^0(t_i)$$

$$\lim_{x\to t_{i+1}^+} B_i^0(x) = 0 = B_i^0(t_{i+1})$$

Since B_i^0 is a piecewise constant fn., it is a spline of degree 0. Also

$$B_i^0(x) \geq 0 \quad \forall x \ \& \ \forall i$$

$$\sum_{i=-\infty}^{\infty} B_i^0(x) = 1 \quad \forall x.$$

With the fns B_i^0 as a starting point, we now generate all higher degree B splines by a simple recursive definition:

$$B_i^k(x) = \left(\frac{x - t_i}{t_{i+k} - t_i}\right) B_i^{k-1}(x) + \left(\frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}}\right) B_{i+1}^{k-1}(x) \qquad (k \geq 1)$$

where $k = 1, 2, 3 \ldots$ & $i = 0, \pm 1, \pm 2 \ldots$

Hence B_i^1 is defined as

$$B_i^1(x) = \left(\frac{x - t_i}{t_{i+1} - t_i}\right) B_i^0(x) + \left(\frac{t_{i+2} - x}{t_{i+2} - t_{i+1}}\right) B_{i+1}^0(x)$$

$$= \begin{cases} 0 & (x \geq t_{i+2} \text{ or } t_i \geq x) \\ \dfrac{x - t_i}{t_{i+1} - t_i} & (t_i \leq x < t_{i+1}) \\ \dfrac{t_{i+2} - x}{t_{i+2} - t_{i+1}} & (t_{i+1} \leq x < t_{i+2}) \end{cases}$$

Also $\displaystyle\sum_{i=-\infty}^{\infty} B_i^1(x) = 1 \quad \forall x$

The fns B_i^k defined above are called B splines of degree k.

Interpolation & approximation by B splines –

we are given

$$x : t_0 \; t_1 \; \ldots \; t_n$$

$$y : y_0 \; y_1 \; \ldots \; y_n$$

We want to find coefficients A_i s.t

$$S(x) = \sum_{i=-\infty}^{\infty} A_i \, B_{i-k}^{k}(x)$$

& $$S(t_i) = y_i \qquad (0 \le i \le n)$$

for zero degree spline, it is very simple

$$S(x) = \sum_{i=0}^{\infty} y_i \, B_i^0(x)$$

$$\because \; B_i^0(t_j) = \delta_{ij} = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}$$

for degree 1 spline, we have

$$S(x) = \sum_{i=0}^{n} y_i \, B_{i-1}^{1}(x)$$

$$\because \; B_{i-1}^{1}(t_j) = \delta_{ij}$$

for degree 2 spline,

$$S(x) = \sum_{i=-\infty}^{\infty} A_i \, B_{i-k}^{2}(x)$$

where

$$A_j(t_{j+1} - t_j) + A_{j+1}(t_j - t_{j-1}) = y_j(t_{j+1} - t_{j-1}) \qquad (0 \le j \le n)$$

This is a system of $(n+1)$ linear eqns in $(n+2)$ unknowns $A_0, A_1, A_2 \ldots A_{n+1}$, that can be solved recursively.

Numerical solutions of Ordinary differential Equations

A number of problems in science & technology can be formulated into differential equations. The analytical methods of solving differential equations are applicable only to a limited type of eq. Quite often differential equations appearing in physical problems do not belong to any of these familiar types & one is obliged to resort to numerical methods. These methods are of even greater importance when we realise that computing machines are now readily available which reduce numerical work considerably.

Solution of a differential Equation – The solution of an ordinary differential equation means finding an explicit expression for y in terms of a finite number of elementry functions of x. Such a solution of a differential equation is known as the closed or finite form of solution. In the absence of such a solution, we have recourse to numerical methods of solution.

Let us consider the first order differential equation

$$\frac{dy}{dx} = f(x,y) \quad ; \text{ given } y(x_0) = y_0$$

to study the various numerical methods of solving such equations. In most of these methods, we replace the differential equation by a difference equation & then solve it. These methods yield solutions either as a power series in x from which the values of y can be found by direct substitution or a set of values of x & y. The Picard's method belong to the former class of solutions. These methods are referred to as single step methods. The methods of Euler, Runga-Kutta & Milne predictor belong to the latter class of solutions. These methods are arealso known as step-by step methods.

Initial & boundary conditions – An ordinary differential equation of the nth order is of the form

$$F(x, y, dy/dx, d^2y/dx^2, \ldots\ldots d^ny/dx^n) = 0$$

Its general solution contains n arbitrary constants & is of form

$$\phi(x, y, C_1, C_2 \ldots C_n) = 0$$

To obtain its particular solution, n conditions must be given so that the constants $C_1, C_2 \ldots C_n$ can be determined. If these conditions are prescribed at one point (say x_0) then the differential equation together with the condition constitute an initial value problem of nth order. If the conditions are prescribed at two or more points, then the problem is termed as boundary value problem. We will only consider initial value problems here.

Picard's Method –

Consider the first order equation

$$\frac{dy}{dx} = f(x, y) \qquad \text{— (1)}$$

It is required to find that particular solution of (1) which assumes the value y_0 when $x = x_0$.

Integrating (1) between limits, we get

$$\int_{y_0}^{y} dy = \int_{x_0}^{x} f(x, y)\, dx$$

$$y = y_0 + \int_{x_0}^{x} f(x, y)\, dx. \qquad \text{— (2)}$$

This is an integral equation equivalent to (1), for it contains the unknown y under the integral sign.

As a first approximation y_1 to the solution, we put $y = y_0$ in $f(x, y)$ & integrating (2), we get

$$y_1 = y_0 + \int_{x_0}^{x} f(x, y_0)\, dx.$$

For a second approximation y_2, we put $y = y_1$ in $f(x, y)$

Integrate ②, giving

$$y_2 = y_0 + \int_{x_0}^{x} f(x, y_1)\,dx.$$

Similarly third approximation is

$$y_3 = y_0 + \int_{x_0}^{x} f(x, y_2)\,dx$$

Continuing, this process, we obtain $y_4, y_5, \ldots y_n$ where

$$y_n = y_0 + \int_{x_0}^{x} f(x, y_{n-1})\,dx.$$

Hence this method gives a sequence of approximations $y_1, y_2, \ldots y_n$. each giving a better result than the preceding one.

Note: This can be applied only to a limited class of eqs in which the successive integrations can be performed easily.

e.g: Using Picard's method, obtain a solution upto fifth approximation of the equation $dy/dx = y+x$. such that $y=1$ when $x=0$.

solu: we have $y = 1 + \int_0^x f(x,y)\,dx$

$$y = 1 + \int_0^x (x+y)\,dx$$

First approx: Put $y=1$ in $(y+x)$, we get

$$y_1 = 1 + \int_0^x (1+x)\,dx = 1 + \left[x + \frac{x^2}{2}\right]_0^x = 1 + x + \frac{x^2}{2}.$$

second approx: Put $y = 1 + x + \frac{x^2}{2}$ in $(y+x)$, we get

$$y_2 = 1 + \int_0^x \left(1 + x + \frac{x^2}{2} + x\right)dx$$

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^2}{2} = 1 + x + x^2 + \frac{x^3}{6}.$$

third approx: Put $y = 1 + x + x^2 + \frac{x^3}{6}$ in $(y+x)$, we get

$$y_3 = 1 + \int_0^x \left(1 + x + x^2 + \frac{x^3}{6} + x\right) dx$$

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{24} + \frac{x^2}{2}$$

$$= 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{24}.$$

Fourth approx: Putting $y = y_3$; we get

$$y_4 = 1 + \int_0^x \left(1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{24} + x\right) dx$$

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{12} + \frac{x^5}{120} + \frac{x^2}{2}.$$

$$= 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{12} + \frac{x^5}{120}$$

Fifth approx: Putting $y = y_4$ in $(y + x)$, we get

$$y_5 = 1 + \int_0^x \left(1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{12} + \frac{x^5}{120} + x\right) dx$$

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{12} + \frac{x^5}{60} + \frac{x^6}{720} + \frac{x^2}{2}$$

$$= 1 + x + x^2 + \frac{x^3}{3} + \frac{x^4}{12} + \frac{x^5}{60} + \frac{x^6}{720}.$$ is the solu.

e.g: Find an approximate value of y when $x = 0.1$, if $dy/dx = x - y^2$ and $y = 1$ at $x = 0$ using Picard's method.

Solu:

$$y = 1 + \int_0^x (x - y^2)\, dx$$

First approx: Put $y = 1$, we get

$$y_1 = 1 + \int_0^x (x - 1)\, dx = 1 + \frac{x^2}{2} - x = 1 - x + \frac{x^2}{2}$$

econd approx : Put $y = 1 - x + \frac{x^2}{2}$, we get

$$y_2 = 1 + \int_0^x \left[x - \left[1 - x + \frac{x^2}{2}\right]^2\right] dx$$

$$= 1 + \int_0^x x - \left(1 + x^2 + \frac{x^4}{4} - 2x - x^3 + x^2\right) dx$$

$$= 1 + \int_0^x \left(-1 + 3x + 2x^2 + (-x^3) + \frac{x^4}{4}\right) dx$$

$$= 1 - x + \frac{3x^2}{2} + \frac{2x^3}{3} - \frac{x^4}{4} + \frac{x^5}{20}.$$

$$y_2\big|_{0.1} = 0.9155$$

Ans: 0.9138.

e.g: Use picard's method to solve $dy/dx = 2x - y$ given that $y = 0.9$ at $x = 0$. Find value of y at $x = 0.2, 0.4, 0.6$.

Ans : 0.7743, 0.7439, 0.7915.

Euler's Method –

This method is very elementry and should not be used for practical solutions. However, the principal involved is of vital importance. More accurate methods which can be used in practice are given later.

Let $x_0, x_1, x_2 \ldots$ be equispaced values of x at a spacing h and let $y_0, y_1, y_2 \ldots$ be the corresponding values of y on a curve satisfying the differential eq.

$$\frac{dy}{dx} = f(x, y), \quad y(x_0) = y_0 \qquad \text{—— (1)}$$

suppose the solution of (1) is $y = \phi(x)$.

For the first step we find the value of y at $x = x_0 + h$, & call it y_1.

then $y_1 = \phi(x_1) = \phi(x_0 + h)$

$$= \phi(x_0) + h\phi'(x_0) + \frac{h^2}{2!}\phi''(x_0) + \cdots$$

$$= y_0 + hf(x_0, y_0) + \frac{h^2}{2!} f'(x_0, y_0) \qquad \text{(using ①)}$$

Hence if h is small, neglecting terms involving h^2 & higher powers of h, we have

$$y_1 = y_0 + hf(x_0, y_0) \qquad \text{— ②}$$

On the same lines, for the second step, we take (x_1, y_1) as the starting point and get

$$y_2 = y_1 + hf(x_1, y_1)$$

Continuing the process, we have in general

$$y_{n+1} = y_n + hf(x_n, y_n) \qquad \text{— ③}$$

This formula can be used to determine y_{n+1} when y_n is known. Putting in ③, $n = 0, 1, 2, \ldots$ we can successively determine $y_1, y_2, y_3, \ldots$ provided the starting point (x_0, y_0) is known.

e.g: Applying Euler's method, solve

$$\frac{dy}{dx} = x + y; \quad y(0) = 0$$

Choosing the step length $= 0.2$. Carry out 6 steps.

Solu: we take $n = 5$; $h = 0.2$. The calculations are arranged as follows:

x	y	$x+y = dy/dx$	old y + 0.2 (dy/dx) = new
0.0	0.0	0.0	0.0 + 0.2 × 0.0 = 0.0
0.2	0.0	0.2	0.0 + 0.2 × 0.2 = 0.04
0.4	0.04	0.4 + 0.04 = 0.44	0.04 + 0.2 × 0.44 = 0.128
0.6	0.128	0.6 + 0.128 = 0.728	0.128 + 0.2 × 0.728 = 0.27
0.8	0.2736	0.8 + 0.2736 = 1.0736	0.2736 + 0.2 × 1.0736 = 0.488
1.0	0.48832	1.0 + 0.48832 = 1.48832	0.48832 + 0.2 × 1.48 = 0.785

7859 $\quad 1.2+0.7859=1.9859 \quad 0.7859+0.2\times 1.9859$

$= 1.18308.$

Ans:

g : Using Euler's method, find an approx value of y corresponding to $x=1$, given that $dy/dx = x+y$ & $y=1$ when $x=0$.

Solu : we take $n=10$, $h=0.1$. The various calculations are as follows:

x	y	$x+y = dy/dx$	$old\,y + 0.1 \times dy/dx = New\,y$
0.0	1.00	1.00	$1.00+0.1\times(1.00) = 1.10$
0.1	1.10	1.20	$1.10+0.1(1.20) = 1.22$
0.2	1.22	1.42	$1.22+0.1(1.42) = 1.36$
0.3	1.36	1.66	$1.36+0.1(1.66) = 1.53$
0.4	1.53	1.93	$1.53+0.1(1.93) = 1.72$
0.5	1.72	2.22	$1.72+0.1(2.22) = 1.94$
0.6	1.94	2.54	$1.94+0.1(2.54) = 2.19$
0.7	2.19	2.89	$2.19+0.1(2.89) = 2.48$
0.8	2.48	3.29	$2.48+0.1(3.29) = 2.81$
0.9	2.81	3.71	$2.81+0.1(3.71) = 3.18.$
1.0	3.18		

$\therefore$ the req. approx value of $y = 3.18$.

Modified Euler's Method – According to modified Euler's method we find a better approximation of y_1 as $y_1^{(1)}$ by taking

$$y_1^{(1)} = y_0 + h\left[\frac{f(x_0,y_0)+f(x_0+h,y_1)}{2}\right] \quad \text{— (1)}$$

So the eq. $y_1 = y_0 + hf(x_0,y_0)$ is called the predictor whereas eq (1) is called corrector of y_1. Now insert $y_1^{(1)}$ as

y_1 in ① as we get a still better value of y_1 as $y_1^{(2)}$ g

by

$$y_1^{(2)} = y_0 + h\left[\frac{f(x_0, y_0) + f(x_0+h, y_1^{(1)})}{2}\right]$$

The process is repeated until two consecutive values of y agree. Once y_1 is obtained to the desired degree of accuracy, y corresponding to (x_0+2h) is found from predictor

$$y_2 = y_1 + hf(x_0+h, y_1)$$

& a better approximation $y_2^{(1)}$ is obtained by corrector

$$y_2^{(1)} = y_1 + h\left[\frac{f(x_0+h, y_1) + f(x_0+2h, y_2)}{2}\right]$$

The step is repeated till we get y_2 correct to desired degree of accuracy. Next we compute y_3 as above & so on. This predictor-corrector method is of great importance in practical problems.

e.g: Using modified Euler's method, find an approximate value of y when $x=0.2$ given that $dy/dx = x+y$ & $y=1$ when $x=0$.

Solu: To compute $y(0.2)$ we have $x_0 = 0$, $y_0 = 1$, $h = 0.1$.

$$f(x, y) = x+y \quad ; \quad f(x_0, y_0) = 0+1 = 1.$$

By predictor formula

$$y_1 = y_0 + h f(x_0, y_0)$$
$$= 1 + 0.1 \times [0+1] = 1.1$$

Then applying corrector formula

$$y_1^{(1)} = y_0 + h\left[\frac{f(x_0, y_0) + f(x_0+h, y_1)}{2}\right]$$
$$= 1 + 0.1\left[\frac{1 + (0.1+1.1)}{2}\right] = 1 + \frac{0.1 \times 2.2}{2} = 1 + 0.11$$
$$= 1.11$$

Again applying corrector.

$$y_1^{(2)} = y_0 + h\left[\frac{f(x_0, y_0) + f(x_0+h, y_1^{(1)})}{2}\right]$$

$$= 1 + 0.1\left[\frac{1 + (0.1 + 1.11)}{2}\right] = 1 + 0.1 \times \frac{2.21}{2} = 1.1105$$

Again applying corrector

$$y_1^{(3)} = y_0 + h\left[\frac{f(x_0, y_0) + f(x_0+h, y_1^{(2)})}{2}\right]$$

$$= 1 + 0.1\left[\frac{1 + (0.1 + 1.1105)}{2}\right] = 1 + 0.1 \times \frac{2.2105}{2} = 1.1105$$

Since $y_1^{(2)}$ & $y_1^{(3)}$ are equal upto third place at least we have

$y(0.1) = 1.1105.$

Now we compute $y(0.2)$. By predictor formula

$$y_1 = y_0 + hf(x_0, y_0)$$

Here we take $x_0 = 0.1$; $h = 0.1$, $y_0 = 1.1105$,

$$f(x_0, y_0) = 0.1 + 1.1105 = 1.2105.$$

$$\therefore \quad y_1 = 1.1105 + 0.1 \times 1.2105 = 1.2316$$

Using corrector.

$$y_1^{(1)} = y_0 + h\left[\frac{f(x_0, y_0) + f(x_0+h, y_1)}{2}\right]$$

$$= 1.1105 + 0.1\left[\frac{1.2105 + (0.2 + 1.2316)}{2}\right]$$

$$= 1.2426$$

$$y_1^{(2)} = y_0 + h\left[\frac{f(x_0, y_0) + f(x_0+h, y_1^{(1)})}{2}\right]$$

$$= 1.1105 + 0.1\left[\frac{1.2105 + (0.2 + 1.2426)}{2}\right]$$

$$= 1.2432.$$

$$y_1^{(2)} = y_0 + h\left[\frac{f(x_0,y_0) + f(x_0+h, y_1^{(1)})}{2}\right]$$

$$= 1.1105 + 0.1\left[\frac{1.2105 + (0.2 + 1.2432)}{2}\right]$$

$$= 1.2432$$

So $y(0.2) = 1.2432$ approx.

IIIly we can find $y(0.3) = 1.4004$. approx.

Runge Kutta's method of fourth order –

This method do not require the calculations of higher order derivatives and give greater accuracy. This method possess the advantage of requiring only the function values at some selected points. This method is most commonly used & is often referred to as Runge-Kutta method only.

Working rule for finding the increment k of y corresponding to an increment h of x by Runge Kutta method from

$$\frac{dy}{dx} = f(x,y) \; ; \; y(x_0) = y_0$$

is as follows :-

Calculate successively $k_1 = hf(x_0, y_0)$

$$k_2 = hf(x_0 + \tfrac{1}{2}h, y_0 + \tfrac{1}{2}k_1)$$

$$k_3 = hf(x_0 + \tfrac{1}{2}h, y_0 + \tfrac{1}{2}k_2)$$

$$k_4 = hf(x_0 + h, y_0 + k_3)$$

& finally compute $k = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$

which gives the required approximate value as –

$$y_1 = y_0 + k.$$

Note: One of the advantages of these methods is that the operation is identical whether the differential equation is linear or non linear.

e.g: Apply Runge-Kutta method (fourth order) to find an approximate value of y when $x=0.2$ given that $dy/dx = x+y$ and $y=1$, when $x=0$

Solu:

$$\frac{dy}{dx} = x+y \qquad y(0)=1.$$

$$x_0 = 0, \quad y_0 = 1, \quad f(x,y) = x+y, \quad h = 0.2$$

$$k_1 = h\,f(x_0, y_0) = 0.2[0+1] = 0.2$$

$$k_2 = h\,f(x_0 + \tfrac{1}{2}h,\ y_0 + \tfrac{1}{2}k_1)$$

$$= 0.2\left[\left(0+\tfrac{1}{2}0.2\right) + \left(1+\tfrac{1}{2}0.2\right)\right]$$

$$= 0.2[0.1 + 1.1] = 0.24.$$

$$k_3 = h\left[f\left(x_0 + \tfrac{1}{2}h,\ y_0 + \tfrac{1}{2}k_2\right)\right]$$

$$= 0.2\left[\left(0+\tfrac{1}{2}0.2\right) + \left(1+\tfrac{1}{2}0.24\right)\right]$$

$$= 0.2[0.1 + 1.12] = 0.2 \times 1.22 = 0.244$$

$$k_4 = h[f(x_0 + h,\ y_0 + k_3)]$$

$$= 0.2[(0+0.2) + (1+0.244)]$$

$$= 0.2[1.444] = 0.2888.$$

Finally $k = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$

$$= \frac{1}{6}[0.2 + 2\times 0.24 + 2\times 0.244 + 0.2888]$$

$$= \frac{1}{6}[0.2 + 0.48 + 0.488 + 0.2888]$$

$$= 0.2428.$$

$$\therefore\ y(0.2) = y_0 + k = 1 + 0.2428 = 1.2428.$$

e.g : Use Runge Kutta's method to approximate y when x given $y(1) = 1.2$ & $dy/dx = 3x + y^2$.

Solu : Given $\frac{dy}{dx} = 3x + y^2$; $x_0 = 1$, $y_0 = 1.2$; $h = 0.1$; $f(x,y) = 3$

Now $k_1 = hf(x_0, y_0)$

$= 0.1\left[3(1) + (1.2)^2\right] = 0.1\left[3 + 1.44\right] = 0.444.$

$k_2 = hf(x_0 + \frac{1}{2}h,\ y_0 + \frac{1}{2}k_1)$

$= 0.1\left[3(x_0 + \frac{1}{2}h) + (y_0 + \frac{1}{2}k_1)^2\right]$

$= 0.1\left[3(1 + \frac{1}{2}0.1) + (1.2 + \frac{1}{2}0.444)^2\right]$

$= 0.1\left[3.15 + 2.022\right] = 0.5172.$

$k_3 = hf(x_0 + \frac{1}{2}h,\ y_0 + \frac{1}{2}k_2)$

$= 0.1\left[3(x_0 + \frac{1}{2}h) + (y_0 + \frac{1}{2}k_2)^2\right]$

$= 0.1\left[3.15 + (1.2 + \frac{1}{2}0.5172)^2\right]$

$= 0.1\left[3.15 + 2.1275\right] = 0.5277$

$k_4 = hf(x_0 + h,\ y_0 + k_3)$

$= 0.1\left[3(x_0 + h) + (y_0 + k_3)^2\right]$

$= 0.1\left[3(1 + 0.1) + (1.2 + 0.5277)^2\right]$

$= 0.1\left[3.3 + 2.9849\right] = 0.6284.$

Now $k = \frac{1}{6}\left[k_1 + 2k_2 + 2k_3 + k_4\right]$

$= \frac{1}{6}\left[0.444 + 2 \times 0.5172 + 2 \times 0.5277 + 0.6284\right]$

$= \frac{1}{6} \times 2.1068 = 0.5270$

$\therefore\ y(1.1) = y_0 + k = 1.2 + 0.5270$

$= 1.7270.$

Taylor Series Method—

Consider the first order equation

$$\frac{dy}{dx} = f(x,y) \quad \text{— (1)}$$

Differentiating (1), we have

$$\frac{d^2y}{dx^2} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} \quad \text{— (2)}$$

i.e $y'' = f_x + f_y f'$

Differentiating this successively, we can get y''', $y^{(iv)}$ etc. Putting $x = x_0 = 0$ & $y = 0$ the values of $y'(0)$, $y''(0)$, $y'''(0)$ etc can be obtained. Hence the Taylor's series

$$y = y_0 + (x-x_0)(y')_0 + \frac{(x-x_0)^2}{2!}(y'')_0 + \frac{(x-x_0)^3}{3!}(y''')_0 + \cdots \quad \text{— (3)}$$

gives the value of y for every value of x for which (3) converges.

Note: This is a single step method but is useful only if the derivatives of higher orders can be calculated easily. If $f(x,y)$ is complicated & the calculation of higher orders derivatives becomes difficult then this method is not much useful. This is the main drawback of this method. However it can be used to find starting values for the calculation & application of powerful methods like Runge Kutta method, Milne predictor method etc which we will discuss later.

eg: Find an approximate value of y when $x = 0.1$ if $\frac{dy}{dx} = x - y^2$ & $y^{(0)} = 1$ using Taylor's method.

Solu: Since $\frac{dy}{dx} = x - y^2$

$$y' = x - y^2$$
$$y'' = 1 - 2yy'$$
$$y''' = 0 - 2y'y' - 2yy'' = -2(y')^2 - 2yy''$$
$$y^{iv} = -4y'y'' - 2y'y'' - 2yy'''$$
$$= -6y'y'' - 2yy'''$$

At $x=0$; $y=1$ i.e $y(0)=1$

$$y'(0) = -1$$
$$y''(0) = 3$$
$$y'''(0) = -8$$
$$y^{iv}(0) = 34.$$

Putting these values in the Taylor's series, we have.

$$y = y(0) + (x-0)y'(0) + \frac{(x-0)^2}{2!}y''(0) + \frac{(x-0)^3}{3!}y''(0) + \frac{(x-0)^4}{4!}y^{iv}(0) + \dots$$

$$= 1 + x(-1) + \frac{x^2}{2!}(3) + \frac{x^3}{3!}(-8) + \frac{x^4}{4!}(34) + \dots$$

So $y(0.1) = 1 + (0.1)(-1) + \frac{(0.1)^2}{2!}(3) + \frac{(0.1)^3}{3!}(-8) + \frac{(0.1)^4}{4!}(34) + \dots$

$$= 1 - 0.1 + 0.015 - 0.00133 + 0.000141$$
$$= 0.913.$$

eg: Find the value of y at $x = 0.1$ and $x = 0.2$ to five places of decimal from $\frac{dy}{dx} = x^2y - 1$; $y(0) = 1$

Solu: $\frac{dy}{dx} = x^2y - 1$

$y' = x^2y - 1$ $\qquad$ $y'(0) = -1$

$$y'' = 2xy + x^2y' \qquad y''(0) = 0$$

$$y''' = 2y + 4xy' + x^2y'' \qquad y'''(0) = 2$$

$$y^{iv} = 6y' + 6xy'' + x^2y''' \qquad y^{iv}(0) = -6$$

Putting these in Taylor series

$$y(x) = y(0) + (x-0)y'(0) + \frac{(x-0)^2}{2!}y''(0) + \dots$$

$$= 1 + x(-1) + \frac{x^2}{2!}(0) + \frac{x^3}{3!}(2) + \frac{x^4}{4!}(-6) + \dots$$

Now $y(x) = 1 - x + \frac{2x^3}{3!} - \frac{6x^4}{4!} - \dots$

$$y(0.1) = 0.90033$$

$$y(0.2) = 0.80227.$$

Milne's Predictor Corrector Method:

Let $\frac{dy}{dx} = f(x,y)$

$y(x_0) = y_0$ is given.

we first compute

$y_1 = y(x_0+h)$

$y_2 = y(x_0+2h)$

& $y_3 = y(x_0+3h)$ by Picard's method.

Next we calculate

$f_0 = f(x_0, y_0)$

$f_1 = f(x_0+h, y_1)$

$f_2 = f(x_0+2h, y_2)$

$f_3 = f(x_0+3h, y_3)$

(This we will do by putting $x = x_0, x_0+h, x_0+2h$ & x_0+3h in place of x & $y = y_0, y_1, y_2, y_3$ in place of y, respectively in $f(x,y)$).

Then we find $y_4 = y(x_0+4h)$ by using

$$y_4 = y_0 + \frac{4h}{3}(2f_1 - f_2 + 2f_3)$$

which is called predictor.

Next we obtain

$$f_4 = f(x_0+4h, y_4)$$

Then a better value of y_4 is found by using

$$y_4 = y_2 + \frac{h}{3}(f_2 + 4f_3 + f_4)$$

which is called corrector.

Then an improved value of f_4 is computed & again the corrector is applied to find a still better value of y_4. We repeat this step until y_4 remains unchanged.

Once y_4 & f_4 are obtained to desired degree of accuracy y_5 is found from predictor as

$$y_5 = y_1 + \frac{4h}{3}(2f_2 - 4f_3 + 2f_4).$$

& $f_5 = f(x_0 + 5h, y_5)$ is calculated.

Then a better approximation of the values of y_5 is obtained from corrector as

$$y_5 = y_3 + \frac{h}{3}(f_3 + 4f_4 + f_5)$$

We repeat this step till y_5 becomes stationary & then we proceed to calculate y_6 as before.

e.g: Apply Milne's predictor corrector method to find a solution of the differential equation

$$\frac{dy}{dx} = x - y^2 \text{ in the range } 0 \le x \le 1 \text{ for the boundary}$$

condition $y = 0$ at $x = 0$. $h = 0.2$

Solu: Using Picard's method

$$y = y(0) + \int_0^x f(x,y)\,dx \quad \text{when } f(x,y) = x - y^2$$

To get the first approx, we put $y = 0$ in $f(x,y)$

$$y_1 = 0 + \int_0^x x\,dx = \frac{x^2}{2}.$$

For second approx, we put $y = y_1$ in $f(x,y)$

$$y_2 = 0 + \int_0^x \left[x - \left(\frac{x^2}{2}\right)^2\right] dx = \frac{x^2}{2} - \frac{x^5}{20}$$

For third approx, we put $y = y_2$ in $f(x,y)$

$$y_3 = 0 + \int_0^x \left[x - \left(\frac{x^2}{2} - \frac{x^5}{20}\right)^2\right] dx$$

$$= \frac{x^2}{2} - \frac{x^5}{20} + \frac{x^8}{160} - \frac{x^{11}}{4400} \quad \text{— (1)}$$

Now let us determine the starting values of the Milne's method from (1) by choosing $h = 0.2$.

$x_0 = 0.0$ $\quad y_0 = 0.0$ $\quad f_0 = 0.0000$

$x_1 = 0.2$ $\quad y_1 = 0.020$ $\quad f_1 = 0.1996$

$x_2 = 0.4 \qquad y_2 = 0.0795 \qquad f_2 = 0.3937$

$x_3 = 0.6 \qquad y_3 = 0.1762 \qquad f_3 = 0.5689$

Using the predictor,

$$y_4 = y_0 + \frac{4h}{3}(2f_1 - f_2 + 2f_3)$$

$$= 0.3049$$

$\therefore$

$x_4 = 0.8 \qquad y_4 = 0.3049 \qquad f_4 = 0.7070.$

And the corrector

$$y_4 = y_2 + \frac{h}{3}[f_2 + 4f_3 + f_4] \text{ yields}$$

$$y_4 = 0.3046, \quad f_4 = 0.7072.$$

Again using corrector

$$y_4 = 0.3046$$

Since value of y_4 is stationary $\therefore$ we now jump to y_5.

Now using the predictor

$$y_5 = y_1 + \frac{4h}{3}(2f_2 - f_3 + 2f_4) = 0.4554$$

$x = 1.0 \qquad y_5 = 0.4554 \qquad f_5 = 0.7926$

and the corrector

$$y_5 = y_3 + \frac{h}{3}(f_3 + 4f_4 + f_5), \text{ gives}$$

$$y_5 = 0.4555 \qquad f_5 = 0.7925$$

Again using corrector

$y_5 = 0.4555$, a value which is the same as last one

$\therefore$

$y(1) = 0.4555$: (Ans)

Adam Bashforth method —

Given $\frac{dy}{dx} = f(x,y)$ & $y_0 = y(x_0)$, we compute

$y_{-1} = y(x_0 - h)$

$y_{-2} = y(x_0 - 2h)$

$y_{-3} = y(x_0 - 3h)$

by Taylor's series or Euler's or Runge kutta method.

Next we calculate

$$f_{-1} = f(x_0 - h, y_{-1})$$

$$f_{-2} = f(x_0 - 2h, y_{-2})$$

$$f_{-3} = f(x_0 - 3h, y_{-3})$$

Then to find y_1, we substitute Newton's backward Interpolation formula;

$$f(x,y) = f_0 + n\nabla f_0 + \frac{n(n+1)}{2!}\nabla^2 f_0 + \frac{(n)(n+1)(n+2)}{3!}\nabla^3 f_0 + \dots$$

in $$y_1 = y_0 + \int_{x_0}^{x_0+h} f(x,y)\,dx \qquad (1)$$

$$y_1 = y_0 + \int_{x_0}^{x_1} \left[f_0 + n\nabla f_0 + \frac{n(n+1)}{2!}\nabla^2 f_0 + \dots\right] dx$$

Put $x = x_0 + nh$

$dx = h\,dn$

$$= y_0 + h\int_0^1 \left[f_0 + n\nabla f_0 + \frac{n(n+1)}{2!}\nabla^2 f_0 - \dots\right] dn$$

$$= y_0 + h\left[f_0 + \frac{1}{2}\nabla f_0 + \frac{5}{12}\nabla^2 f_0 + \frac{3}{8}\nabla^3 f_0 \dots\right]$$

Neglecting fourth & higher order differences & expressing ∇f_0, $\nabla^2 f_0$ & $\nabla^3 f_0$ in terms of function values, we get

$$y_1 = y_0 + \frac{h}{24}(55 f_0 - 59 f_{-1} + 37 f_{-2} - 9 f_{-3}) \quad (2)$$

This is called Adams-Bashforth predictor formula.

Now we find

$$f_1 = f(x_0 + h, y_1)$$

Then to find a better value of y_1, we derive a corrector formula by substituting Newton's Backward formula at f_1 i.e

$$f(x, y) = f_1 + n\nabla f_1 + \frac{n(n+1)}{2!}\nabla^2 f_1 + \frac{n(n+1)(n+2)}{3!}\nabla^3 f_1 \ldots$$

in (1)

$$y_1 = y_0 + \int_{x_0}^{x_1} \left(f_1 + n\nabla f_1 + \frac{n(n+1)}{2!}\nabla^2 f_1 \ldots\right) dx$$

Put $x = x_1 + nh$

$dx = h\,dn$

$$= y_0 + \int_{-1}^{0} \left(f_1 + n\nabla f_1 + \frac{n(n+1)}{2!}\nabla^2 f_1 \ldots\right) dn$$

$$= y_0 + h\left[f_1 - \frac{1}{2}\nabla f_1 - \frac{1}{12}\nabla^2 f_1 - \frac{1}{24}\nabla^3 f_1 - \ldots\right]$$

Neglecting fourth & higher order differences & expressing ∇f_1, $\nabla^2 f_1$ & $\nabla^3 f_1$ in terms of function values, we obtain

$$y_1 = y_0 + \frac{h}{24}\left[9f_1 + 19f_0 - 5f_{-1} + f_{-2}\right] \quad - (3)$$

which is called Adam-Moulton corrector formula.

Then an improved value of f_1 is calculated & again the corrector (3) is applied to find a still better value of y_1. This step is repeated till y_1 remains unchanged & then we proceed to calculate y_2 as above.

eg: Solve the initial value problem $\frac{dy}{dx} = x - y^2$, $y(0) = 1$ to find $y(0.4)$ by Adam's method. Starting solutions required are to be obtained using Runge-Kutta method of order 4 using step value $h = 0.1$.

Solu: We have $f(x,y) = x - y^2$

To find $y(0.1)$:

$x_0 = 0 \quad y_0 = 1 \quad h = 0.1$

$K_1 = hf(x_0 y_0) = (0.1) f(0,1) = -0.1$

$K_2 = hf(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}K_1) = (0.1) f(0.05, 0.95) = -0.08525$

$K_3 = hf(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}K_2) = (0.1) f(0.05, 0.9574) = -0.08666$

$K_3 = hf(x_0 + h, y_0 + K_3) = (0.1) f(0.1, 0.9137) = -0.07348$

$K = \frac{1}{6}\left[K_1 + 2K_2 + 2K_3 + K_4\right] = = -0.08621$

$y(0.1) = y_1 = y_0 + K = 1 - 0.08621$
$= 0.9137$

To find $y(0.2)$

Here $x_1 = 0.1$ $y_1 = 0.9137$, $h = 0.1$

$$K_1 = hf(x_1, y_1) = (0.1)f(0.1, 0.9137)$$

$$K_2 = hf(x_1 + \tfrac{1}{2}h, y_1 + \tfrac{1}{2}K_1) = (0.1)f(0.15, 0.8751)$$

$$K_3 = hf(x_1 + \tfrac{1}{2}h, y_1 + \tfrac{1}{2}K_2) = (0.1)f(0.15, 0.8809)$$

$$K_4 = hf(x_1 + h, y_1 + K_3) = (0.1)f(0.2, 0.8491)$$

$$K = \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

$$y(0.2) = y_2 = y_1 + K = 0.8494.$$

To find $y(0.3)$

$x_2 = 0.2$ $y_2 = 0.8494$ $y = 0.1$

$$K_1 = hf(x_2, y_2) = (0.1)f(0.2, 0.8494)$$

$$K_2 = hf(x_2 + \tfrac{1}{2}h, y_2 + \tfrac{1}{2}K_1)$$

$$= (0.1)f(0.25, 0.8233)$$

$$K_3 = hf(x_2 + \tfrac{1}{2}h, y_2 + \tfrac{1}{2}K_2)$$

$$= (0.1)f(0.25, 0.828)$$

$$K_4 = hf(x_2 + h, y_2 + K_3)$$

$$= (0.1)f(0.3, 0.8058)$$

$$K = \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

$$y(0.3) = y_3 = y_2 + K = 0.8061$$

Now the starting values of Adam's methods are

$x = 0.0 \qquad y_{-3} = 1.000$

$x = 0.1 \qquad y_{-2} = 0.9137$

$x = 0.2 \qquad y_{-1} = 0.8494$

$x = 0.3 \qquad y_0 = 0.8061$

We find $f_{-3} = f(0.0, 1.0) = 0.0 - (1.0)^2 = -1.000$

$f_{-2} = f(0.1, 0.9137) = 0.1 - (0.9137)^2 = -0.7312$

$f_{-1} = f(0.2, 0.8494) = 0.2 - (0.8494)^2 = -0.5215$

$f_0 = f(0.3, 0.8061) = 0.3 - (0.8061)^2 = -0.3498$

Using the predictor formula

$$y_1 = y_0 + \frac{h}{24}\left(55 f_0 - 59 f_{-1} + 37 f_{-2} - 9 f_{-3}\right)$$

$x = 0.4$

$$y_1 = 0.8061 + \frac{0.1}{24}\Big[55 \times (-0.3498) - 59(-0.5215) + 37(-0.7312) - 9(-1.00)\Big]$$

$$= 0.7789$$

$$\therefore f(0.4, 0.7789) = f_1 = 0.4 - (0.7789)^2$$

$$= -0.2067$$

Using corrector

$$y_1 = y_0 + \frac{h}{24}\left(9 f_1 + 19 f_0 - 5 f_{-1} + f_{-2}\right)$$

$$= 0.8061 + \frac{0.1}{24}\Big[9 \times (-0.2067) + 19(-0.3498) - 5(-0.5215) + (-0.7312)\Big]$$

$$= 0.7785$$

$\therefore\ y(0.4) = 0.7785$

(correct upto 3 decimal places)

11

NUMERICAL SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS

1. Introduction
2. Classification of second order equations
3. Finite-difference approximations to partial derivatives
4. Elliptic equations
5. Solution of Laplace equation
6. Solution of Poisson's equation
7. Solution of elliptic equations by relaxation method
8. Parabolic equations
9. Solution of one-dimensional heat equation
10. Solution of two-dimensional heat equation
11. Hyperbolic equations
12. Solution of wave equation

11.1. INTRODUCTION

Partial differential equations arise in the study of many branches of applied mathematics *e.g.* in fluid dynamics, heat transfer, boundary layer flow, elasticity, quantum mechanics and electro-magnetic theory. Only a few of these equations can be solved by analytical methods which are also complicated requiring use of advanced mathematical techniques. In most of the cases, it is easier to develop approximate solutions by numerical methods. Of all the numerical methods available for the solution of partial differential equations, the method of finite differences is most commonly used. In this method, the derivatives appearing in the equation and the boundary conditions are replaced by their finite difference approximations. Then the given equation is changed to a system of linear equations which are solved by iterative procedures. This process is slow but produces good results in many boundary value problems. An added advantage of this method is that the computation can be carried by electronic computers. To accelerate the solution, sometimes the method of relaxation proves quite effective.

Besides discussing finite difference method, we shall briefly describe the relaxation method also in this chapter.

11.2. CLASSIFICATION OF SECOND ORDER EQUATIONS

The general linear partial differential equation of the second order in two independent variables is of the form

$$A(x, y)\frac{\partial^2 u}{dx^2} + B(x, y)\frac{\partial^2 u}{\partial x \partial y} + C(x, y)\frac{\partial^2 u}{\partial y^2} + F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0 \quad \text{...(1)}$$

Such a partial differential equation is said to be

(*i*) **elliptic** *if* $B^2 - 4AC < 0$, (*ii*) **parabolic** *if* $B^2 - 4AC = 0$,

and (*iii*) **hyperbolic** *if* $B^2 - 4AC > 0$.

Example 11.1. *Classify the following equations :*

(*i*) $\dfrac{\partial^2 u}{\partial x^2} + 4\dfrac{\partial^2 u}{\partial x \partial y} + 4\dfrac{\partial^2 u}{\partial y^2} - \dfrac{\partial u}{\partial x} + 2\dfrac{\partial u}{\partial y} = 0$

(*ii*) $x^2\dfrac{\partial^2 u}{\partial x^2} + (1 - y^2)\dfrac{\partial^2 u}{\partial y^2} = 0, -\infty < x\ \infty, -1 < y < 1$

(*iii*) $(1 + x^2)\dfrac{\partial^2 u}{\partial x} + (5 + 2x^2)\dfrac{\partial^2 u}{\partial x \partial t} + (4 + x^2)\dfrac{\partial^2 u}{\partial t} = 0.$

Sol. (*i*) Comparing this equation with (1) above, we find that

$$A = 1, B = 4, C = 4$$

$$\therefore \quad B^2 - 4AC = (4)^2 - 4 \times 1 \times 4 = 0$$

So the equation is parabolic.

(*ii*) Here $A = x^2, B = 0, C = 1 - y^2$

$$\therefore \quad B^2 - 4AC = 0 - 4x^2(1 - y^2) = 4x^2(y^2 - 1)$$

For all x between $-\infty$ and ∞, x^2 is positive

For all y between -1 and 1, $y^2 < 1$

$$B^2 - 4AC < 0$$

Hence the equation is elliptic.

(*iii*) Here $A = 1 + x^2, B = 5 + 2x^2, C = 4 + x^2$

$$\therefore \quad B^2 - 4AC = (5 + 2x^2)^2 - 4(1 + x^2)(4 + x^2) = 9 \text{ i.e. } > 0$$

So the equation is hyperbolic.

PROBLEMS 11.1

1. What is the classification of the equation $f_{xx} + 2f_{xy} + f_{yy} = 0$. (*Madras B.E., 2001*)
2. Determine whether the following equation is elliptic or hyperbolic ?
$$(x + 1)u_{xx} - 2(x + 2)u_{xy} + (x + 3)u_{yy} = 0.$$
3. Classify the equation (*i*) $y^2\dfrac{\partial^2 u}{\partial x^2} - 2y\dfrac{\partial^2 u}{\partial x \partial y} + \dfrac{\partial^2 u}{\partial y^2} - \dfrac{\partial u}{\partial y} = 8y.$

   (*ii*) $y^2 u_{xx} - 2xyu_{xy} + x^2 u_{yy} + 2u_x - 3u = 0.$ (*Madras, B.E., 2003*)
4. In which parts of the (x, y) plane is the following equation elliptic ?
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x \partial y} + (x^2 + 4y^2)\frac{\partial^2 u}{\partial y^2} = 2\sin(xy).$$

11.3. FINITE DIFFERENCE APPROXIMATIONS TO PARTIAL DERIVATIVES

Consider a rectangular region R in the x, y plane. Divide this region into a rectangular network of sides $\Delta x = h$ and $\Delta y = k$ as shown in Fig. 11.1. The points of intersection of the dividing lines are called *mesh points, nodal points or grid points.*

Fig. 11.1

Then we have the finite difference approximations for the partial derivatives in x-direction (§ 10.17) :

$$\frac{\partial u}{\partial x} = \frac{u(x+h, y) - u(x, y)}{h} + O(h) = \frac{u(x, y) - u(x-h, y)}{h} + O(h)$$

$$= \frac{u(x+h, y) - u(x-h, y)}{2h} + O(h^2)$$

and
$$\frac{\partial^2 u}{\partial x^2} = \frac{u(x-h, y) - 2u(x, y) + u(x+h, y)}{h^2} + O(h^2)$$

Writing $u(x, y) = u(ih, jk)$ as simply $u_{i,j}$, the above approximations become

$$u_x = \frac{u_{i+1,j} - u_{i,j}}{h} + O(h) \quad \text{...(1)}$$

$$= \frac{u_{i,j} - u_{i-1,j}}{h} + O(h) \quad \text{...(2)}$$

$$= \frac{u_{i+1,j} - u_{i-1,j}}{2h} + O(h^2) \quad \text{...(3)}$$

and
$$u_{xx} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} + O(h^2) \quad \text{...(4)}$$

Similarly we have the approximations for the derivatives w.r.t. y :

$$u_y = \frac{u_{i,j+1} - u_{i,j}}{k} + O(k) \quad \text{...(5)}$$

$$= \frac{u_{i,j} - u_{i,j-1}}{k} + O(k) \qquad ...(6)$$

$$= \frac{u_{i,j+1} - u_{i,j-1}}{2k} + O(k^2) \qquad ...(7)$$

and

$$u_{yy} = \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{k^2} + O(k^2) \qquad ...(8)$$

Replacing the derivatives in any partial differential equation by their corresponding difference approximations (1) to (8), we obtain the finite-difference analogues of the given equation.

11.4. ELLIPTIC EQUATIONS

The *Laplace equation* $\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$ and the *Poisson's equation* $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$ are examples of elliptic partial differential equations. Laplace equation arises in steady-state flow and potential problems. Poisson's equation arises in fluid mechanics, electricity and magnetism and torsion problems.

The solution of these equations is a function $u(x, y)$ which is satisfied at every point of a region R subject to certain boundary conditions specified on the closed curve C (Fig. 11.2).

Fig. 11.2

In general, problems concerning steady viscous flow, equilibrium stresses in elastic structures etc., lead to elliptic type of equations.

11.5. SOLUTION OF LAPLACE EQUATION

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \qquad ...(1)$$

Consider a rectangular region R for which $u(x, y)$ is known at the boundary. Divide this region into a network of square mesh of side h, as shown in Fig. 11.3 (assuming that an exact sub-division of R is possible). Replacing the derivatives in (1) by their difference approximations, we have

$$\frac{1}{h^2}[u_{i-1,j} - 2u_{i,j} + u_{i+1,j}] + \frac{1}{h^2}[u_{i,j-1} - 2u_{i,j} + u_{i,j+1}] = 0$$

or

$$u_{i,j} = \frac{1}{4}[u_{i-1,j} + u_{i+1,j} + u_{i,j+1} + u_{i,j-1}] \qquad ...(2)$$

Fig. 11.3

This shows that the value of u at any interior mesh point is the average of its values at four neighbouring points to the left, right, above and below. (2) is called the **standard 5-point formula** which is exhibited in Fig. 11.4.

Sometimes a formula similar to (2) is used which is given by

$$u_{i,j} = \frac{1}{4}(u_{i-1,j+1} + u_{i+1,j-1} + u_{i+1,j+1} + u_{i-1,j-1}) \qquad ...(3)$$

This shows that the value of $u_{i,j}$ is the average of its values at the four neighbouring diagonal mesh points. (3) is called the **diagonal 5-point formula** which is represented in Fig. 11.5. Although (3) is less accurate than (2), yet it serves as a reasonably good approximation for obtaining the starting values at the mesh points.

Fig. 11.4

Fig. 11.5

Now to find the initial values of u at the interior mesh points, we first use the diagonal five point formula (3) and compute $u_{3,3}$, $u_{2,4}$, $u_{4,4}$, $u_{4,2}$ and $u_{2,2}$, in this order. Thus we get,

$$u_{3,3} = \frac{1}{4}(b_{1,5} + b_{5,1} + b_{5,5} + b_{1,1})\,;\; u_{2,4} = \frac{1}{4}(b_{1,5} + u_{3,3} + b_{3,5} + b_{1,3})$$

$$u_{4,4} = \frac{1}{4}(b_{3,5} + b_{5,3} + b_{3,5} + u_{3,3})\ ;\ u_{4,2} = \frac{1}{4}(u_{3,3} + b_{5,1} + b_{3,1} + b_{5,3})$$

$$u_{2,2} = \frac{1}{4}(b_{1,3} + b_{3,1} + u_{3,3} + b_{1,1})$$

The values at the remaining interior points *i.e.* $u_{2,3}$, $u_{3,4}$, $u_{4,3}$ and $u_{3,2}$ are computed by the standard five-point formula (2). Thus, we obtain

$$u_{2,3} = \frac{1}{4}(b_{1,3} + u_{3,3} + u_{2,4} + u_{2,2}),\ u_{3,4} = \frac{1}{4}(u_{2,4} + u_{4,4} + b_{3,5} + u_{3,3})$$

$$u_{4,3} = \frac{1}{4}(u_{3,3} + b_{5,3} + u_{4,4} + u_{4,2}),\ u_{3,2} = \frac{1}{4}(u_{2,2} + u_{4,2} + u_{3,3} + u_{3,1})$$

Having found all the nine values of $u_{i,j}$ once, their accuracy is improved by either of the following iterative methods. In each case, the method is repeated till the difference between two consecutive iterates becomes negligible.

(*i*) **Jacobi's method.** Denoting the nth iterative value of $u_{i,j}$, by $u^{(n)}_{i,j}$, the iterative formula to solve (2) is

$$u^{(n+1)}_{i,j} = \frac{1}{4}[u^{(n)}_{i-1,j} + u^{(n)}_{i+1,j} + u^{(n)}_{i,j+1} + u^{(n)}_{i,j-1}] \qquad ...(4)$$

It gives improved values of $u_{i,j}$ at the interior mesh points and is called the *point Jacobi's formula*.

(*ii*) **Gauss-Seidal method.** In this method, the iteration formula is

$$u^{(n+1)}_{i,j} = \frac{1}{4}[u^{(n+1)}_{i-1,j} + u^{(n)}_{i+1,j} + u^{(n+1)}_{i,j+1} + u^{(n)}_{i,j-1}]$$

It utilises the latest iterative value available and scans the mesh points symmetrically from left to right along successive rows.

Obs. Gauss-Seidal method is simple and can be adapted to computer calculations. Its convergence being slow, the working is somewhat lengthy. It can however, be shown that the Gauss-Seidal scheme converges twice as fast as Jacobi's scheme.

The accuracy of calculations depends on the mesh-size *i.e.* smaller the h, better the accuracy. But if h is too small, it may increase rounding-off errors and also increases the labour of computation.

Example 11.2. *Solve the elliptic equation $u_{xx} + u_{yy} = 0$ for the following square mesh with boundary values as shown in Fig. 11.6.* (*Rohtak, B.E., 2003*)

Sol. Let $u_1, u_2, \ldots\ldots, u_9$ be the values of u at the interior mesh-points. Since the boundary values of u are symmetrical about AB, $\therefore\ u_7 = u_1, u_8 = u_2, u_9 = u_3$.

Also the values of u being symmetrical about CD. $u_3 = u_1, u_6 = u_4, u_9 = u_7$.

Thus it is sufficient to find the values u_1, u_2, u_4 and u_5.

Now we find their initial values in the following order :

$$u_5 = \frac{1}{4}(2000 + 2000 + 1000 + 1000) = 1500 \qquad \text{(Std. formula)}$$

$$u_1 = \frac{1}{4}(0 + 1500 + 1000 + 2000) = 1125 \qquad \text{(Diag. formula)}$$

Fig. 11.6

$$u_2 = \frac{1}{4}(1125 + 1125 + 1000 + 1500) \approx 1188 \qquad \text{(Std. formula)}$$

$$u_4 = \frac{1}{4}(2000 + 1500 + 1125 + 1125) \approx 1438 \qquad \text{(Std. formula)}$$

We carry out the iteration process using the formulae :

$$u_1^{(n+1)} = \frac{1}{4}[1000 + u_2^{(n)} + 500 + u_4^{(n)}]$$

$$u_2^{(n+1)} = \frac{1}{4}[u_1^{(n+1)} + u_1^{(n)} + 1000 + u_5^{(n)}]$$

$$u_4^{(n+1)} = \frac{1}{4}[2000 + u_5^{(n)} + u_1^{(n+1)} + u_1^{(n)}]$$

$$u_5^{(n+1)} = \frac{1}{4}[u_4^{(n+1)} + u_4^{(n)} + u_2^{(n+1)} + u_2^{(n)}]$$

First iteration : (put $n = 0$)

$$u_1^{(1)} = \frac{1}{4}(1000 + 1188 + 500 + 1438) \approx 1032$$

$$u_2^{(1)} = \frac{1}{4}(1032 + 1125 + 1000 + 1500) = 1164$$

$$u_4^{(1)} = \frac{1}{4}(2000 + 1500 + 1032 + 1125) = 1414$$

$$u_5^{(1)} = \frac{1}{4}(1414 + 1438 + 1164 + 1188) = 1301$$

Second iteration : (put $n = 1$)

$$u_1^{(2)} = \frac{1}{4}(1000 + 1164 + 500 + 1414) = 1020$$

$$u_2^{(2)} = \frac{1}{4}(1020 + 1032 + 1000 + 1301) = 1088$$

$$u_4^{(2)} = \frac{1}{4}(2000 + 1301 + 1020 + 1032) = 1338$$

$$u_5^{(2)} = \frac{1}{4}(1338 + 1414 + 1088 + 1164) = 1251$$

Third iteration :

$$u_1^{(3)} = \frac{1}{4}(1000 + 1088 + 500 + 1338) = 982$$

$$u_2^{(3)} = \frac{1}{4}(982 + 1020 + 1000 + 1251) = 1063$$

$$u_4^{(3)} = \frac{1}{4}(2000 + 1251 + 982 + 1020) = 1313$$

$$u_5^{(3)} = \frac{1}{4}(1313 + 1338 + 1063 + 1088) = 1201$$

Fourth iteration :

$$u_1^{(4)} = \frac{1}{4}(1000 + 1063 + 500 + 1313) = 969$$

$$u_2^{(4)} = \frac{1}{4}(969 + 982 + 1000 + 1201) = 1038$$

$$u_4^{(4)} = \frac{1}{4}(2000 + 1201 + 969 + 982) = 1288$$

$$u_5^{(4)} = \frac{1}{4}(1288 + 1313 + 1038 + 1063) = 1176$$

Fifth iteration :

$$u_1^{(5)} = \frac{1}{4}(1000 + 1038 + 500 + 1288) = 957$$

$$u_2^{(5)} = \frac{1}{4}(957 + 969 + 1000 + 1176) \approx 1026$$

$$u_4^{(5)} = \frac{1}{4}(2000 + 1176 + 957 + 969) \approx 1276$$

$$u_5^{(5)} = \frac{1}{4}(1276 + 1288 + 1026 + 1038) = 1157$$

Similarly,

$$u_1^{(6)} = 951,\ u_2^{(6)} = 1016,\ u_4^{(6)} = 1266,\ u_5^{(6)} = 1146$$
$$u_1^{(7)} = 946,\ u_2^{(7)} = 1011,\ u_4^{(7)} = 1260,\ u_5^{(7)} = 1138$$
$$u_1^{(8)} = 943,\ u_2^{(8)} = 1007,\ u_4^{(8)} = 1257,\ u_5^{(8)} = 1134$$
$$u_1^{(9)} = 941,\ u_2^{(9)} = 1005,\ u_4^{(9)} = 1255,\ u_5^{(9)} = 1131$$
$$u_1^{(10)} = 940,\ u_2^{(10)} = 1003,\ u_4^{(10)} = 1253,\ u_5^{(10)} = 1129$$
$$u_1^{(11)} = 939,\ u_2^{(11)} = 1002,\ u_4^{(11)} = 1252,\ u_5^{(11)} = 1128$$
$$u_1^{(12)} \approx 939,\ u_2^{(12)} \approx 1001,\ u_4^{(12)} \approx 1251,\ u_5^{(12)} = 1126$$

Thus there is negligible difference between the values obtained in the 11th and 12th iterations.

Hence $u_1 = 939$, $u_2 = 1001$, $u_4 = 1251$ and $u_5 = 1126$.

Example 11.3. *Given the values of $u(x, y)$ on the boundary of the square in the Fig. 11.7, evaluate the function $u(x, y)$ satisfying the Laplace equation $\nabla^2 u = 0$ at the pivotal points of this figure by*

(a) Jacobi's method.

(b) Gauss-Seidal method. *(Madras, B.E., 2003)*

Sol. To get the initial values of u_1, u_2, u_3, u_4, we assume that $u_4 = 0$. Then

$$u_1 = \frac{1}{4}(1000 + 0 + 1000 + 2000) = 1000 \qquad \text{(Diag. formula)}$$

$$u_2 = \frac{1}{4}(1000 + 500 + 1000 + 0) = 625 \qquad \text{(Std. formula)}$$

$$u_3 = \frac{1}{4}(2000 + 0 + 1000 + 500) = 875 \qquad \text{(Std. formula)}$$

$$u_4 = \frac{1}{4}(875 + 0 + 625 + 0) = 375 \qquad \text{(Std. formula)}$$

Fig. 11.7

(*a*) We carry out the successive iterations, using Jacobi's formulae :

$$u_1^{(n+1)} = \frac{1}{4}[2000 + u_2^{(n)} + 1000 + u_3^{(n)}]$$

$$u_2^{(n+1)} = \frac{1}{4}[u_1^{(n)} + 500 + 1000 + u_4^{(n)}]$$

$$u_3^{(n+1)} = \frac{1}{4}[2000 + u_4^{(n)} + u_1^{(n)} + 500]$$

$$u_4^{(n+1)} = \frac{1}{4}[u_3^{(n)} + 0 + u_2^{(n)} + 0]$$

First iteration : (put $n = 0$)

$$u_1^{(1)} = \frac{1}{4}(2000 + 625 + 1000 + 875) = 1125$$

$$u_2^{(1)} = \frac{1}{4}(1000 + 500 + 1000 + 375) \approx 719$$

$$u_3^{(1)} = \frac{1}{4}(2000 + 375 + 1000 + 500) \approx 969$$

$$u_4^{(1)} = \frac{1}{4}(875 + 0 + 625 + 0) \approx 375$$

Second iteration : (put $n = 1$)

$$u_1^{(2)} = \frac{1}{4}(2000 + 719 + 1000 + 969) = 1172$$

$$u_2^{(2)} = \frac{1}{4}(1125 + 500 + 1000 + 375) = 750$$

$$u_3^{(2)} = \frac{1}{4}(2000 + 375 + 1125 + 500) = 1000$$

$$u_4^{(2)} = \frac{1}{4}(969 + 0 + 719 + 0) = 422$$

Similarly, $u_1^{(3)} \approx 1188, u_2^{(3)} \approx 774, u_3^{(3)} \approx 1024, u_4^{(3)} \approx 438$

$u_1^{(4)} \approx 1200, u_2^{(4)} \approx 782, u_3^{(4)} \approx 1032, u_4^{(4)} \approx 450$

$u_1^{(5)} \approx 1204, u_2^{(5)} \approx 788, u_3^{(5)} \approx 1038, u_4^{(5)} \approx 454$

$u_1^{(6)} \approx 1206.5, u_2^{(6)} \approx 790, u_3^{(6)} \approx 1040, u_4^{(6)} \approx 456.5$

$u_1^{(7)} \approx 1208, u_2^{(7)} \approx 791, u_3^{(7)} \approx 1041, u_4^{(7)} \approx 458$

and $u_1^{(8)} \approx 1208, u_2^{(8)} \approx 791.5, u_3^{(8)} \approx 1041.5, u_4^{(8)} \approx 458.$

Thus there is no significant difference between the seventh and eighth iteration values.

Hence $u_1 = 1208, u_2 = 792, u_3 = 1042$ and $u_4 = 458$.

(*b*) We carry out the successive iterations, using Gauss-Seidal formulae

$$u_1^{(n+1)} = \frac{1}{4}[2000 + u_2^{(n)} + 1000 + u_3^{(n)}]$$

$$u_2^{(n+1)} = \frac{1}{4}[u_1^{(n+1)} = 500 + 1000 + u_4^{(n)}]$$

$$u_3^{(n+1)} = \frac{1}{4}[2000 + u_4^{(n)} + u_1^{(n+1)} = 500]$$

$$u_4^{(n+1)} = \frac{1}{4}[u_3^{(n+1)} + 0 + u_2^{(n+1)} + 0]$$

First iteration : (put $n = 0$)

$$u_1^{(1)} = \frac{1}{4}(2000 + 625 + 1000 + 875) = 1125$$

$$u_2^{(1)} = \frac{1}{4}(1125 + 500 + 1000 + 375) = 750$$

$$u_3^{(1)} = \frac{1}{4}(2000 + 375 + 1125 + 500) = 1000$$

$$u_4^{(1)} = \frac{1}{4}(1000 + 0 + 750 + 0) \approx 438$$

Second iteration : (put $n = 1$)

$$u_1^{(2)} = \frac{1}{4}(2000 + 750 + 1000 + 1000) \approx 1188$$

$$u_2^{(2)} = \frac{1}{4}(1188 + 500 + 1000 + 438) \approx 782$$

$$u_3^{(2)} = \frac{1}{4}(2000 + 438 + 1188 + 500) \approx 1032$$

$$u_4^{(2)} = \frac{1}{4}(1032 + 0 + 782 + 0) \approx 454$$

Similarly $u_1^{(3)} \approx 1204,\ u_2^{(3)} \approx 789,\ u_3^{(3)} \approx 1040,\ u_4^{(3)} \approx 458$

$u_1^{(4)} \approx 1207,\ u_2^{(4)} \approx 791,\ u_3^{(4)} \approx 1041,\ u_4^{(4)} = 458$

and $u_1^{(5)} = 1208,\ u_2^{(5)} \approx 791.5,\ u_3^{(5)} \approx 1041.5,\ u_4^{(5)} \approx 458.25$

Thus there is no significant difference between the fourth and fifth iteration values.

Hence $u_1 = 1208,\ u_2 = 792,\ u_3 = 1042$ and $u_4 = 458$.

11.6. SOLUTION OF POISSON'S EQUATION

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y) \qquad ...(1)$$

Its method of solution is similar to that of the Laplace equation. Here the standard 5-point formula for (1) takes the form

$$u_{i-1,j} + u_{i+1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = h^2 f(ih, jh) \qquad ...(2)$$

By applying (2) at each interior mesh point, we arrive at linear equations in the nodal values $u_{i,j}$. These equations can be solved by Gauss-Seidal method.

Obs. The error in replacing u_{xx} by the finite difference approximation is of the order $O(h^2)$. Since $k = h$, the error in replacing u_{yy} by the difference approximation is also of the order $O(h^2)$. Hence the error in solving Laplace and Poisson's equations by finite difference method is of the order $O(h^2)$.

Example 11.4. *Solve the equation $\nabla^2 u = -10(x^2 + y^2 + 10)$ over the square with sides $x = 0 = y$, $x = 3 = y$ with $u = 0$ on the boundary and mesh length = 1.*

(Anna, B.E., 2004)

Fig. 11.8

Sol. Here $h = 1$.

$\therefore$ The standard 5-point formula for the given equation is

$$u_{i-1,j} + u_{i+1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = -10(i^2 + j^2 + 10) \qquad ...(i)$$

For u_1 $(i = 1, j = 2)$, (i) gives $0 + u_2 + 0 + u_3 - 4u_1 = -10(1 + 4 + 10)$

i.e, $$u_1 = \frac{1}{4}(u_2 + u_3 + 150) \qquad ...(ii)$$

For u_2 $(i = 2, j = 2)$, (i) gives $\quad u_2 = \frac{1}{4}(u_1 + u_4 + 180)$ ...(*iii*)

For $u_3 (i = 1, j = 1)$, we have $\quad u_3 = \frac{1}{4}(u_1 + u_4 + 120)$ ...(*iv*)

For u_4 $(i = 2, j = 1)$, we have $\quad u_4 = \frac{1}{4}(u_2 + u_3 + 150)$ ...(*v*)

Equations (*ii*) and (*v*) show that $u_4 = u_1$. Thus the above equations reduce to

$$u_1 = \frac{1}{4}(u_2 + u_3 + 150),\ u_2 = \frac{1}{4}(u_2 + 90),\ u_3 = \frac{1}{4}(u_1 + 60)$$

Now let us solve these equations by Gauss-Seidal iteration method.

First iteration : Starting from the approximations $u_2 = 0$, $u_3 = 0$, we obtain $u_1^{(1)} = 37.5$

Then $\quad u_2^{(1)} = \frac{1}{2}(37.5 + 90) \approx 64$

$$u_2^{(1)} = \frac{1}{2}(37.5 + 60) \approx 49$$

Second iteration : $\quad u_1^{(2)} = \frac{1}{4}(64 + 49 + 150) \approx 66,\ u_2^{(2)} = \frac{1}{2}(66 + 90) = 78$

$$u_3^{(2)} = \frac{1}{2}(66 + 60) = 63$$

Third iteration : $\quad u_1^{(3)} = \frac{1}{4}(78 + 63 + 150) \approx 73,\ u_2^{(3)} = \frac{1}{2}(73 + 90) \approx 82$

$$u_3^{(3)} = \frac{1}{2}(73 + 60) \approx 67$$

Fourth iteration : $\quad u_1^{(4)} = \frac{1}{2}(82 + 67 + 150) \approx 75,\ u_2^{(4)} = \frac{1}{2}(75 + 90) = 82.5$

$$u_3^{(4)} = \frac{1}{2}(75 + 60) = 67.5$$

Fifth iteration : $\quad u_1^{(5)} = \frac{1}{4}(82.5 + 67.5 + 150) = 75,\ u_2^{(5)} = \frac{1}{2}(75 + 90) = 82.5$

$$u_3^{(5)} = \frac{1}{2}(75 + 90) = 67.5$$

Since these values are the same as those of fourth iteration, we have $u_1 = 75$, $u_2 = 82.5$, $u_3 = 67.5$ and $u_4 = 75$.

PROBLEMS 11.2

1. Solve the equation $u_{xx} + u_{yy} = 0$ for the square mesh with the boundary values as shown in Fig. 11.9. *(Delhi, B.E., 2002)*

6. Solve the Laplace's equation $u_{xx} + u_{yy} = 0$ in the domain of Fig. 11.14 by (*a*) Jacobi's metho (*b*) Gauss-Seidal method.
7. Solve the Laplace's equation $\nabla^2 u = 0$ in the domain of the Fig. 11.15.
8. Solve the Poisson's equation $\nabla^2 u = 8x^2y^2$ for the square mesh of Fig. 11.16 with $u(x, y) =$ on the boundary and mesh length = 1.

Fig. 11.15

Fig. 11.16

11.7. (1) SOLUTION OF ELLIPTIC EQUATIONS BY RELAXATION METHOD

If the equations for all the mesh points are written using (2) of § 11.6, we get a system of equations which can be solved by any method. For this purpose, the method of relaxation i particularly well-suited. Here we shall describe this method in relation to elliptic equations

Consider the Laplace equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \qquad \text{...(1}$$

We take a square region and divide it into a square net of mesh size h. Let the value o u at A be u_0 and its values at the four adjacent points be u_1, u_2, u_3, u_4 (Fig. 11.17). Then

$$\frac{\partial^2 u}{\partial x^2} \approx \frac{u_1 + u_3 - 2u_0}{h^2} \quad \text{and} \quad \frac{\partial^2 u}{\partial y^2} \approx \frac{u_2 + u_4 - 2u_0}{h^2}$$

If (1) is satisfied at A, then

$$\frac{u_1 + u_3 - 2u_0}{h^2} + \frac{u_2 + u_4 - 2u_0}{h^2} \approx 0$$

or

$$u_1 + u_2 + u_3 + u_4 - 4u_0 \approx 0$$

If r_0 be the residual (discrepency) at the mesh point A, then

$$r_0 = u_1 + u_2 + u_3 + u_4 - 4u_0 \qquad \text{...(2)}$$

Similarly the residual at the point B, is given by

$$r_1 = u_0 + u_5 + u_6 + u_7 - 4u_1 \text{ and so on} \qquad \text{...(3)}$$

Fig. 11.17

The main aim of the relaxation process is to reduce all the residuals to zero by making them as small as possible step by step. We, therefore, try to adjust the value of u at an internal mesh point so as to make the residual there at zero. But when the value of u is changing at a mesh point, the values of the residuals at the neighbouring interior points will also be changed. If u_0 is given an increment 1, then

(*i*) (2) shows that r_0 is changed by -4.

(*ii*) (3) shows that r_1 is changed by 1.

i.e. if the value of the function is increased by 1 at a mesh point (shown by a double ring), *then the residual at that point is decreased by 4 while the residuals at the adjacent interior points* (shown by a single ring), *get increased each by 1.* This relaxation pattern is shown in Fig. 11.18.

Fig. 11.18

(2) Working procedure *to solve an equation by relaxation method :*

I. *Write down by trial, the initial values of u at the interior mesh points by diagonal-averaging or cross-averaging.*

II. *Calculate the residuals at each of these points by (2) above.* If we apply this formula at a point near the boundary, one or more end points get chopped off since there are no residuals at the boundary.

III. *Write the residuals at a mesh-point on the right of this point and the value of u on its left.*

IV. *Obtain the solution by reducing the residuals to zero, one by one, by giving suitable increments to u and using Fig. 11.18. At each step, we reduce the numerically largest residual to zero and record the increment of u on the left* (below the earlier value thereat) *and the modified residual on the right* (below the earlier residual).

V. *When a round of relaxation is completed, the value of u and its increments are added at each point. Using these values, calculate all the residuals afresh. If some of the recalculated residuals are large, liquidate these again.*

VI. *Stop the relaxation process, when the current values of the residuals are quite small. The solution will be the current value of u at each of the nodes.*

Obs. Relaxation method combines simplicity with the speed of convergence. *Its only drawback is its unsuitability for computer calculations.*

Example 11.5. *Solve by relaxation method, the Laplace equation* $\dfrac{\partial^2 u}{\partial x^2}+\dfrac{\partial^2 u}{\partial y^2}=0$, *inside the square bounded by the lines* $x = 0, x = 4, y = 0, y = 4$, *given that* $u = x^2y^2$ *on the boundary.*

Sol. Taking $h = 1$, we find u on the boundary from $u = x^2y^2$. The initial values of u at the nine mesh-points are estimated to be 24, 56, 104 ; 16, 32, 56 ; 8, 16, 24 as shown on the left of the points in Fig. 11.19.

$\therefore$ Residual at A *i.e.* $r_A = 0 + 56 + 16 + 16 - 4 \times 24 = -8$

Similarly $r_B = 0, r_C = -16, r_D = 0, r_E = 16, r_F = 0, r_G = 0, r_H = 0, r_I = -8$.

(*i*) The numerically largest residual is 16 at E. To liquidate it, we increase u by 4 so that the residual becomes zero and the residuals at neighbouring nodes get increased by 4.

Fig. 11.19

(*ii*) Next, the numerically largest residual is – 16 at *C*. To reduce it to zero, we increase *u* by – 4 so that the residuals at the adjacent nodes are increased by – 4.

(*iii*) Now, the numerically largest residual is – 8 at *A*. To liquidate it, we increase *u* by – 2 so that the residuals at the adjacent nodes are increased by – 2.

(*iv*) Finally, the largest residual is – 8 at *I*. To liquidate it, we increase *u* by – 2 so that the residuals at the adjacent points are increased by – 2.

(*v*) The numerically largest current residual being 2, we stop the relaxation process. Hence the final values of *u* are

$$u_A = 22, \quad u_B = 56, \quad u_C = 100;$$
$$u_D = 16, \quad u_E = 36, \quad u_F = 56;$$
$$u_G = 8, \quad u_H = 16, \quad u_I = 22.$$

Example 11.6. *Solve by relaxation method Example 11.3.*

Sol. (*i*) The initial values of *u* at *A*, *B*, *C* and *D* are estimated to be 1000, 625, 875 and 375 [Fig. 11.20 (*i*)].

$\therefore \quad r_A = 500, r_B = 375, r_C = 375, r_D = 0$

To liquidate r_A, increase *u* by 125

To liquidate r_B, increase *u* by 94

To liquidate r_C, increase *u* by 94

(*ii*) Modified values of *u* are 1125, 719, 969, 375 [Fig. (*ii*)]

$\therefore \quad r_A = 188, r_B = 124, r_C = 124, r_D = 188.$

To liquidate r_A, r_D, r_B, r_C increase *u* by 47, 47, 31, 31 in turn.

Fig. 11.20 (*i*)

(*ii*)

(*iii*)

(*iii*) Revised values of u are 1172, 750, 1000, 422 [Fig. (*iii*)]

$\therefore \qquad r_A = 62, r_B = 84, r_C = 84, r_D = 62$

To liquidate r_B, r_C, r_A, r_D increase u by 21, 21, 15, 15 respectively.

(*iv*) Improved values of u are 1187, 771, 1021, 437 [Fig. (*iv*)]

$\therefore \qquad r_A = 44, r_B = 40, r_C = 40, r_D = 44.$

To liquidate r_A, r_D, r_B, r_C increase u by 11, 11, 10, 10 respectively.

(*iv*)

(*v*)

(*v*) Modified values of u are 1198, 781, 1031, 448 [Fig. (*v*)]

$\therefore \qquad r_A = 20, r_B = 22, r_C = 22, r_D = 20.$

To liquidate r_B, r_C, r_A, r_D increase u by 5, 5, 5, 5 respectively.

(*vi*) Revised values of u are 1203, 786, 1036, 453 [Fig. (*vi*)]

$\therefore \qquad r_A = 10, r_B = 12, r_C = 12, r_D = 10$

To liquidate r_B, r_C, r_A, r_D increase u by 3, 3, 2, 2 respectively.

(vi)

(vii)

(*vii*) Improved values of u are 1205, 789, 1039, 455 [Fig. (*vii*)]

$\therefore$ $r_A = 8, r_B = 4, r_C = 4, r_D = 8.$

To liquidate r_A, r_D, r_B, r_C increase u by 2, 2, 1, 1.

(*viii*) Finally the current residuals being 1, 0, 0, 1, we stop the relaxation process.

Hence the values of u at A, B, C, D are 1207, 790, 1040, 457.

PROBLEMS 11.3

1. Given that $u(x, y)$ satisfies the equation $\nabla^2 u = 0$ and the boundary conditions are $u(0,$ $= 0, u(4, y) = 8 + 2y, u(x, 0) = \frac{1}{2}x^2, u(x, 4) = x^2$, find the values $u(i, j), i = 1, 2, 3; j = 1, 2, 3$ relaxation method.
2. Apply relaxation method to solve the equation $\nabla^2 u = -400$, when the region of u is square bounded by $x = 0, y = 0, x = 4$ and $y = 4$ and u is zero on the boundary of the sq
3. Solve by relaxation method, the equation $\nabla^2 u = 0$ in the square region with square m (Fig. 11.21) starting with the initial values $u_1 = u_2 = u_3 = u_4 = 1$.

Fig. 11.21

11.8. PARABOLIC EQUATIONS

The one-dimensional heat conduction equation $\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$ is a well-known example of parabolic partial differential equations. The solution of this equation is a temperature function $u(x, t)$ which is defined for values of x from 0 to l and for values of time t from 0 to ∞. The solution is not defined in a closed domain but advances in an open-ended region from initial values, satisfying the prescribed boundary conditions (Fig. 11.22).

Fig. 11.22

In general, the study of pressure waves in a fluid, propagation of heat and unsteady state problems lead to parabolic type of equations.

11.9. SOLUTION OF ONE DIMENSIONAL HEAT EQUATION

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2} \qquad ...(1)$$

where $c^2 = k/s\rho$ is the diffusivity of the substance (cm²/sec.)

(1) Schmidt method. Consider a rectangular mesh in the x-t plane with spacing h along x direction and k along time t direction. Denoting a mesh point $(x, t) = (ih, jk)$ as simply i, j, we have

$$\frac{\partial u}{\partial t} = \frac{u_{i,j+1} - u_{i,j}}{k} \qquad \text{[by (5) § 11.3}$$

and

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} \qquad \text{[by (4) § 11.3}$$

Substituting these in (1), we obtain $u_{i,j+1} - u_{i,j} = \frac{kc^2}{h^2}[u_{i-1,j} - 2u_{i,j} + u_{i+1,j}]$

or

$$u_{i,j+1} = \alpha u_{i-1,j} + (1 - 2\alpha)\, u_{i,j} + \alpha u_{i+1,j} \qquad ...(2)$$

where $\alpha = kc^2/h^2$ is the mesh ratio parameter.

This formula enables us to determine the value of u at the $(i, j+1)$th mesh point in terms of the known function values at the points x_{i-1}, x_i and x_{i+1} at the instant t_j. It is a relation between the function values at the two time levels $j+1$ and j and is therefore, called a *2-level formula*. In schematic form (2) is shown in Fig. 11.23.

Fig. 11.23

Hence (2) is called the *Schmidt explicit formula* which is valid only for $0 < \alpha \le \frac{1}{2}$.

Obs. In particular when $\alpha = \frac{1}{2}$, (2) reduces to

$$u_{i,j+1} = \frac{1}{2}(u_{i-1,j} + u_{i+1,j}) \qquad \text{...(3)}$$

which shows that *the value of u at x_i at time t_{j+1} is the mean of the u-values at x_{i-1} and x_{i+1} at time t_j*. This relation, known as *Bendre-Schmidt recurrence relation*, gives the values of u at the internal mesh points with the help of boundary conditions.

(2) Crank-Nicolson method. We have seen that the Schmidt scheme is computationally simple and for convergent results $\alpha \le \frac{1}{2}$ *i.e.* $k \le h^2/2c^2$. To obtain more accurate results, h should be small *i.e.* k is necessarily very small. This makes the computations exceptionally lengthy as more time-levels would be required to cover the region. A method that does not restrict α and also reduces the volume of calculations was proposed by Crank and Nicolson in 1947.

According to this method, $\partial^2 u/\partial x^2$ is replaced by the average of its central-difference approximations on the jth and $(j+1)$th time rows. Thus (1) is reduced to

$$\frac{u_{i,j+1} - u_{i,j}}{h} = c^2 \cdot \frac{1}{2}\left\{\frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2}\right\} + \left\{\frac{u_{i-1,j+1} - 2u_{i,j+1} + u_{i+1,j+1}}{h^2}\right\}$$

or
$$-\alpha u_{i-1,j+1} + (2+2\alpha)u_{i,j+1} - \alpha u_{i+1,j+1} = \alpha u_{i-1,j} + (2-2\alpha)u_{i,j} + \alpha u_{i+1,j} \qquad \text{...(4)}$$

where $\alpha = kc^2/h^2$.

Clearly the left side of (4) contains three unknown values of u at the $(j+1)$th level while all the three values on the right are known values at the jth level. Thus (4) is a *2-level implicit relation* and is known as *Crank-Nicolson formula*. It is convergent for all finite values of α. Its computational model is given in Fig. 11.24.

Fig. 11.24

If there are n internal mesh points on each row, then the relation (4) gives n simultaneous equations for the n unknown values in terms of the known boundary values. These equations can be solved to obtain the values at these mesh points. Similarly, the values at the internal mesh points on all rows can be found. A method such as this in which the calculation of an unknown mesh value necessitates the solution of a set of simultaneous equations, is known as an *implicit scheme*.

(3) Iterative methods of solution *for an implicit scheme.*

From (4), we have

$$(1+\alpha)\, u_{i,j+1} = \tfrac{1}{2}\,\alpha(u_{i-1,j+1} + u_{i+1,j+1}) + u_{i,j} + \tfrac{1}{2}\,\alpha(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}) \qquad ...(5)$$

Here only $u_{i,j+1}$, $u_{i-1,j+1}$ and $u_{i+1,j+1}$ are unknown while all others are known since these were already computed in the jth step.

Writing $\qquad b_i = u_{i,j} + \dfrac{\alpha}{2}\,(u_{i-1,j} - 2u_{i,j} + u_{i+1,j})$

and dropping j's (5) becomes $u_i = \dfrac{\alpha}{2(1+\alpha)}\,(u_{i-1} + u_{i+1}) + \dfrac{b_i}{1+\alpha}$

This gives the iteration formula

$$u_i^{(n+1)} = \frac{\alpha}{2(1+\alpha)}\,\{u_{i-1}^{(n)} + u_{i+1}^{(n)}\} + \frac{b_i}{1+\alpha} \qquad ...(6)$$

which expresses the $(n+1)$th iterates in terms of the nth iterates only. This is known as the *Jacobi's iteration formula*.

As the latest value of u_{i-1} *i.e.* $u_{i-1}^{(n+1)}$ is already available, the convergence of the iteration formula (6) can be improved by replacing $u_{i-1}^{(n)}$ by $u_{i-1}^{(n+1)}$. Accordingly (6) may be written as

$$u_i^{(n+1)} = \frac{\alpha}{2(1+\alpha)}\,\{u_{i-1}^{(n+1)} + u_{i+1}^{(n)}\} + \frac{b_i}{1+\alpha} \qquad ...(7)$$

which is known as *Gauss-Seidal iteration formula*.

Obs. Gauss-Seidal iteration scheme is valid for all finite values of α and converges twice as fast as Jacobi's scheme.

(4) Du Fort and Frankel method. If we replace the derivatives in (1) by the central-difference approximations,

$$\frac{\partial u}{\partial t} = \frac{u_{i,\,j+1} - u_{i,\,j-1}}{2k} \qquad \text{[From (7) § 11.3}$$

and
$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i-1,\,j} - 2u_{i,\,j-1} + u_{i+1,\,j}}{h^2} \qquad \text{[From (4) § 11.3}$$

we obtain
$$u_{i,\,j+1} - u_{i,\,j-1} = \frac{2kc^2}{h^2}\,[u_{i-1,\,j} - 2u_{i,\,j} + u_{i+1,\,j}]$$

i.e.
$$u_{i,\,j+1} = u_{i,\,j-1} + 2\alpha\,[u_{i-1,\,j} - 2u_{i,\,j} + u_{i+1,\,j}] \qquad \text{...(7)}$$

where $\alpha = kc^2/h^2$. This difference equation is called the *Richardson scheme* which is a *3-level method*.

If we replace $u_{i,j}$ by the mean of the values $u_{i,j-1}$ and $u_{i,j+1}$ *i.e.* $u_{i,j} = \frac{1}{2}(u_{i,j-1} + u_{i,j+1})$ in (7), then we get

$$u_{i,\,j+1} = u_{i,\,j-1} + 2\alpha[u_{i-1,\,j} - (u_{i,\,j-1} + u_{i,\,j+1}) + u_{i+1,\,j}]$$

On simplification, it can be written as

$$u_{i,\,j+1} = \frac{1-2\alpha}{1+2\alpha}\,u_{i,\,j-1} + \frac{2\alpha}{1+2\alpha}\,\{u_{i-1,\,j} + u_{i+1,\,j}\} \qquad \text{...(8)}$$

This difference scheme is called *Du Fort-Frankel method* which is a *3-level explicit method*. Its computational model is given in Fig. 11.25.

Fig. 11.25

Example 11.7. *Find the values of u(x, t) satisfying the parabolic equation* $\frac{\partial u}{\partial t} = 4\,\frac{\partial^2 u}{\partial x^2}$ *and the boundary conditions* $u(0, t) = 0 = u(8, t)$ *and* $u(x, 0) = 4x - \frac{1}{2}x^2$ *at the points* $x = i$: $i = 0, 1, 2, \ldots\ldots, 7$ *and* $t = \frac{1}{8}j : j = 0, 1, 2, \ldots\ldots, 5$.

Sol. Here $c^2 = 4$, $h = 1$ and $k = 1/8$. Then $\alpha = c^2k/h^2 = 1/2$.

$\therefore$ We have Bendre-Schmidt's recurrence relation $u_{i,\,j+1} = \frac{1}{2}\,(u_{i-1,\,j} + u_{i+1,\,j})$ ...(1)

Now since $u(0, t) = 0 = u(8, t)$

$\therefore$ $u_{0,\,i} = 0$ and $u_{8,\,j} = 0$ for all values of j, *i.e.* the entries in the first and last column are zero.

Since $\quad u(x, 0) = 4x - \frac{1}{2}x^2$

$$\therefore \quad u_{i,0} = 4i - \frac{1}{2}i^2$$

$$= 0, 3.5, 6, 7.5, 8, 7.5, 6, 3.5 \text{ for } i = 0, 1, 2, 3, 4, 5, 6, 7 \text{ at } t = 0$$

These are the entries of the first row.

Putting $j = 0$ in (i), we have $u_{i,1} = \frac{1}{2}(u_{i-1,0} + u_{i+1,0})$

Taking $i = 1, 2, \ldots, 7$ successively, we get

$$u_{1,1} = \frac{1}{2}(u_{0,0} + u_{2,0}) = \frac{1}{2}(0 + 6) = 3$$

$$u_{2,1} = \frac{1}{2}(u_{1,0} + u_{3,0}) = \frac{1}{2}(3.5 + 7.5) = 5.5$$

$$u_{3,1} = \frac{1}{2}(u_{2,0} + u_{4,0}) = \frac{1}{3}(6 + 8) = 7$$

$$u_{4,1} = 7.5, u_{5,1} = 7, u_{6,1} = 5.5, u_{7,1} = 3.$$

These are the entries in the second row.

Putting $j = 1$ in (1), the entries of the third row are given by

$$u_{1,2} = \frac{1}{2}(u_{i-1,1} + u_{i+1,1})$$

Similarly putting $j = 2, 3, 4$ successively in (i), the entries of the fourth, fifth and sixth rows are obtained.

Hence the values of $u_{i,j}$ are as given in the following table :

j \ i	0	1	2	3	4	5	6	7	8
0	0	3.5	6	7.5	8	7.5	6	3.5	0
1	0	3	5.5	7	7.5	7	5.5	3	0
2	0	2.75	5	6.5	7	6.5	5	2.75	0
3	0	2.5	4.625	6	6.5	6	4.625	2.5	0
4	0	2.3125	4.25	5.5625	6	5.5625	4.25	2.3125	0
5	0	2.125	3.9375	5.125	5.5625	5.125	3.9375	2.125	0

Example 11.8. *Solve the equation* $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2}$ *subject to the conditions* $u(x, 0) = \sin \pi x$, $0 \le x \le 1$; $u(0, t) = u(1, t) = 0$, *using (a) Schmidt method, (b) Crank-Nicolson method, (c) Du Fort-Frankel method. Carryout computations for two levels, taking* $h = 1/3$, $k = 1/36$.

Sol. Here $c^2 = 1$, $h = 1/3$, $k = 1/36$ so that $\alpha = kc^2/h^2 = 1/4$.

Also $u_{1,0} = \sin \pi/3 = \sqrt{3}/2$, $u_{2,0} = \sin 2\pi/3 = \sqrt{3}/2$ and all boundary values are zero as shown in Fig. 11.26.

Fig. 11.26

(*a*) *Schmidt's formula* [(2) of § 11.9]

$$u_{i,j+1} = \alpha\, u_{i-1,j} + (1-2\alpha)\, u_{i,j} + \alpha\, u_{i+1,j}$$

becomes
$$u_{i,j+1} = \tfrac{1}{4}\,[u_{i-1,j} + 2u_{i,j} + u_{i+1,j}]$$

For $i = 1, 2\,; j = 0$:

$$u_{1,1} = \tfrac{1}{4}\,[u_{0,0} + 2u_{1,0} + u_{2,0}] = \tfrac{1}{4}\,(0 + 2 \times \sqrt{3}/2 + 3/2) = 0.65$$

$$u_{2,1} = \tfrac{1}{2}\,[u_{1,0} + 2u_{2,0} + u_{3,0}] = \tfrac{1}{4}\,(\sqrt{3}/2 + 2 \times \sqrt{3}/2 + 0) = 0.65$$

For $i = 1, 2, j = 1$:

$$u_{1,2} = \tfrac{1}{4}\,(u_{0,1} + 2u_{1,1} + u_{2,1}) = 0.49$$

$$u_{2,2} = \tfrac{1}{4}\,(u_{1,1} + 2u_{2,1} + u_{3,1}) = 0.49$$

(*b*) *Crank-Nicolson formula* [(4) of § 11.9] becomes

$$-\tfrac{1}{4}\,u_{i-1,j+1} + \tfrac{5}{2}\,u_{i,j+1} - \tfrac{1}{4}\,u_{i+1,j+1} = \tfrac{1}{4}\,u_{i-1,j} + \tfrac{3}{2}\,u_{i,j} + \tfrac{1}{4}\,u_{i+1,j}$$

For $i = 1, 2\,; j = 0$:

$$-u_{0,1} + 10u_{1,1} - u_{2,1} = u_{0,0} + 6u_{1,0} + u_{2,0}$$

i.e.
$$10\,u_{1,1} - u_{2,1} = 7\sqrt{3}/2$$

$$-u_{1,1} + 10u_{2,1} - u_{3,1} = u_{1,0} + 6u_{2,0} + u_{3,0}$$

i.e.
$$-u_{1,1} + 10u_{2,1} = 7\sqrt{3}/2$$

Solving these equations, we find

$$u_{1,1} = u_{2,1} = 0.67$$

For $i = 1, 2\,; j = 1$:

$$-u_{0,2} + 10u_{1,2} - u_{2,2} = u_{0,1} + 6u_{1,1} + u_{2,1}$$

i.e.
$$10u_{1,2} - u_{2,2} = 4.69$$

$$-u_{1,2} + 10u_{2,2} - u_{3,2} = u_{1,1} + 6u_{2,1} + u_{3,1}$$

i.e.
$$-u_{1,2} + 10u_{2,2} = 4.69$$

Solving these equations, we get $u_{1,2} = u_{2,2} = 0.52$.

(*c*) *Du Fort-Frankel formula* [(8) of § 11.9] becomes $u_{i,j+1} = \tfrac{1}{3}\,(u_{i,j-1} + u_{i-1,j} + u_{i+1,j})$

To start the calculations, we need $u_{1,1}$ and $u_{2,1}$.

We may take $u_{1,1} = u_{2,1} = 0.65$ from Schmidt method.

For $i = 1, 2\,; j = 1$:

$$u_{1,2} = \tfrac{1}{3}\,(u_{1,0} + u_{0,1} + u_{2,1}) = \tfrac{1}{3}\,(\sqrt{3}/2 + 0 + 0.65) = 0.5$$

$$u_{2,2} = \tfrac{1}{3}\,(u_{2,0} + u_{1,1} + u_{3,1}) = \tfrac{1}{3}\,(\sqrt{3}/2 + 0.65 + 0) = 0.5.$$

11.10. SOLUTION OF TWO DIMENSIONAL HEAT EQUATION

$$\frac{\partial u}{\partial t} = c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \qquad \text{...(1)}$$

The methods employed for the solution of one dimensional heat equation can be readily extended to the solution of (1).

Consider a square region $0 \le x \le y \le a$ and assume that u is known at all points within and on the boundary of this square.

If h be the step-size then a mesh point $(x, y, t) = (ih, jh, nl)$ may be denoted as simply (i, j, n).

Replacing the derivatives in (1) by their finite difference approximations, we get

$$\frac{u_{i,j,n+1} - u_{i,j,n}}{l}$$

$$= \frac{c^2}{h^2}\{(u_{i-1,j,n} - 2u_{i,j,n} + u_{i+1,j,n}) + (u_{i,j-1,n} - 2u_{i,j,n} + u_{i,j+1,n})\}$$

i.e. $$u_{i,j,n+1} = u_{i,j,n} + \alpha(u_{i-1,j,n} + u_{i+1,j,n} + u_{i,j+1,n} + u_{i,j-1,n} - 4u_{i,j,n}) \quad ...(2)$$

Fig. 11.27

where $\alpha = lc^2/h^2$. This equation needs the five points available on the nth plane (Fig. 11.27).

The computation process consists of point-by-point evaluation in the $(n + 1)$th plane using the points on the nth plane. It is followed by plane-by-plane evaluation. This method is known as *ADE (Alternating Direction Explicit) method*.

Example 11.9. *Solve the equation* $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$ *subject to the initial conditions* $u(x, y, 0) = \sin 2\pi x \sin 2\pi y$, $0 \le x, y \le 1$, *and the conditions* $u(x, y, t) = 0$, $t > 0$ *on the boundaries, using ADE method with* $h = \frac{1}{3}$ *and* $\alpha = \frac{1}{8}$. *(Calculate the results for one time level).*

Sol. The equation (2) above becomes

$$u_{i,j,n+1} = u_{i,j,n} + \tfrac{1}{8}(u_{i-1,j,n} + u_{i+1,j,n} + u_{i,j+1,n} + u_{i,j-1,n} - 4u_{i,j,n})$$

i.e. $$u_{i,j,n+1} = \tfrac{1}{2}u_{i,j,n} + \tfrac{1}{8}(u_{i-1,j,n} + u_{i+1,j,n} + u_{i,j+1,n} + u_{i,j-1,n}) \quad ...(1)$$

The mesh points and the computational model is given in Fig. 11.28.

At the zeroth level ($n = 0$), the initial and boundary conditions are

$$u_{i,j,0} = \sin\frac{2\pi i}{3}\sin\frac{2\pi j}{3}$$

and $$u_{i,0,0} = u_{0,j,0} = u_{3,j,0} = u_{i,3,0} = 0\,;\ i, j = 0, 1, 2, 3.$$

Now we calculate the mesh values at the first level :

For $n = 0$, (1) gives

$$u_{i,j,1} = \tfrac{1}{2}u_{i,j,0} + \tfrac{1}{8}(u_{i-1,j,0} + u_{i+1,j,0} + u_{i,j+1,0} + u_{i,j-1,0}) \quad ...(2)$$

(*i*) Put $i = j = 1$ in (2) :

$$u_{1,1,1} = \tfrac{1}{2}u_{1,1,0} + \tfrac{1}{8}(u_{0,1,0} + u_{2,1,0} + u_{1,2,0} + u_{1,0,0})$$

$$= \frac{1}{2}\left(\sin\frac{2\pi}{3}\right)^2 + \frac{1}{8}\left(0 + \sin\frac{4\pi}{3}\sin\frac{2\pi}{3} + \sin\frac{2\pi}{3}\sin\frac{4\pi}{3} + 0\right)$$

$$= \frac{3}{8} + \frac{1}{8}\left(-\frac{\sqrt{3}}{2} \times \frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} \times \frac{\sqrt{3}}{2}\right) = \frac{3}{16}$$

(*ii*) Put $i = 2, j = 1$ in (2)

$$u_{2,1,1} = \tfrac{1}{2}\, u_{2,1,0} + \tfrac{1}{8}\,(u_{1,1,0} + u_{3,1,0} + u_{2,2,0} + u_{2,0,0})$$

$$= \frac{1}{2} \sin\frac{4\pi}{3} \sin\frac{2\pi}{3} + \frac{1}{8}\left\{\left(\sin\frac{2\pi}{3}\right)^2 + 0 + \left(\sin\frac{4\pi}{3}\right)^2 + 0\right\}$$

$$= -\frac{1}{2}\left(\frac{\sqrt{3}}{2}\right)^2 + \frac{1}{8}\left\{\left(\frac{\sqrt{3}}{2}\right)^2 + \left(-\frac{\sqrt{3}}{2}\right)^2\right\} = -\frac{3}{16}.$$

Fig. 11.28

(*iii*) Put $i = 1, j = 2$ in (2) :

$$u_{1,2,1} = \tfrac{1}{2}\, u_{1,2,0} + \tfrac{1}{8}\,(u_{0,2,0} + u_{2,2,0} + u_{1,1,0})$$

$$= \frac{1}{2} \sin\frac{2\pi}{3} \sin\frac{4\pi}{3} + \frac{1}{8}\left\{0 + \left(\sin\frac{4\pi}{3}\right)^2 + 0 + \left(\sin\frac{2\pi}{3}\right)^2\right\}$$

$$= -\frac{3}{8} + \frac{1}{8}\left(\frac{3}{4} + \frac{3}{4}\right) = -\frac{3}{16}$$

(*iv*) Put $i = 2, j = 2$ in (2) :

$$u_{2,2,1} = \tfrac{1}{2}\, u_{2,2,0} + \tfrac{1}{8}\,(u_{1,2,0} + u_{3,2,0} + u_{2,3,0} + u_{2,1,0})$$

$$= \frac{1}{2}\left(\sin\frac{4\pi}{3}\right)^2 + \frac{1}{8}\left(\sin\frac{2\pi}{3}\sin\frac{4\pi}{3} + 0 + 0 + \sin\frac{4\pi}{3}\sin\frac{2\pi}{3}\right)$$

$$= \frac{3}{8} + \frac{1}{8}\left(-\frac{3}{4} - \frac{3}{4}\right) = -\frac{3}{16}$$

Similarly the mesh values at the second and higher levels can be calculated.

PROBLEMS 11.4

1. Find the solution of the parabolic equation $u_{xx} = 2u_t$ when $u(0, t) = u(4, t) = 0$ and $u(x, 0) = x(4 - x)$, taking $h = 1$. Find the values upto $t = 5$. *(Madras, B.E., 2001 S)*

2. Solve the equation $\dfrac{\partial^2 u}{\partial x^2} = \dfrac{\partial u}{\partial t}$ with the conditions $u(0, t) = 0$, $u(x, 0) = x(1 - x)$ and $u(1, t) = 0$. Assume $h = 0.1$. Tabulate u for $t = k$, $2k$ and $3k$ choosing an appropriate value of k. *(Anna, B.E., 2004)*

3. Given $\dfrac{\partial^2 f}{\partial x^2} - \dfrac{\partial f}{\partial t} = 0$; $f(0, t) = f(5, t) = 0$, $f(x, 0) = x^2(25 - x^2)$; find the values of f for $x = ih$ $(i = 0, 1, ..., 5)$ and $t = jk$ $(j = 0, 1, ..., 6)$ with $h = 1$ and $k = \frac{1}{2}$, using the explicit method.

4. Solve the boundary value problem $u_t = u_{xx}$, under the conditions $u(0, t) = u(1, t) = 0$ and $u(x, 0) = \sin \pi x$, $0 \le x \le 1$, by the Gauss-Seidal method. (Take $h = 0.2$ and $\alpha = \frac{1}{2}$). *(Rohtak, B.E., 2003)*

5. Solve the heat equation $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2}$ subject to the conditions $u(0, t) = u(1, t) = 0$ and $u(x, 0) = 2x$ for $0 \le x \le 1/2 = 2(1 - x)$ for $1/2 \le x \le 1$
Take $h = 1/4$ and k according to Bandre-Schmidt equation.

6. Solve the 2-dimensional heat equation $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}$ satisfying the initial condition : $u(x, y, 0) = \sin \pi x \sin \pi y$, $0 \le x, y \le 1$ and the boundary conditions : $u = 0$ at $x = 0$ and $x = 1$ for $t > 0$. Obtain the solution upto two time levels with $h = \frac{1}{3}$ and $\alpha = \frac{1}{8}$.

11.11. HYPERBOLIC EQUATIONS

The wave equation $\dfrac{\partial^2 u}{\partial t^2} = c^2 \dfrac{\partial^2 u}{\partial x^2}$ is the simplest example of hyperbolic partial differential equations. Its solution is the displacement function $u(x, t)$ defined for values of x from 0 to

l and for t from 0 to ∞, satisfying the initial and boundary conditions. The solution, as f parabolic equations, advances in an open-ended region (Fig. 11.22). In the case of hyperbol equations however, we have two initial conditions and two boundary conditions.

Such equations arise from convective type of problems in vibrations, wave mechani and gas dynamics.

11.12. SOLUTION OF WAVE EQUATION

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \qquad ...(1)$$

subject to the initial conditions : $u = f(x), \dfrac{\partial u}{\partial t} = g(x), 0 \le x \le 1$ at $t = 0$ ...(2)

and the boundary conditions : $u(0, t) = \phi(t), u(1, t) = \psi(t)$ ...(3)

Consider a rectangular mesh in the x-t plane spacing h along x direction and k along time direction. Denoting a mesh point $(x, t) = (ih, jk)$ as simply i, j, we have

$$\frac{\partial^2 u}{\partial x^2} = \frac{u_{i-1,j} - 2u_{i,j} + u_{i+1,j}}{h^2} \quad \text{and} \quad \frac{\partial^2 u}{\partial t^2} = \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h^2}$$

Replacing the derivatives in (1) by their above approximations, we obtain

$$u_{i,j-1} - 2u_{i,j} + u_{i,j+1} = \frac{c^2k^2}{h^2}(u_{i-1,j} - 2u_{i,j} + u_{i,j+1})$$

or
$$u_{i,j+1} = 2(1 - \alpha^2c^2)\, u_{i,j} + \alpha^2c^2(u_{i-1,j} + u_{i+1,j}) - u_{i,j-1} \qquad ...(4)$$

where
$$\alpha = k/h.$$

Now replacing the derivative in (2) by its central difference approximation, we get

$$\frac{u_{i,j+1} - u_{i,j-1}}{2k} = \frac{\partial u}{\partial t} = g(x) \qquad \text{[See (7) p. 279]}$$

or
$$u_{i,j+1} = u_{i,j-1} + 2kg(x) \text{ at } t = 0$$

i.e.
$$u_{i,1} = u_{i,-1} + 2kg(x) \text{ for } j = 0 \qquad ...(5)$$

Also initial condition $u = f(x)$ at $t = 0$ becomes $u_{i,-1} = f(x)$ ...(6)

Combining (5) and (6), we have $u_{i,1} = f(x) + 2kg(x)$ ...(7)

Also (3) gives $u_{0,j} = \phi(t)$ and $u_{1,j} = \psi(t)$.

Hence the explicit form (4) gives the values of $u_{i,j+1}$ at the $(j + 1)$th level when the nodal values at $(j - 1)$th and jth levels are known from (6) and (7) as shown in Fig 11.29.

Obs. 1. The coefficient of $u_{i,j}$ in (4) will vanish if $\alpha c = 1$ or $k = h/c$. Then (4) reduces to the simple form

$$u_{i,j+1} = u_{i-1,j} + u_{i+1,j} - u_{i,j-1} \qquad ...(8)$$

2. *For* $\alpha = 1/c$, the solution of (4) is stable and coincides with the solution of (1).

For $\alpha < 1/c$, the solution is stable but inaccurate.

For $\alpha > 1/c$, the solution is unstable.

3. The formula (4) converges for $\alpha \le 1$ *i.e.* $k \le h$.

Fig. 11.29

Example 11.10. *Evaluate the pivotal values of the equation $u_{tt} = 16u_{xx}$, taking $\Delta x = 1$ upto $t = 1.25$. The boundary conditions are $u(0, t) = u(5, t) = 0$, $u_t(x, 0) = 0$ and $u(x, 0) = x^2(5 - x)$.*

(Anna, B.E., 2004)

Sol. Here $c^2 = 16$.

$\therefore$ The difference equation for the given equation is

$$u_{i, j+1} = 2(1 - 16\alpha^2)\, u_{i, j} + 16\alpha^2 (u_{i-1, j} + u_{i+1, j}) - u_{i, j-1} \qquad \ldots(i)$$

where $\alpha = k/h$.

Taking $h = 1$ and choosing k so that the coefficient of $u_{i, j}$ vanishes, we have $16\alpha^2 = 1$, *i.e.* $k = h/4 = 1/4$.

$\therefore$ (1) reduces to $u_{i, j+1} = u_{i-1, j} + u_{i+1, j} - u_{i, j-1}$ $\qquad \ldots(ii)$

which gives a convergent solution (since $k/h < 1$). Its solution coincides with the solution of the given differential equation.

Now since $u(0, t) = u(5, t) = 0$, $\therefore$ $u_{0, j} = 0$ and $u_{5, j} = 0$ for all values of j

i.e. the entries in the first and last columns are zero.

Since $\quad u_{(x, 0)} = x^2 (5 - x)$

$\therefore \quad u_{i, 0} = i^2(5 - i) = 4, 12, 18, 16$ for $i = 1, 2, 3, 4$ at $t = 0$.

These are the entries for the first row.

Finally since $u_{t(x, 0)} = 0$ becomes

$$\therefore \quad \frac{u_{i, j+1} - u_{i, j-1}}{2k} = 0, \text{ when } j = 0, \text{ giving } u_{i, 1} = u_{i, -1} \qquad \ldots(iii)$$

Thus the entries of the second row are the same as those of the first row.

Putting $j = 0$ in (*ii*), $\quad u_{i, 1} = u_{i-1, 0} + u_{i+1, 0} - u_{i, -1} = u_{i-1, 0} + u_{i+1, 0} - u_{i, 1}$, using (*iii*)

or $$u_{i, 1} = \tfrac{1}{2}(u_{i-1, 0} + u_{i+1, 0}) \qquad \ldots(iv)$$

Taking $i = 1, 2, 3, 4$ successively, we obtain

$$u_{1, 1} = \tfrac{1}{2}(u_{0, 0} + u_{2, 0}) = \tfrac{1}{2}(0 + 12) = 6$$

$$u_{2, 1} = \tfrac{1}{2}(u_{1, 0} + u_{3, 0}) = \tfrac{1}{2}(4 + 18) = 11$$

$$u_{3,1} = \frac{1}{2}(u_{2,0} + u_{4,0}) = \frac{1}{2}(12 + 16) = 14$$

$$u_{4,1} = \frac{1}{2}(u_{3,0} + u_{5,0}) = \frac{1}{2}(18 + 0) = 9$$

These are the entries of the *second row*.

Putting $j = 1$ in (*ii*), we get $u_{i,2} = u_{i-1,1} + u_{i+1,1} - u_{i,0}$

Taking $i = 1, 2, 3, 4$ successively, we obtain

$$u_{1,2} = u_{0,1} + u_{2,1} - u_{1,0} = 0 + 11 - 4 = 7$$

$$u_{2,2} = u_{1,1} + u_{3,1} - u_{2,0} = 6 + 14 - 12 = 8$$

$$u_{3,2} = u_{2,1} + u_{4,1} - u_{3,0} = 11 + 9 - 18 = 2$$

$$u_{4,2} = u_{3,1} + u_{5,1} - u_{4,0} = 14 + 0 - 16 = -2$$

These are the entries of the *third row*.

Similarly putting $j = 2, 3, 4$ successively in (*ii*), the entries of the fourth, fifth and sixth rows are obtained.

Hence the values of $u_{i,j}$ are as shown in the table below :

$j \backslash i$	0	1	2	3	4	5
0	0	4	12	18	16	0
1	0	6	11	14	9	0
2	0	7	8	2	−2	0
3	0	2	−2	−8	−7	0
4	0	−9	−14	−11	−6	0
5	0	−16	−18	−12	−4	0

Example 11.11. *The transverse displacement u of a point at a distance x from one end and at any time t of a vibrating string satisfies the equation* $\partial^2 u/\partial t^2 = 4\partial^2 u/\partial x^2$, *with boundary conditions* $u = 0$ *at* $x = 0, t > 0$ *and* $u = 0$ *at* $x = 4, t > 0$ *and initial conditions* $u = x(4 - x)$ *and* $\partial u/\partial t = 0,\ 0 \le x \le 4$. *Solve this equation numerically for one half period of vibration, taking* $h = 1$ *and* $k = 1/2$.

Sol. Here, $h/k = 2 = c$.

$\therefore$ The difference equation for the given equation is

$$u_{i,j+1} = u_{i-1,j}\, u_{i+1,j} - u_{i,j-1} \qquad \text{...(i)}$$

which gives a convergent solution (since $k < h$).

Now since $u(0, t) = u(4, t) = 0$,

$\therefore$ $u_{0,j} = 0$ and $u_{4,j} = 0$ for all values of j.

i.e. the entries in the first and last columns are zero.

Since $u_{(x,0)} = x(4 - x)$,

$\therefore$ $u_{i,0} = i(4 - i) = 3, 4, 3$ for $i = 1, 2, 3$ at $t = 0$.

These are the entries of the *first row*.

Also $u_t(x, 0) = 0$ becomes

$$\frac{u_{i,j+1} - u_{i,j-1}}{2k} = 0 \text{ when } j = 0, \text{ giving } u_{i,1} = u_{i,-1} \qquad \text{...(ii)}$$

Putting $j = 0$ in (*i*), $u_{i,1} = u_{i-1,0} + u_{i+1,0} - u_{i,-1} = u_{i-1,0} + u_{i+1,0} - u_{i,1}$, using (*ii*)

or $$u_{i,1} = \frac{1}{2}(u_{i-1,0} + u_{i+1,0}) \qquad ...(iii)$$

Taking $i = 1, 2, 3$ successively, we obtain

$$u_{1,1} = \frac{1}{2}(u_{0,0} + u_{1,0}) = 2\,;\; u_{2,1} = \frac{1}{2}(u_{1,0} + u_{3,0}) = 3,\; u_{3,1} = \frac{1}{2}(u_{2,0} + u_{4,0}) = 2$$

These are the entries of the *2nd row*.

Putting $j = 1$ *in* (*i*), $u_{i,2} = u_{i-1,1} + u_{i+1,1} - u_{i,0}$

Taking $i = 1, 2, 3$, successively, we get

$$u_{1,2} = u_{0,1} + u_{2,1} - u_{1,0} = 0 + 3 - 3 = 0$$
$$u_{2,2} = u_{1,1} + u_{3,1} - u_{2,0} = 2 + 2 - 4 = 0$$
$$u_{3,2} = u_{2,1} + u_{4,1} - u_{3,0} = 3 + 0 - 3 = 0$$

These are the entries of the *3rd row* and so on.

Now the equation of the vibrating string of length l is $u_{tt} = c^2 u_{xx}$.

$\therefore$ Its period of vibration $= \dfrac{2l}{c} = \dfrac{2 \times 4}{2} = 4$ sec. $\qquad [\because\ l = 4 \text{ and } c = 2]$

This shows that we have to compute $u_{(x,t)}$ upto $t = 2$

i.e. similarly we obtain the values of $u_{i,2}$ (4th row) and $u_{i,3}$ (5th row).

Hence the values of $u_{i,j}$ are as shown in the table below :

$j \backslash i$	0	1	2	3	4
0	0	3	4	3	0
1	0	2	3	2	0
2	0	0	0	0	0
3	0	−2	−3	−2	0
4	0	−3	−4	−3	0

PROBLEMS 11.5

1. Solve the boundary value problem $u_{tt} = u_{xx}$ with the conditions $u(0, t) = u(1, t) = 0$, $u(x, 0) = \frac{1}{2}x(1-x)$ and $u_t(x, 0) = 0$, taking $h = k = 0.1$ for $0 \le t \le 0.4$. Compare your solution with the exact solution at $x = 0.5$ and $t = 0.3$. *(V.T.U., B.E., 2000)*

2. The transverse displacement of a point at a distance x from one end and at any time t of a vibrating string satisfies the equation $\dfrac{\partial^2 u}{\partial t^2} = 25\dfrac{\partial^2 u}{\partial x^2}$, with the boundary conditions $u(0, t) = u(5, t) = 0$ and the initial conditions $u(x, 0) = \begin{cases} 20x & \text{for } 0 \le x < 1 \\ 5(5-x) & \text{for } 1 \le x < 5 \end{cases}$ and $u_t(x, 0) = 0$. Solve this equation numerically for one half period of vibration, taking $h = 1$, $k = 0.2$.