# Contents

# LIST OF PROGRAMS

# USEFUL RESULTS

## I. BASIC INFORMATION & ERRORS

### 1. Useful Data

| | | | |
|---|---|---|---|
| $e = 2.7183$ | $1/e = 0.3679$ | $\log_e 2 = 0.6931$ | $\log_e 3 = 1.0986$ |
| $\pi = 3.1416$ | $1/\pi = 0.3183$ | $\log_e 10 = 2.3026$ | $\log_{10} e = 0.4343$ |
| $\sqrt{2} = 1.4142$ | $\sqrt{3} = 1.732$ | $1$ rad. $= 57° \, 17' \, 45''$ | $1° = 0.0174$ rad. |

### 2. Conversion Factors

| | |
|---|---|
| 1 ft. = 30.48 cm = 0.3048 m | 1 m = 100 cm = 3.2804 ft. |
| 1 ft$^2$ = 0.0929 m$^2$ | 1 acre = 4840 yd$^2$ = 4046.77 m$^2$ |
| 1 ft$^3$ = 0.0283 m$^3$ | 1 m$^3$ = 35.32 ft$^3$ |
| 1 m/sec = 3.2804 ft/sec. | 1 mile/h = 1.609 km/h. |

### 3. Some Notations

| | | | |
|---|---|---|---|
| $\in$ | belongs to | $\cup$ | union |
| $\notin$ | doesnot belong to | $\cap$ | intersection |
| $\Rightarrow$ | implies | $\ni$ | such that |
| $\Leftrightarrow$ | implies and implied by | | |

*Factorial n i.e.* $n! = n(n-1)(n-2) \ldots\ldots 3 \,.\, 2 \,.\, 1.$

*Double factorials :* $(2n)!! = 2n(2n-2)(2n-4) \ldots\ldots 6 \,.\, 4 \,.\, 2.$

$$(2n-1)!! = (2n-1)(2n-3)(2n-5) \ldots\ldots 5 \,.\, 3 \,.\, 1.$$

*Stirling's approximation.* When $n$ is large $n! \sim \sqrt{2\pi n} \,.\, n^n e^{-n}$.

**4.** If $X$ is the true value of a quantity and $X'$ is its approximate value, then

(i) *Absolute error* $= |X - X'|$

(ii) *Relative error* $= \left| \dfrac{X - X'}{X} \right|$

(iii) *Percentage error* $= 100 \left| \dfrac{X - X'}{X} \right|$

**5.** *If $\delta y$ is the error in the function $y = f(x_1, x_2, \ldots\ldots, x_n)$ corresponding to the errors $\delta x_1, \delta x_2, \ldots\ldots \delta x_n$, then*

$$\delta y = \frac{\partial y}{\partial x_1} \delta x_1 + \frac{\partial y}{\partial x_2} \delta x_2 + \ldots\ldots + \frac{\partial y}{\partial x_n} \delta x_n.$$

**6.** *Relative error of a product of n numbers*

*= Algebraic sum of their relative errors approximately*

## II. SOLUTION OF ALGEBRAIC AND TRANCENDENTAL EQUATIONS

**1. Intermediate value property :** *If $f(x)$ is continuous in the interval $[a, b]$ and $f(a)$, $f(b)$ have different signs, then the equation $f(x) = 0$ has at least one root between $x = a$ and $x = b$.*

2. **Descartes rule of signs :** *The equation $f(x) = 0$ cannot have more positive roots than the change of signs in $f(x)$ and cannot have more negative roots than the change of signs in $f(-x)$.*

3. *If $\alpha_1, \alpha_2, \alpha_3$ ...... be the roots of the equation $a_0x^n + a_1x^{n-1} + a_2x^{n-2} + a_3x^{n-3} + \ldots\ldots = 0$, then*

$$\Sigma\alpha_1 = -\frac{a_1}{a_0}\ ;\ \Sigma\alpha_1\alpha_2 = \frac{a_2}{a_0}\ ;\ \Sigma\alpha_1\alpha_2\alpha_3 = -\frac{a_3}{a_0}\ ;\ \text{etc.}$$

4. **Bisection method :** *Iteration formula is* $x_3 = \frac{1}{2}(x_1 + x_2)$

   This process is continued till the difference between two consecutive values is negligible.

5. **Method of false-position or Regula falsi method :** *Iteration formula is*

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

   This process is repeated till the difference between two consecutive values is negligible.

6. **Secant method :**

   *Iteration formula* is $x_2 = x_1 - \dfrac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_1)$

   **Obs.** If secant method once converges, its rate of convergence is 1.6 which is faster than that of method of false position.

7. **Iteration method :** *Writing $f(x) = 0$ as $x = \phi(x)$ and taking $x_0$ as the initial root of the given equation, the approximations to the root are $x_i = \phi(x_i)$ such that $\phi'(x) < 1$.*

8. **Newton-Raphson method** *alogrithm is*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}\quad (n = 0, 1, 2, \ldots\ldots)$$

   **Obs.** Condition for its convergence is $|f(x)f''(x)| < |f'(x)|^2$. Newton's method has a second order of convergence. If this method once converges, it converges faster than the Regula-falsi method and is preferred.

9. **Method of Least squares :** (*i*) *Curve of best fit* $y = a + bx$

   *Normal equations :* $\Sigma y = na + b\Sigma x$, $\Sigma xy = a\Sigma x + b\Sigma x^2$

   To find $a$, $b$, solve these equations.

   (*ii*) *Curve of best fit* $y = a + bx + cx^2$

   *Normal equations :* $\Sigma y = na + b\Sigma x + c\Sigma x^2$

   $\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$, $\Sigma x^2y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$.

   To find $a$, $b$, $c$, solve these equations.

## III. SOL. OF SIMULTANEOUS ALGEBRAIC EQUATIONS

1. **Numerical solution of linear simultaneous equations are**

   (*i*) *Direct methods* (*ii*) *Indirect (or Iterative) methods*

   Matrix Inversion method, Gauss-elimination method, Gauss-Jordan method and Factorization method are *direct methods* ; Gauss-Jacobi method, Gauss-Seidal method and Relaxation methods are *indirect methods*.

2. **Matrix Inversion method.** For the equations :

$$a_1x + b_1y + c_1z = d_1,\ a_2x + b_2y + c_2z = d_2,\ a_3x + b_3y + c_3z = d_3,$$

if $A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$, $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $D = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$

then $$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{|A|} \begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} \times \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

where $A_1$, $B_1$, etc. are the cofactors of $a_1$, $b_1$ etc. in the determinant $|A|$.

3. **In Gauss-elimination method**, *the coefficient matrix is transformed to* **upper triangular matrix.**

4. **In Gauss-Jordan method**, *the coefficient matrix is transformed to* **diagonal matrix.**

5. **Gauss-Jordan method of finding the inverse of a matrix A.** *The matrices A and I are written side by side and the same transformations are performed on both. As soon as A is reduced to I, the other matrix represents* $A^{-1}$.

6. **The convergence in Gauss-Seidal method** is thrice as fast as in Jacobi's method.

7. **The condition for Gauss-Jacobi's method to converge is** *that the coefficient matrix should be diagonally dominant.*

## IV. FINITE DIFFERENCES AND INTERPOLATION

1. **Forward differences :** $\Delta y_r = y_{r+1} - y_r$.
   **Backward differences :** $\nabla y_r = y_r - y_{r-1}$
   **Central differences :** $\delta y_{x-1/2} = y_x - y_{x-1}$

2. **Relations between operators :**
   (i) $\Delta = E - 1$ (ii) $\nabla = 1 - E^{-1}$
   (iii) $\delta = E^{1/2} - E^{-1/2}$ (iv) $\mu = \frac{1}{2}(E^{1/2} + E^{-1/2})$
   (v) $\Delta = E\nabla = \nabla E = \delta E^{1/2}$ (vi) $E = e^{hD}$

3. **Factorial notation.** The product $x(x-1)(x-2) \ldots\ldots (x-r+1)$ *is denoted by* $[x]^r$ *and is called a factorial.*
   **Factorial polynomial** *is defined as* $[x]^n = x(x-h)(x-2h) \ldots\ldots [x-(n-1)h]$.

4. **Newton's forward interpolation formula :**

$$y_p = y_0 + p\,\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \ldots\ldots$$

where $p = (x - x_0)/h$.

5. **Newton's backward interpolation formula :**

$$y_p = y_n + p\,\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \ldots \text{ where } p = (x - x_n)/h$$

6. **Gauss forward interpolation formula :**

$$y_p = y_0 + p\,\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-1} + \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \ldots\ldots$$

7. **Gauss's backward interpolation formula :**

$$y_p = y_0 + p\,\Delta y_{-1} + \frac{(p+1)p}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-2} + \frac{(p+2)(p+1)p(p-1)}{4!}\Delta^4 y_{-2} + \ldots\ldots$$

8. **Stirling's formula :**

$$y_p = y_0 + p\left(\frac{\Delta y_0 + \Delta y_{-1}}{2}\right) + \frac{p^2}{2!}\Delta^2 y_{-1} + \frac{p(p^2-1)}{3!}\left(\frac{\Delta^3 y_{-1} + \Delta^3 y_{-2}}{2}\right)$$
$$+ \frac{p^2(p^2-1)}{4!}\Delta^4 y_{-2} + \ldots$$

9. **Bessel's formula :**

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\,\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{\left(p-\frac{1}{2}\right)p(p-1)}{3!}\Delta^3 y_{-1}$$
$$+ \frac{(p+1)p(p-1)(p-2)}{4!}\,\frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \ldots\ldots$$

10. **Laplace-Everett's formula :**

$$y_p = qy_0 + \frac{q(q^2-1^2)}{3!}\Delta^2 y_{-1} + \frac{q(q^2-1^2)(q^2-2^2)}{5!}\Delta^4 y_{-2} + \ldots\ldots$$
$$+ py_1 + \frac{p(p^2-1^2)}{3!}\Delta^2 y_0 + \frac{p(p^2-1^2)(p^2-2^2)}{5!}\Delta^4 y_{-1} + \ldots\ldots \quad [q = 1 - p$$

11. **Lagrange's interpolation formula :**

$$y = \frac{(x-x_1)(x-x_2)\ldots\ldots(x-x_n)}{(x_0-x_1)(x_0-x_2)\ldots\ldots(x_0-x_n)}\,y_0 + \frac{(x-x_0)(x-x_2)\ldots\ldots(x-x_n)}{(x_1-x_0)(x_1-x_2)\ldots\ldots(x_1-x_n)}\,y_1$$
$$+ \ldots\ldots + \frac{(x-x_0)(x-x_1)\ldots\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\ldots\ldots(x_n-x_{n-1})}\,y_n.$$

12. **Lagrange's inverse interpolation formula :**

$$x = \frac{(y-y_1)(y-y_2)\ldots\ldots(y-y_n)}{(y_0-y_1)(y_0-y_2)\ldots\ldots(y_0-y_n)}\,x_0 + \frac{(y-y_0)(y-y_2)\ldots\ldots(y-y_n)}{(y_1-y_0)(y_1-y_2)\ldots\ldots(y_1-y_n)}\,x_1$$
$$+ \ldots\ldots + \frac{(y-y_0)(y-y_1)\ldots\ldots(y-y_{n-1})}{(y_n-y_0)(y_n-y_1)\ldots\ldots(y_n-y_{n-1})}\,x_n.$$

13. **Newton's divided difference formula :**

$$y = f(x) = y_0 + (x-x_0)\,[x_0, x_1] + (x-x_0)(x-x_1)\,[x_0, x_1, x_2]$$
$$+ (x-x_0)(x-x_1)(x-x_2)\,[x_0, x_1, x_2, x_3] + \ldots\ldots$$

14. **Hermite interpolation formula :**

$$P(x) = [1 - 2(x - x_0)\, L'_0(x_0)]\, [L_0(x)]^2\, y(x_0) + (x - x_0)\, [L_0(x)]^2\, y'(x_0)$$
$$+ [1 - 2(x - x_1)\, L'_1(x_1)]\, [L_1(x)]^2\, y(x_1) + (x - x_1)\, [L_1(x)]^2\, y'(x_1)$$
$$+ [1 - 2]\, (x - x_2) L'_2(x_2)]\, [L_2(x)]^2\, y(x_2) + (x - x_2)\, [L_2(x)]^2\, y'(x_2)$$
$$+ \ldots\ldots$$

15. **Cubic Spline interpolation formula :**

$$f(x) = \frac{1}{6h}\, [(x_{i+1} - x)^3 M_i + (x - x_i)^3 M_{i+1}]$$
$$+ \frac{1}{h}\left[(x_{i+1} - x)\left(y_i - \frac{h^2}{6} M_i\right) + (x - x_i)\left(y_{i+1} - \frac{h^2}{6} M_{i+1}\right)\right]$$

where $\quad M_{i-1} + 4M_i + M_{i+1} = \dfrac{6}{h^2}(y_{i-1} - 2y_i + y_{i+1}),\ i = 1, 2, 3, \ldots, (n-1)$

and $\quad M_0 = 0,\ M_n = 0.$

## V. NUMERICAL DIFFERENTIATION AND INTEGRATION

1. **Forward difference formulae :**

$$\left(\frac{dy}{dx}\right)_{x_0} = \frac{1}{h}\left[\Delta y_0 - \frac{1}{2}\Delta^2 y_0 + \frac{1}{3}\Delta^3 y_0 - \frac{1}{4}\Delta^4 y_0 + \ldots\ldots\right]$$

$$\left(\frac{d^2 y}{dx^2}\right)_{x_0} = \frac{1}{h^2}\left[\Delta^2 y_0 - \Delta^3 y_0 + \frac{11}{12}\Delta^4 y_0 - \ldots\ldots\right] \text{ and so on.}$$

2. **Backward difference formulae :**

$$\left(\frac{dy}{dx}\right)_{x_n} = \frac{1}{h}\left[\nabla y_n + \frac{1}{2}\nabla^2 y_n + \frac{1}{3}\nabla^3 y_n + \frac{1}{4}\nabla^4 y_n + \ldots\ldots\right]$$

$$\left(\frac{d^2 y}{dx^2}\right)_{x_n} = \frac{1}{h^2}\left[\nabla^2 y_n + \nabla^3 y_n + \frac{11}{12}\nabla^4 y_n + \ldots\ldots\right] \text{ and so on.}$$

3. **Central difference formulae :**

*Stirling's formula* gives

$$\left(\frac{dy}{dx}\right)_{x_0} = \frac{1}{h}\left[\frac{\Delta y_0 + \Delta y_{-1}}{2} - \frac{1}{6}\,\frac{\Delta^3 y_{-1} + \Delta^3 y_{-2}}{2} + \frac{1}{30}\,\frac{\Delta^5 y_{-2} + \Delta^5 y_{-3}}{2} + \ldots\ldots\right]$$

$$\left(\frac{d^2 y}{dx^2}\right)_{x_0} = \frac{1}{h^2}\left[\Delta^2 y_{-1} - \frac{1}{12}\Delta^4 y_{-2} + \frac{1}{90}\Delta^6 y_{-3} - \ldots\ldots\right]$$

4. **Trapezoidal rule :**

$$\int_{x_0}^{x_0+nh} f(x)\,dx = \frac{h}{2}\,[(y_0 + y_n) + 2(y_1 + y_2 + \ldots\ldots + y_{n-1})]$$

5. **Simpson's 1/3rd rule :**

$$\int_{x_0}^{x_0+nh} f(x)\,dx = \frac{h}{3}\,[(y_0 + y_n) + 4(y_1 + y_3 + \ldots\ldots + y_{n-1}) + 2(y_2 + y_4 + \ldots\ldots + y_{n-2})]$$

(Number of sub-intervals should be taken as even)

6. **Simpson's 3/8th rule :**

$$\int_{x_0}^{x_0+nh} f(x)\,dx = \frac{3h}{8}\,[(y_0 + y_n) + 3(y_1 + y_2 + y_4 + y_5 + \ldots\ldots + y_{n-1}) + 2(y_3 + y_6 + \ldots\ldots + y_{n-3})]$$

(Number of sub-intervals should be taken as a multiple of 3)

7. **Weddle's rule :**

$$\int_{x_0}^{x_0+nh} f(x)\,dx = \frac{3h}{10}\,[y_0 + 5y_1 + y_2 + 6y_3 + y_4 + 5y_5 + 2y_6 + 5y_7 + \ldots\ldots]$$

(Number of sub-intervals should be taken as a multiple of 6)

8. **Errors :**

| *Rule* | *No. of intervals (multiples of)* | *Error* | *Order of error* |
|---|---|---|---|
| Trapezoidal | Any | $-\frac{h^2}{12}\,y''$ | $h^2$ |
| Simpson's 1/3 | 2 | $-\frac{h^4}{180}\,y^{iv}$ | $h^4$ |
| Simpson's 3/8 | 3 | $-\frac{3h^5}{80}\,y^{iv}$ | $h^5$ |
| Weddle's | 6 | $-\frac{h^7}{140}\,y_0^{vi}$ | $h^7$ |

9. **Romberg's method :**

$$I(h, h/2) = \frac{1}{3}\,[4I(h/2) - I(h)]$$

The computation is continued till two successive values are equal.

10. **Gaussian integration :**

*(i) two point formula :* $\int_{-1}^{1} f(x)\,dx = f\left(\frac{-1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right)$

*(ii) three point formula :* $\int_{-1}^{1} f(x)\,dx = \frac{8}{9}f(0) + \frac{5}{9}\left[f\left(-\sqrt{\frac{3}{5}}\right) + f\left(\sqrt{\frac{3}{5}}\right)\right]$

*(iii) To apply Gaussian integration, the limits of integration a, b are changed to – 1, 1 by the transformation* $x = \frac{1}{2}(b-a)u + \frac{1}{2}(b+a)$.

11. **Double integration :**

(*i*) *Trapezoidal rule* .

$$I = \frac{hk}{4} [(f_{00} + f_{0m}) + 2(f_{01} + f_{02} + \ldots + f_{0, m-1})$$
$$+ (f_{n0} + f_{nm}) + 2(f_{n1} + f_{n2} + \ldots\ldots + f_{n, m-1})$$
$$+ 2\sum_{i=1}^{n-1} \{[(f_{i0} + f_{im}) + 2(f_{i1} + f_{i2} + \ldots\ldots + f_{i,m-1})]\} \text{ where } f_{ij} = f(x_i, y_j)$$

(*ii*) *Simpson's rule :*

$$\int_{y_{j-1}}^{y_{j+1}} \int_{x_{i-1}}^{x_{i+1}} f(x, y)\, dx\, dy = \frac{hk}{9} [(f_{i-1, j-1} + 4f_{i-1, j} + f_{i-1, j+1}) + 4(f_{i, j-1} + 4f_{i, j} + 4f_{i, j+1})$$
$$+ (f_{i+1, j-1} + 4f_{i+1, j} + f_{i+1, j+1})]$$

Adding all such intervals, we get $I$.

## VI. NUM. SOL. OF ORDINARY DIFFERENTIAL EQUATIONS

1. **Picard's method :** $y_1 = y_0 + \int_{x_0}^{x} f(x, y_0)\, dx$

$$y_2 = y_0 + \int_{x_0}^{x} f(x, y_1)\, dx \text{ etc.}$$

2. **Taylor's method :**

$$y = y_0 + (x - x_0)(y')_0 + \frac{(x - x')^2}{2!}(y'')_0 + \frac{(x - x')^3}{3!}(y''')_0 + \ldots\ldots$$

3. **Euler's method :** $y_2 = y_1 + hf(x_0 + h, y_1)$
Repeat this process till $y_2$ is stationary. Then calculate $y_3$ and so on.

4. **Modified Euler's method :** $y_2 = y_1 + \frac{h}{2} [(f(x_0 + h, y_1) + f(x_0 + 2h, y_2)]$

Repeat this step, till $y_2$ becomes stationary. Then calculate $y_3$ and so on and so on.

5. **Runge Kutta method :** $y_1 = y_0 + k$, where $k = \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$

such that $k_1 = hf(x_0, y_0)$, $k_2 = hf(x_0 + h/2, y_0 + k_{1/2})$
$k_3 = hf(x_0, h/2, y_0 + k_2/2)$, $k_4 = hf(x_0 + h, y_0 + k_3)$

6. **Milne's method :**

(*i*) *Predictor formula :* $y_4 = y_0 + \frac{4h}{3}(2f_1 - f_2 + 2f_3)$

(*ii*) *Corrector formula :* $y_4 = y_2 + \frac{h}{3}(f_2 + 4f_3 + f_4)$

7. **Adams-Bashforth method :**

(i) *Predictor formula :* $y_1 = y_0 + \frac{h}{24}(55 f_0 - 59 f_{-1} + 37 f_{-2} - 9f_{-3})$

(ii) *Corrector formula :* $y_1 = y_0 + \frac{h}{24}(9f_1 + 19f_0 - 5f_{-1} + f_{-2})$

*(Four prior values are required to find the next values by Milne's or Adams-Bashforth method)*

8. **Central-difference approximations :**

$$y_i' = \frac{1}{2h}(y_{i+1} - y_{i-1})$$

$$y_i'' = \frac{1}{h^2}(y_{i+1} - 2y_i + y_{i-1})$$

$$y_i''' = \frac{1}{2h^3}(y_{i+2} - 2y_{i+1} + 2y_{i-1} - y_{i-2})$$

$$y_i^{\text{iv}} = \frac{1}{h^4}(y_{i+2} - 4y_{i+1} + 6y_i - 4y_{i-1} + y_{i-2})$$

## VII. NUM. SOL. OF PARTIAL DIFFERENTIAL EQUATIONS

1. **Classification of second order equation :**

$$A(x, y)\frac{\partial^2 u}{\partial x^2} + B(x, y)\frac{\partial^2 u}{\partial x \partial y} + C(x, y)\frac{\partial^2 u}{\partial y^2} + F\left(x, y, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) = 0$$

is said to be

(i) *elliptic* if $B^2 - 4AC < 0$

(ii) *parabolic* if $B^2 - 4AC = 0$

(iii) *hyperbolic* if $B^2 - 4AC > 0$.

2. **Laplace equation :** $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$

(i) *Standard 5-point formula :*

$$u_{i,j} = \frac{1}{4}[u_{i-1,j} + u_{i+1,j} + u_{i,j+1} + u_{i,j-1}]$$

(ii) *Diagonal 5-point formula :*

$$u_{i,j} = \frac{1}{4}[u_{i-1,j+1} + u_{i+1,j-1} + u_{i+1,j+1} + u_{i-1,j-1}]$$

(Four conditions are required to solve Laplace equation.)

3. **Poisson's equation :** $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y)$

*Standard 5-point formula :*

$$u_{i-1,j} + u_{i+1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j} = h^2 f(ih, jh)$$

**4. One-dimensional Heat equation :**

$$\frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}$$

(*i*) *Schmidt formula* : $u_{i,j+1} = \alpha u_{i-1,j} + (1-2\alpha)u_{i,j} + \alpha u_{i+1,j}$, where $\alpha = kc^2/h^2$

(*ii*) *Bendre-Schmidt relation* ; $u_{i,j+1} = \frac{1}{2}(u_{i-1,j} + u_{i+1,j})$

[when $\alpha = \frac{1}{2}$, (*i*) reduces to (*ii*)]

(*iii*) *Crank-Nicolson formula* :

$$\alpha(u_{i+1,j+1} + u_{i-1,j+1}) - 2(\alpha+1)u_{i,j+1} = 2(\alpha-1)u_{i,j} - \alpha(u_{i+1,j} + u_{i-1,j})$$

**5. Wave equation :** $\dfrac{\partial^2 u}{\partial t^2} = c^2 \dfrac{\partial^2 u}{\partial x^2}$

(*i*) *Explicit formula for solution is*

$$u_{i,j+1} = 2(1-\alpha^2c^2)\,u_{i,j} + \alpha^2c^2\,(u_{i-1,j} + u_{i+1,j}) - u_{i,j-1} \text{ where } \alpha = k/h$$

(*ii*) If $\alpha$ is so choosen that coefficient of $u_{i,j}$ is zero, then

$\alpha(= k/h) = 1/c$ *i.e.* $k = h/c$, then (*i*) takes the simplified form

$$u_{i,j+1} = u_{i-1,j} + u_{i+1,j} - u_{i,j-1}$$

# 1

# APPROXIMATIONS AND ERRORS IN COMPUTATION



## 1.1. INTRODUCTION

The limitations of analytical methods in practical applications have led scientists and engineers to evolve numerical methods. We know that exact methods often fail in drawing plausible inferences from a given set of tabulated data or in finding roots of transcendental equations or in solving non-linear differential equations. There are many more such situations where analytical methods are unable to produce desirable results. Even if analytical solutions are available, these are not amenable to direct numerical interpretation. *The aim of numerical analysis is therefore, to provide constructive methods for obtaining answers to such problems in a numerical form.*

With the advent of high speed computers and increasing demand for numerical solution to various problems, numerical techniques have become indispensible tools in the hands of engineers and scientists.

The input information is rarely exact since it comes from some measurement or the other and the method also introduces further error. As such, the error in the final result may be due to error in the initial data or in the method or both. Our effort will be to minimize these errors, so as to get best possible results. We therefore begin by explaining various kinds of approximations and errors which may occur in a problem and derive some results on error propagation in numerical calculations.

## 1.2. ACCURACY OF NUMBERS

**(1)** *Approximate numbers.* There are two types of numbers *exact* and *approximate*. *Exact numbers* are 2, 4, 9, 13, 7/2, 6.45, ... etc. But there are numbers such as 4/3 ( = 1.33333 ...), $\sqrt{2}$ (= 1.414213 ...) and $\pi$ ( = 3.141592 ...) which cannot be expressed by a finite number of digits. These may be approximated by numbers 1.3333, 1.4142 and 3.1416 respectively. Such numbers which represent the given numbers to a certain degree of accuracy are called *approximate numbers*.

**(2)** *Significant figures.* The digits used to express a number are called *significant digits (figures).* Thus each of the numbers 7845, 3.589, 0.4758 contains four significant figures while the numbers 0.00386, 0.000587 and 0.0000296 contain only three significant figures since zeros only help to fix the position of the decimal point. Similarly the numbers 45000 and 7300.00 have two significant figures only.

**(3)** *Rounding off.* There are numbers with large number of digits *e.g.*, 22/7 = 3.142857143. In practice, it is desirable to limit such numbers to a manageable number of digits such as 3.14 or 3.143. This process of dropping unwanted digits is called *rounding off.*

**(4) Rule** *to round off a number to n significant figures :*

(*i*) *Discard all digits to the right of the nth digit.*

(*ii*) *If this discarded number is*

(*a*) *less than half a unit in the nth place, leave the nth digit unchanged ;*

(*b*) *greater than half a unit in the nth place, increase the nth digit by unity ;*

(*c*) *exactly half a unit in the nth place, increase the nth digit by unity if it is odd otherwise leave it unchanged.*

For instance, the following numbers rounded off to three significant figures are :

| | | | | | |
|---|---|---|---|---|---|
| 7.893 | to | 7.89 | 3.567 | to | 3.57 |
| 12.865 | to | 12.9 | 84767 | to | 84800 |
| 6.4356 | to | 6.44 | 5.8254 | to | 5.82 |

Also the numbers 6.284359, 9.864651, 12.464762 rounded off to four places of decimal at 6.2844, 9.8646, 12.4648 respectively.

**Obs.** The numbers thus rounded off to *n* significant figures (or *n* decimal places) are said to be *correct to n significant figures (or n decimal places).*

## 1.3. ERRORS

In any numerical computation, we come across the following types of errors :

**(1)** *Inherent errors.* Errors which are already present in the statement of a problem before its solution, are called *inherent errors.* Such errors arise either due to the given data being approximate or due to the limitations of mathematical tables, calculators or the digital computer. Inherent errors can be minimized by taking better data or by using high precision computing aids.

**(2)** *Rounding errors* arise from the process of rounding off the numbers during the computation. Such errors are unavoidable in most of the calculations due to the limitations of the computing aids. Rounding errors can, however, be reduced :

(*i*) by changing the calculation procedure so as to avoid subtraction of nearly equal numbers or division by a small number ;

or (*ii*) by retaining at least one more significant figure at each step than that given in the data and rounding off at the last step.

**(3)** *Truncation errors* are caused by using approximate results or on replacing an infinite process by a finite one. If we are using a decimal computer having a fixed word length of 4 digits, rounding off of 13.658 gives 13.66 whereas truncation gives 13.65.

For example, if $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \infty = X$ (say) is replaced by $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} = X'$ (say), then the truncation error is $X - X'$.

Truncation error is a type of algorithm error.

(4) *Absolute, Relative and Percentage errors*. If $X$ is the true value of a quantity and $X'$ is its approximate value, then $|X - X'|$ *i.e.* | Error | is called the *absolute error* $E_a$.

The *relative error* is defined by $E_r = \left|\frac{X - X'}{X}\right|$ *i.e.* $\frac{|\text{Error}|}{|\text{True value}|}$

and the *percentage error* is $E_p = 100\, E_r = 100\left|\frac{X - X'}{X}\right|$.

If $\overline{X}$ be such a number that $|X - X'| \le \overline{X}$, then $\overline{X}$ is an upper limit on the magnitude of absolute error and measures the *absolute accuracy*.

**Obs. 1.** The relative and percentage errors are independent of the units used while absolute error is expressed in terms of these units.

**Obs. 2.** *If a number is correct to n decimal places then the error* $= \frac{1}{2}10^{-n}$.

For example, if the number is 3.1416 correct to 4 decimal places, then the error

$$= \frac{1}{2}10^{-4} = 0.00005.$$

## 1.4. USEFUL RULES FOR ESTIMATING ERRORS

To estimate the errors which creep in when the numbers in a calculation are truncated or rounded off to a certain number of digits, the following rules are useful.

If the approximate value of a number $X$ having $n$ decimal digits is $X'$, then

(1) *Absolute error due to truncation to k digits*

$$= |X - X'| < 10^{n-k}$$

(2) *Absolute error due to rounding off to k digits*

$$= |X - X'| < \frac{1}{2}10^{n-k}$$

(3) *Relative error due to truncation to k digits*

$$= \left|\frac{X - X'}{X}\right| < 10^{1-k}$$

(4) *Relative error due to rounding off to k digits*

$$= \left|\frac{X - X'}{X}\right| < \frac{1}{2}10^{1-k}$$

**Obs. 1.** *If a number is correct to n significant digits, then the maximum relative error* $\le \frac{1}{2}10^{-n}$.

*If a number is correct to d decimal places, then the absolute error* $\le \frac{1}{2}10^{-d}$.

**Obs. 2.** *If the first significant figure of a number is k and the number is correct to n significant figures, then the relative error* $< 1/(k \times 10^{n-1})$.

Let us verify this result by finding the relative error in the number 864.32 correct to five significant figures.

Here $k = 8$, $n = 5$ and

absolute error $\not> 0.01 \times \frac{1}{2} = 0.005$.

$$\therefore \text{ Relative error} \leq \frac{0.005}{864.32} = \frac{5}{864320} = \frac{1}{2 \times 86432} < \frac{1}{2 \times 80000} = \frac{1}{2 \times 8 \times 10^4}$$

$$< \frac{1}{8 \times 10^4} \quad i.e. \quad \frac{1}{k \times 10^{n-1}}.$$ Hence the result is verified.

**Example 1.1.** *Round off the numbers 865250 and 37.46235 to four significant figures and compute $E_a$, $E_r$, $E_p$ in each case.*

**Sol.** (*i*) Number rounded off to four significant figures = 865200

$$\therefore \quad E_a = |X - X_1| = |865250 - 865200| = 50$$

$$E_r = \left|\frac{X - X_1}{X}\right| = \frac{50}{865250} = 6.71 \times 10^{-5}$$

$$E_p = E_r \times 100 = 6.71 \times 10^{-3}$$

(*ii*) Number rounded off to four significant figures = 37.46

$$\therefore \quad E_a = |X - X_1| = |37.46235 - 37.46000| = 0.00235$$

$$E_r = \left|\frac{X - X_1}{X}\right| = \frac{0.00235}{37.46235} = 6.27 \times 10^{-5}$$

$$E_p = E_r \times 100 = 6.27 \times 10^{-3}.$$

**Example 1.2.** *Find the absolute error if the number $X = 0.00545828$ is*

(*i*) *truncated to three decimal digits.*

(*ii*) *rounded off to three decimal digits.*

**Sol.** We have $X = 0.00545828 = 0.545828 \times 10^{-2}$

(*i*) After truncating to three decimal places, its approximate value $X' = 0.545 \times 10^{-2}$

$$\therefore \text{ Absolute error} = |X - X'| = 0.000828 \times 10^{-2}$$

$$= 0.828 \times 10^{-5} < 10^{-2-3}$$

This proves rule (1).

(*ii*) After rounding off to three decimal places, its approximate value $X' = 0.546 \times 10^{-2}$

$$\therefore \text{ Absolute error} = |X - X'|$$

$$= |0.545828 - 0.546| \times 10^{-2}$$

$$= 0.000172 \times 10^{-2} = 0.172 \times 10^{-5}$$

which is $< 0.5 \times 10^{-2-3}$. This proves rule (2).

**Example 1.3.** *Find the relative error if the number $X = 0.004997$ is*

(*i*) *truncated to three decimal digits*

(*ii*) *rounded off to three decimal digits.*

**Sol.** We have $X = 0.004997 = 0.4997 \times 10^{-2}$

(*i*) After truncating to three decimal places, its approximate value $X' = 0.499 \times 10^{-2}$.

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

$$= 0.140 \times 10^{-2} < 10^{1-3}$$

This proves rule (3).

(*ii*) After rounding off to three decimal places, the approximate value of the given number

$$X' = 0.500 \times 10^{-2}$$

$$\therefore \quad \text{Relative error} = \left|\frac{X - X'}{X}\right| = \left|\frac{0.4997 \times 10^{-2} - 0.500 \times 10^{-2}}{0.4997 \times 10^{-2}}\right|$$

$$= 0.600 \times 10^{-3} = 0.06 \times 10^{-3+1}$$

which is less than $0.5 \times 10^{-3+1}$. This proves rule (4).

## PROBLEMS 1.1

1. Round off the following numbers correct to four significant figures : 3.26425, 35.46735, 4985561, 0.70035, 0.00032217, 18.265101.
2. Round off the number 75462 to four significant digits and then calculate the absolute error and percentage error. *(U.P.T.U., B. Tech., 2004)*
3. If 0.333 is the approximate value of 1/3, find the absolute and relative errors. *(Bhopal, B.E., 2007)*
4. Find the percentage error if 625.483 is approximated to three significant figures.
5. Find the relative error in taking $\pi = 3.141593$ as 22/7. *(V.T.U. MCA, 2007)*
6. The height of an observation tower was estimated to be 47 m, whereas its actual height was 45 m. Calculate the percentage relative error in the measurement.
7. Suppose that you have a task of measuring the lengths of a bridge and a rivet, and come up with 9999 and 9 cm, respectively. If the true values are 10,000 and 10 cm respectively, compute the percentage relative error in each case. *(Pune, B. Tech., 2004)*
8. Find the value of $e^x$ using series expansion $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$ for $x = 0.5$ with an absolute error less than 0.005.
9. $\sqrt{29} = 5.385$ and $\sqrt{\pi} = 3.317$ correct to 4 significant figures. Find the relative errors in their sum and difference.
10. Given : $a = 9.00 \pm 0.05$, $b = 0.0356 \pm 0.0002$, $c = 15300 \pm 100$, $d = 62000 \pm 500$. Find the maximum value of absolute error in $a + b + c + d$. *(P.T.U., B. Tech., 2001)*
11. Two numbers are 3.5 and 47.279 both of which are correct to the significant figures given. Find their product.
12. Find the absolute error and the relative error in the product of 432.8 and 0.12584 using four digit mantissa. *(Kerala B. Tech., 2003)*
13. The discharge $Q$ over a notch for head $H$ is calculated by the formula $Q = kH^{5/2}$ where $k$ is a given constant. If the head is 75 cm and an error of 0.15 cm is possible in its measurement, estimate the percentage error in computing the discharge.
14. If the number $p$ is correct to 3 significant digits, what will be the maximum relative error ?

## 1.5. ERROR PROPAGATION

A number of computational steps are carried out for the solution of a problem. It is necessary to understand the way the error propagates with progressive computation.

If the approximate values of two numbers $X$ and $Y$ be $X'$ and $Y'$ respectively, then the absolute error

$$E_{ax} = X - X' \quad \text{and} \quad E_{ay} = Y - Y'$$

**(1)** *Absolute error in addition operation*

$$X + Y = (X' + E_{ax}) + (Y' + E_{ay})$$
$$= X' + Y' + E_{ax} + E_{ay}$$

$$\therefore \quad |(X + Y) - (X' + Y')| = |E_{ax} + E_{ay}| \le |E_{ax}| + |E_{ay}|$$

Thus the absolute error in taking $(X' + Y')$ as an approximation to $(X + Y)$ is less than or equal to the sum of the absolute errors in taking $X'$ as an approximation to $X$ and $Y'$ as an approximation to $Y$.

**(2)** *Absolute error in subtraction operation*

$$X - Y = (X' + E_{ax}) - (Y' + E_{ay})$$
$$= (X' - Y') + (E_{ax} - E_{ay})$$

$$\therefore \quad |(X - Y) - (X' - Y')| = |E_{ax} - E_{ay}| \le |E_{ax}| + |E_{xy}|$$

Thus the absolute error in taking $(X' - Y')$ as an approximation to $(X - Y)$ is less than or equal to the sum of the absolute errors in taking $X'$ as an approximation to $X$ and $Y'$ as an approximation to $Y$.

**(3)** *Absolute error in multiplication operation*

To find the absolute error $E_a$ in the product of two numbers $X$ and $Y$, we write

$$E_a = (X + E_{ax})(Y + E_{ay}) - XY$$

where $E_{ax}$ and $E_{ay}$ are the absolute errors in $X$ and $Y$ respectively. Then

$$E_a = XE_{ay} + YE_{ax} + E_{ax}E_{ay}$$

Assuming $E_{ax}$ and $E_{ay}$ are reasonably small so that $E_{ax}\, E_{xy}$ can be ignored.

Thus $E_a = XE_{ay} + YE_{ax}$ approximately.

**(4)** *Absolute error in division operation*

Similarly the absolute error $E_a$ in the quotient of two numbers $X$ and $Y$ is given by

$$E_a = \frac{X + E_{ax}}{Y + E_{ay}} - \frac{X}{Y} = \frac{Y\,E_{ax} - X\,E_{ay}}{Y(Y + E_{ay})}$$

$$= \frac{yE_{ax} - XE_{ay}}{Y^2\,(1 + E_{ay}/Y)}$$

$$= \frac{YE_{ax} - XE_{ay}}{Y^2}, \text{ assuming } E_{ay/y} \text{ to be small.}$$

$$= \frac{X}{Y}\left(\frac{E_{ax}}{X} - \frac{E_{ay}}{Y}\right)$$

**Example 1.4.** *Find the absolute error and relative error in $\sqrt{6}+\sqrt{7}+\sqrt{8}$ correct to 4 significant digits.*

**Sol.** We have $\sqrt{6} = 2.449$, $\sqrt{7} = 2.646$, $\sqrt{8} = 2.828$

$$\therefore \qquad S = \sqrt{6}+\sqrt{7}+\sqrt{8} = 7.923.$$

Then the absolute error $E_a$ in $S$, is

$$E_a = 0.0005 + 0.0007 + 0.0004 = 0.0016$$

This shows that $S$ is correct to 3 significant digits only. Therefore, we take $S = 7.92$

Then the relative error $E_r$ is

$$E_r = \frac{0.0016}{7.92} = 0.0002.$$

**Example 1.5.** *The area of cross-section of a rod is desired upto 0.2% error. How accurately should the diameter be measured?* *(Pune, B. Tech., 2003)*

**Sol.** If $A$ is the area and $D$ is the diameter of the rod, then $A = \pi\left(\frac{D}{2}\right)^2 = \frac{\pi}{4} D \,.\, D$.

Now error in area $A$ is 0.2% *i.e.*, 0.002 which is due to the error in the product $D \times D$.

We know that if $E_a$ is the absolute error in the product of two numbers $X$ and $Y$, then

$$E_a = X_{aY}E + YE_{aX}$$

Here $X = Y = D$ and $E_{aX} = E_{aY} = E_D$, therefore

$$E_a = DE_D + DE_D \quad \text{or} \quad 0.002 = 2DE_D$$

Thus $E_D = 0.001/D$ *i.e.*, the error in the diameter should not exceed $0.001\ D^{-1}$.

**Example 1.6.** *Find the product of the numbers 3.7 and 52.378 both of which are correct to given significant digits.*

**Sol.** Since the absolute error is greatest in 3.7, therefore we round off the other number to 3 significant figures *i.e.* 52.4.

$\therefore$ Their product $P = 3.7 \times 52.4 = 193.88 = 1.9388 \times 10^2$

Since the first number contains only two significant figures, therefore retaining only two significant figures in the product, we get

$$P = 1.9 \times 10^2$$

## 1.6. ERROR IN THE APPROXIMATION OF A FUNCTION

Let $y = f(x_1, x_2)$ be a function of two variables $x_1, x_2$. If $\delta x_1, \delta x_2$ be the errors in $x_1, x_2$, then the error $\delta y$ in $y$ is given by

$$y + \delta y = f(x_1 + \delta x_1, x_2 + \delta x_2)$$

Expanding the right hand side by Taylor's series, we get

$$y + \delta y = f(x_1, x_2) + \left(\frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2\right) \qquad ...(i)$$

+ terms involving higher powers of $\delta x_1$ and $\delta x_2$

If the errors $\delta x_1, \delta x_2$ be so small that their squares and higher powers can be neglected, then (*i*) gives

$$\delta y = \frac{\partial f}{\partial x_1}\delta x_1 + \frac{\partial f}{\partial x_2}\delta x_2 \text{ approximately.}$$

Hence $\delta y \approx \frac{\partial y}{\partial x_1}\delta x_1 + \frac{\partial y}{\partial x_2}\delta x_2$

In general, the error $\delta y$ in the function $y = f(x_1, x_2, \ldots\ldots, x_n)$ corresponding to the errors $\delta x_i$ in $x_i$ $(i = 1, 2, \ldots\ldots n)$ is given by

$$\delta y \approx \frac{\partial y}{\partial x_1}\delta x_1 + \frac{\partial y}{\partial x_2}\delta x_2 + \ldots\ldots + \frac{\partial y}{\partial x_n}\delta x_n$$

and the relative error in $y$ is $E_r = \frac{\delta y}{y} \approx \frac{\partial y}{\partial x_1}\frac{\delta x_1}{y} + \frac{\partial y}{\partial x_2}\frac{\delta x_2}{y} + \ldots.. + \frac{\partial y}{\partial x_n}\frac{\delta x_n}{y}$.

**Example 1.7.** *If $u = 4x^2y^3/z^4$ and errors in x, y, z be 0.001, compute the relative maximum error in u when $x = y = z = 1$.*

**Sol.** Since $\frac{\partial u}{\partial x} = \frac{8xy^3}{z^4}, \frac{\partial u}{\partial y} = \frac{12x^2y^2}{z^4}, \frac{\partial u}{\partial z} = -\frac{16x^2y^3}{z^5}$

$$\therefore \quad \delta u = \frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y + \frac{\partial u}{\partial z}\delta z = \frac{8xy^3}{z^4}\delta x + \frac{12x^2y^2}{z^4}\delta y - \frac{16x^2y^3}{z^5}\delta z$$

Since the errors $\delta x$, $\delta y$, $\delta z$ may be positive or negative, we take the absolute values of the terms on the right side, giving

$$(\delta u)_{max} \approx \left|\frac{8xy^3}{z^4}\delta x\right| + \left|\frac{12x^2y^2}{z^4}\delta y\right| + \left|\frac{16x^2y^3}{z^5}\delta z\right|$$

$$= 8(0.001) + 12(0.001) + 16(0.001) = 0.036$$

Hence the maximum relative error $= (\delta u)_{max}/u = 0.036/4 = 0.009$.

**Example 1.8.** *Find the relative error in the function $y = ax_1^{m_1} x_2^{m_2} \ldots\ldots x_n^{m_n}$.*

**Sol.** We have $\log y = \log a + m_1 \log x_1 + m_2 \log x_2 + \ldots + m_n \log x_n$

$$\therefore \quad \frac{1}{y}\frac{\partial y}{\partial x_1} = \frac{m_1}{x_1}, \frac{1}{y}\frac{\partial y}{\partial x_2} = \frac{m_2}{x_2} \text{ etc.}$$

Hence $E_r \approx \frac{\partial y}{\partial x_1}\frac{\delta x_1}{y} + \frac{\partial y}{\partial x_2}\frac{\delta x_2}{y} + \ldots. + \frac{\partial y}{\partial x_n}\frac{\delta x_n}{y} = m_1\frac{\delta x_1}{x_1} + m_2\frac{\delta x_2}{x_2} + \ldots + m_n\frac{\delta x_n}{x_n}$

Since the errors $\delta x_1, \delta x_2, \ldots\ldots, \delta x_n$ may be positive or negative, we take the absolute values of the terms on the right side. This gives :

$$(E_r)_{max} \le m_1\left|\frac{\delta x_1}{x_1}\right| + m_2\left|\frac{\delta x_2}{x_2}\right| + \ldots + m_n\left|\frac{\delta x_n}{x_n}\right|$$

**Cor.** Taking $a = 1, m_1 = m_2 = \ldots\ldots = m_n = 1$, we have

$$y = x_1x_2 \ldots\ldots x_n.$$

then $E_r = \frac{\delta x_1}{x_1} + \frac{\delta x_2}{x_2} + \ldots\ldots + \frac{\delta x_n}{x_n}$

Thus *the relative error of a product of n numbers is approximately equal to the algebraic sum of their relative errors.*

## 1.7. ERROR IN A SERIES APPROXIMATION

We know that the Taylor's series for $f(x)$ at $x = a$ with a remainder after $n$ terms is

$$f(x) = f(a + \overline{x - a}) = f(a) + (x - a)f'(a) + \frac{(x-a)^2}{2!} f''(a) + \ldots + \frac{(x-a)^{n-1}}{(n-1)!} f^{n-1}(a) + R_n(x)$$

where $R_n(x) = \dfrac{(x-a)^n}{n!} f^n(\theta),\ a < \theta < x.$

If the series is convergent, $R_n(x) \to 0$ as $n \to \infty$ and hence if $f(x)$ is approximated by the first $n$ terms of this series, then the maximum error will be given by the remainder term $R_n(x)$. On the other hand, if the accuracy required in a series approximation is preassigned, then we can find $n$, the number of terms which would yield the desired accuracy.

**Example 1.9.** *Find the number of terms of the exponential series such that their sum gives the value of $e^x$ correct to six decimal places at $x = 1$.*

**Sol.** We have $\quad e^x = 1 + x + \dfrac{x^2}{2!} + \dfrac{x^3}{3!} + \ldots\ldots + \dfrac{x^{n-1}}{(n-1)!} + R_n(x) \qquad \ldots(i)$

where $\quad R_n(x) = \dfrac{x^n}{n!} e^\theta,\ 0 < \theta < x.$

$\therefore$ Maximum absolute error (at $\theta = x$) $= \dfrac{x^n}{n!} e^x$ and the maximum relative error $= \dfrac{x^n}{n!}$

Hence $(E_r)_{max}$ at $x = 1$ is $\dfrac{1}{n!}$.

For a six decimal accuracy at $x = 1$, we have

$$\frac{1}{n!} < \frac{1}{2} 10^{-6}, \quad i.e. \quad n! > 2 \times 10^6 \qquad \text{which gives } n = 10.$$

Thus we need 10 terms of the series $(i)$ in order that its sum is correct to 6 decimal places.

**Example 1.10.** *The function $f(x) = \tan^{-1} x$ can be expanded as*

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots\ldots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + \ldots\ldots,$$

*find $n$ such that the series determine $\tan^{-1} x$ correct to eight significant digits at $x = 1$.*

*(U.P.T.U., B. Tech. 2007)*

**Sol.** If we retain $n$ terms in the expansion of $\tan^{-1} x$, then $(n + 1)$th term

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

$$= \frac{(-1)^n}{2n+1} \quad \text{for } x = 1.$$

To determine $\tan^{-1}(1)$ correct to eight significant digits accuracy $\left|\dfrac{(-1)^n}{2n+1}\right| < \dfrac{1}{2} \times 10^{-8}$

*i.e.,*
$$2n + 1 > 2 \times 10^8 \quad \text{or} \quad n > 10^8 - \frac{1}{2}$$

Hence $n = 10^8 + 1$.

## 1.8. ORDER OF APPROXIMATION

We often replace a function $f(h)$ with its approximation $\phi(h)$ and the error bound is known to be $\mu(h^n)$, $n$ being a positive integer so that

$$| f(h) - \phi(h) | \le \mu \, | h^n | \qquad \text{for sufficiently small } h.$$

Then we say that $\phi(h)$ *approximates* $f(h)$ *with order of approximation* $O(h^n)$ and write $f(h) = \phi(h) + O(h^n)$.

For instance,
$$\frac{1}{1-h} = 1 + h + h^2 + h^3 + h^4 + h^5 + \ldots\ldots$$

is written as
$$\frac{1}{1-h} = 1 + h + h^2 + h^3 + O(h^4) \qquad \ldots(i)$$

to the 4th order of approximation.

Similarly
$$\cos(h) = 1 - \frac{h^2}{2!} + \frac{h^4}{4!} - \frac{h^6}{6!} + \frac{h^8}{8!} - \ldots\ldots$$

to the 6th order of approximation becomes

$$\cos(h) = 1 - \frac{h^2}{2!} + \frac{h^4}{4!} + O(h^6) \qquad \ldots(ii)$$

The sum of $(i)$ and $(ii)$ gives

$$(1-h)^{-1} + \cos(h) = 2 + h + \frac{h^2}{2!} + h^3 + O(h)^4 + \frac{h^4}{4!}\, O(h^6) \qquad \ldots(iii)$$

Since
$$O(h^4) + \frac{h^4}{4!} = O(h^4) \text{ and } O(h^4) + O(h^6) = O(h^4)$$

$\therefore$ $(iii)$ takes the form $(1-h)^{-1} + \cos(h) = 2 + h + \dfrac{h^2}{2} + h^3 + O(h^4)$, which is of the 4th order of approximation.

Similarly the product of $(i)$ and $(ii)$ yields

$$(1-h)^{-1}\cos(h) = (1 + h + h^2 + h^3)\left(1 - \frac{h^2}{2!} + \frac{h^4}{4!}\right) + (1 + h + h^2 + h^3)\, O(h^6)$$
$$+ \left(1 - \frac{h^2}{2!} + \frac{h^4}{4!}\right) O(h^4) + O(h^4)\, O(h^6)$$
$$= 1 + h + \frac{h^2}{2} + \frac{h^3}{2} - \frac{11h^4}{24} + \frac{11}{24}h^5 + \frac{h^6}{24} + \frac{h^7}{24} + O(h^4) + O(h^6) + O(h^4)\, O(h^6) \qquad \ldots(iv)$$

Since $O(h^4)\,O(h^6) = O(h^{10})$

and $$-\frac{11h^4}{24} + \frac{11}{24}h^5 + \frac{h^6}{24} + \frac{h^7}{24} + O(h^4) + O(h^6) + O(h^{10}) = O(h^4)$$

$\therefore$ (iv) is reduced to $(1-h)^{-1}\cos(h) = 1 + h + \dfrac{h^2}{2} + \dfrac{h^3}{2} + O(h^4)$, which is of the 4th order of approximation.

## 1.9. GROWTH OF ERROR

Let $e(n)$ represent the growth of error after $n$ steps of a computation process.

*If* $|\,e(n)\,| \sim n\,\varepsilon$, *we say that the growth of error is* **linear.**

*If* $|\,e(n)\,| \sim \delta^n\,\varepsilon$, *we say that the growth of error is* **exponential**.

If $\delta > 1$, the exponential error grows indefinitely as $n \to \infty$, and

if $0 < \delta < 1$, the exponential error decreases to zero as $n \to \infty$.

## PROBLEMS 1.2

1. Find the smaller root of the equation $x^2 - 400x + 1 = 0$, correct to 4 decimal places.
2. If $r = h(4h^5 - 5)$, find the percentage error in $r$ at $h = 1$, if the error in $h$ is 0.04. (*W.B.T.U., 2005*)
3. If $R = 10\,x^3y^2z^2$ and errors in $x, y, z$ are 0.03, 0.01, 0.02 respectively at $x = 3, y = 1, z = 2$. Calculate the absolute error and % relative error in evaluating $R$?
4. If $R = 4xy^2/z^3$ and errors in $x, y, z$ be 0.001, show that the maximum relative error at $x = y = z = 1$ is 0.006.
5. If $V = \dfrac{1}{2}\left(\dfrac{r^2}{h} + h\right)$ and the error in $V$ is at the most 0.4%, find the percentage error allowable in $r$ and $h$ when $r = 5.1$ cm and $h = 5.8$ cm.
6. Find the value of $I = \displaystyle\int_0^{0.8} \frac{\sin x}{x}\,dx$ correct to 4 decimal places.
7. Using the series $\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \ldots\ldots$, evaluate sin 25° with an accuracy of 0.001.
8. Determine the number of terms required in the series for $\log(1 + x)$ to evaluate log 1.2 correct to six decimal places.
9. Use the series $\log_e\left(\dfrac{1+x}{1-x}\right) = 2\left(x + \dfrac{x^3}{3} + \dfrac{x^5}{5} + \ldots\ldots\right)$ to compute the value of log (1.2) correct to seven decimal places and find the number of terms retained. (*U.P.T.U., B.Tech., 2003*)
10. Find the order of approximation for the sum and product of the following expansions:

$$e^h = 1 + h + \frac{h^2}{2} + \frac{h^3}{3!} + O(h^4) \quad \text{and} \quad \cos(h) = 1 - \frac{h^2}{2!} + \frac{h^4}{4!} + O(h^6).$$

**11.** Given the expansions :

$$\sin(t) = t - \frac{t^3}{3!} + \frac{t^5}{5!} + O(t^7) \quad \text{and} \quad \cos(t) = 1 - \frac{t^2}{2!} + \frac{t^3}{4!} + O(t^6)$$

Determine the order of approximation for their sum and product.

## 1.10. OBJECTIVE TYPE OF QUESTIONS

### PROBLEMS 1.3

*Select the correct answer or fill up the blanks in the following questions :*

**1.** If $x$ is the true value of a quantity and $x_1$ is its approximate value, then the relative error is

(a) $|x_1 - x|/x_1$ (b) $|x - x_1|/x$

(c) $|x_1/x|$ (d) $x/|x_1 - x|$.

**2.** The relative error in the number 834.12 correct to five significant figures is ...... .

**3.** If a number is rounded to $k$ decimal places, then the absolute error is

(a) $\frac{1}{2}\,10^{k-1}$ (b) $\frac{1}{2}\,10^{-k}$

(c) $\frac{1}{3}\,10^{k}$ (d) $\frac{1}{4}\,10^{-k}$.

**4.** If $\pi$ is taken = 3.14 in place of 3.14156, then the relative error is ......

**5.** Given $x = 1.2$, $y = 25.6$ and $z = 4.5$, then the relative error in evaluating $w = x^2 + y/z$ is ...... .

**6.** Round off values of 43.38256, 0.0326457 and 0.2537623 to four significant digits are ...... .

**7.** Round relative maximum error in $3x^2y/z$ when $\delta x = \delta y = \delta z = 0.001$ at $x = y = z = 1$ is ...... .

**8.** If both the digits of the number 8.6 are correct, then the relative error is ...... .

**9.** If a number is correct to $n$ significant digits, then the relative error is

(a) $\frac{1}{2}\,10^{n}$ (b) $\frac{1}{2}\,10^{n-1}$

(c) $\leq \frac{1}{2}\,10^{-n}$ (d) $< \frac{1}{2}\,10^{n-1}$.

**10.** If $(\sqrt{3} + \sqrt{5} + \sqrt{7})$ is rounded to four significant digits, then the absolute error is ...... .

**11.** $(\sqrt{102} - \sqrt{101})$ correct to three significant figures is ...... .

**12.** Approximate values of 1/3 are given as 0.3, 0.33 and 0.34. Out of these the best approximation is ...... .

**13.** The relative error if $\frac{2}{3}$ is approximated to 0.667, is ...... .

**14.** If the first significant digit of a number is $p$ and the number is correct to $n$ significant digits, then the relative error is ...... . *(U.P.T.C., MCA, 2009)*

# 2

# SOLUTION OF ALGEBRAIC AND TRANSCENDENTAL EQUATIONS



## 2.1. INTRODUCTION

An expression of the form $f(x) = a_0 x^n + a_1 x^{n-1} + \ldots\ldots + a_{n-1} x + a_n$ where $a$'s are constants ($a_0 \neq 0$) and $n$ is a positive integer, is called a *polynomial* in $x$ of degree $n$. The polynomial $f(x) = 0$ is called an *algebraic equation* of degree $n$. If $f(x)$ contains some other functions such as trigonometric, logarithmic, exponential etc., then $f(x) = 0$ is called a *transcendental equation*.

**Def.** *The value $\alpha$ of $x$ which satisfies $f(x) = 0$* ...(1)
*is called a* **root** *of $f(x) = 0$*. Geometrically, a root of (1) is that value of $x$ where the graph of $y = f(x)$ crosses the $x$-axis.

The process of finding the roots of an equation is known as the *solution of that equation*. This is a problem of basic importance in applied mathematics.

If $f(x)$ is a quadratic, cubic or a biquadratic expression, algebraic solutions of equations are available. But the need often arises to solve higher degree or transcendental equations for which no direct methods exist. Such equations can best be solved by approximate methods. In this chapter, we shall discuss some numerical methods for the solution of algebraic and transcendental equations.

## 2.2. BASIC PROPERTIES OF EQUATIONS

*I. If $f(x)$ is exactly divisible by $x - \alpha$, then $\alpha$ is a root of $f(x) = 0$.*

*II. Every equation of the nth degree has only n roots* (real or imaginary).

Conversely if $\alpha_1, \alpha_2, \ldots\ldots, \alpha_n$ be the roots of the $n$th degree equation $f(x) = 0$, then

$$f(x) = A(x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n)$$

where $A$ is a constant.

**Obs.** If a polynomial of degree $n$ vanishes for more than $n$ values of $x$, it must be identically zero.

**Example 2.1.** *Solve the equation $2x^3 + x^2 - 13x + 6 = 0$.*

**Sol.** By inspection, we find $x = 2$ satisfies the given equation.

$\therefore$ 2 is its root, *i.e.* $x - 2$ is a factor of $2x^3 + x^2 - 13x + 6$.

Dividing this polynomial by $x - 2$, we get the quotient $2x^2 + 5x - 3$ and remainder 0.

Equating this quotient to zero, we get $2x^2 + 5x - 3 = 0$.

Solving this quadratic, we get

$$x = \frac{-5 \pm \sqrt{[5^2 - 4.2.(-3)]}}{2.2} = -3,\ 1/2.$$

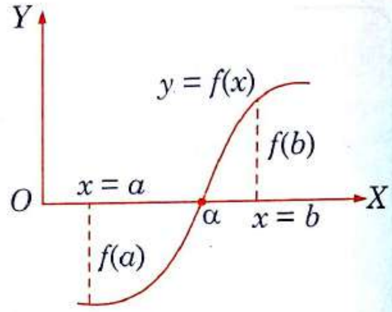


Fig. 2.1

Hence the roots of the given equation are 2, – 3, 1/2.

*III.* **Intermediate value property.** *If $f(x)$ is continuous in the interval $[a, b]$ and $f(a)$, $f(b)$ have different signs, then the equation $f(x) = 0$ has atleast one root between $x = a$ and $x = b$.*

Since $f(x)$ is continuous between $a$ and $b$, so while $x$ changes from $a$ to $b$, $f(x)$ must pass through all the values from $f(a)$ to $f(b)$ [Fig. 2.1]. But one of these quantities $f(a)$ or $f(b)$ is positive and the other negative, it follows that atleast for one value of $x$ (say $\alpha$) lying between $a$ and $b$, $f(x)$ must be zero. Then $\alpha$ is the required root.

*IV. In an equation with real coefficients, imaginary roots occur in conjugate pairs, i.e.* if $\alpha + i\beta$ is a root of the equation $f(x) = 0$, then $\alpha - i\beta$ must also be its root.

Similarly if $a + \sqrt{b}$ is an irrational root of an equation, then $a - \sqrt{b}$ must also be its root.

**Obs.** *Every equation of the odd degree has atleast one real root.*

This follows from the fact that imaginary roots occur in conjugate pairs.

**Example 2.2.** *Solve the equation $3x^3 - 4x^2 + x + 88 = 0$, one root being $2 + \sqrt{7}i$.*

**Sol.** Since one root is $2 + \sqrt{7i}$, the other root must be $2 - \sqrt{7i}$.

$\therefore$ The factors corresponding to these roots are $(x - 2 - \sqrt{7i})$ and $(x - 2 + \sqrt{7i})$

or $$(x - 2 - \sqrt{7i})(x - 2 + \sqrt{7i}) = (x - 2)^2 + 7 = x^2 - 4x + 11 \quad \ldots(i)$$

is a divisor of $3x^3 - 4x^2 + x + 88$

$\therefore$ Division of (*i*) by $x^2 - 4x + 11$ gives $3x + 8$ as the quotient.

Thus the depressed equation is $3x + 8 = 0$. Its root is $-8/3$.

Hence the roots of the given equation are $2 \pm \sqrt{7i}$, $-8/3$.

V. **Descarte's rule of signs.** *The equation $f(x) = 0$ cannot have more positive roots than the changes of signs in $f(x)$ ; and more negative roots than the changes of signs in $f(-x)$.*

For instance, consider the equation $f(x) = 2x^7 - x^5 + 4x^3 - 5 = 0$ ...(*i*)

Signs of $f(x)$ are $\quad + \quad - \quad + \quad -$

Clearly $f(x)$ has 3 changes of signs (from + to − or − to +).

Thus (*i*) cannot have more than 3 positive roots.

Also
$$f(-x) = 2(-x)^7 - (-x)^5 + 4(-x)^3 - 5$$
$$= -2x^7 + x^5 - 4x^3 - 5$$

This shows that $f(x)$ has 2 changes of signs.

Thus (*i*) cannot have more than 2 negative roots.

**Obs. Existence of imaginary roots.** *If an equation of the nth degree has at the most p positive roots and at the most q negative roots, then it follows that the equation has at least $n - (p + q)$ imaginary roots.*

Evidently (*i*) above is an equation of the 7th degree and has at the most 3 positive roots and 2 negative roots. Thus (*i*) has atleast 2 imaginary roots.

VI. **Relations between roots and coefficients.** *If $\alpha_1, \alpha_2, \alpha_3, \ldots\ldots, \alpha_n$ be the roots of the equation*

$$a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots\ldots + a_{n-1}x + a_n = 0 \qquad \text{...(1)}$$

*then*
$$\Sigma\alpha_1 = -\frac{a_1}{a_0},$$
$$\Sigma\alpha_1\alpha_2 = \frac{a_2}{a_0},$$
$$\Sigma\alpha_1\alpha_2\alpha_3 = -\frac{a_3}{a_0}, \ldots\ldots$$
$$\alpha_1\alpha_2\alpha_3 \ldots\ldots \alpha_n = (-1)^n \frac{a_n}{a_0}.$$

**Example 2.3.** *Solve the equation $x^3 - 7x^2 + 36 = 0$, given that one root is double of another.*

**Sol.** Let the roots be $\alpha, \beta, \gamma$ such that $\beta = 2\alpha$.

Also $\alpha + \beta + \gamma = 7, \quad \alpha\beta + \beta\gamma + \gamma\alpha = 0, \quad \alpha\beta\gamma = -36$

$\therefore$
$$3\alpha + \gamma = 7 \qquad \text{...}(i)$$
$$2\alpha^2 + 3\alpha\gamma = 0 \qquad \text{...}(ii)$$
$$2\alpha^2\gamma = -36 \qquad \text{...}(iii)$$

Solving (*i*) and (*ii*), we get $\alpha = 2, \gamma = -2$.

[The values $\alpha = 0, \gamma = 7$ are inadmissible, as they do not satisfy (*iii*)].

Hence the roots are 3, 6 and − 2.

**Example 2.4.** *Solve the equation $x^4 - 2x^3 + 4x^2 + 6x - 21 = 0$, given that the sum of two its roots is zero.* (*Madras B.E., 2003*)

**Sol.** Let the roots be $\alpha, \beta, \gamma, \delta$ such that $\alpha + \beta = 0$.

Also $\alpha + \beta + \gamma + \delta = 2, \quad \therefore \quad \gamma + \delta = 2$

Thus the quadratic factor corresponding to $\alpha$, $\beta$ is of the form $x^2 - 0x + p$ and that corresponding to $\gamma$, $\delta$ is of the form of $x^2 - 2x + q$.

$\therefore$ $$x^4 - 2x^3 + 4x^2 + 6x - 21 = (x^2 + p)(x^2 - 2x + q) \quad ...(i)$$

Equating coefficients of $x^2$ and $x$ from both sides of $(i)$, we get

$$4 = p + q, \qquad 6 = -2p.$$

$\therefore$ $$p = -3, \qquad q = 7.$$

Hence the given equation is equivalent to

$$(x^2 - 3)(x^2 - 2x + 7) = 0$$

$\therefore$ The roots are $x = \pm\sqrt{3}, 1 \pm i\sqrt{6}$.

**Example 2.5.** *Find the condition that the cubic $x^3 - lx^2 + mx - n = 0$ should have its roots in*

*(a) Arithmetical progression* *(b) Geometrical progression.*

**Sol.** (a) Let the roots be $a - d$, $a$, $a + d$ so that the sum of the roots $= 3a = l$ *i.e.* $a = l/3$.

Since $a$ is the root of the given equation $a^3 - la^2 + ma - n = 0$

Substituting $a = l/3$, we get $2l^3 - 9lm + 27n = 0$ which is the required condition.

(b) Let the roots be $a/r$, $a$, $ar$, then the product of the roots $= a^3 = n$.

Since $a$ is a root of the given equation

$\therefore$ $$a^3 - la^2 + ma - n = 0$$

Putting $a = (n)^{1/3}$, we get $n - ln^{2/3} + mn^{1/3} - n = 0$ or $m = ln^{1/3}$

Cubing both sides, we get $m^3 = l^3 n$

which is the required condition.

**Example 2.6.** *If $\alpha$, $\beta$, $\gamma$ be the roots of the equation $x^3 + px + q = 0$, find the value of (a) $\Sigma\alpha^2\beta$, (b) $\Sigma\alpha^4$.*

**Sol.** We have $\alpha + \beta + \gamma = 0 \quad ...(i)$

$$\alpha\beta + \beta\gamma + \gamma\alpha = p \quad ...(ii)$$

$$\alpha\beta\gamma = -q \quad ...(iii)$$

(a) Multiplying $(i)$ and $(ii)$, we get

$$\alpha^2\beta + \alpha^2\gamma + \beta^2\gamma + \beta^2\alpha + \gamma^2\alpha + \gamma^2\beta + 3\alpha\beta\gamma = 0$$

or $$\Sigma\alpha^2\beta = -3\alpha\beta\gamma = 3q \quad \text{[by } (iii)]$$

(b) Multiplying the given equation by $x$, we get

$$x^4 + px^2 + qx = 0$$

Putting $x = \alpha, \beta, \gamma$ successively and adding, we get

$$\Sigma\alpha^4 + p\Sigma\alpha^2 + q\Sigma\alpha = 0 \quad \text{or} \quad \Sigma\alpha^4 = -p\Sigma\alpha^2 - q(0) \quad ...(iv)$$

Now squaring $(i)$, we get

$$\alpha^2 + \beta^2 + \gamma^2 + 2(\alpha\beta + \beta\gamma + \gamma\alpha) = 0 \quad \text{or} \quad \Sigma\alpha^2 = -2p \quad \text{[by } (ii)]$$

Hence, substituting the value of $\Sigma\alpha^2$ in $(iv)$, we obtain

$$\Sigma\alpha^4 = -p(-2p) = 2p^2.$$

## PROBLEMS 2.1

1. Form the equation of the fourth degree whose roots are $3 + i$ and $\sqrt{7}$. *(Madras, B. Tech., 2000 S)*
2. Solve the equation :
   (*i*) $x^3 + 6x + 20 = 0$, one root being $1 + 3i$.
   (*ii*) $x^4 - 2x^3 - 22x^2 + 62x - 15 = 0$ given that $2 + \sqrt{3}$ is a root.
3. Show that $x^7 - 3x^4 + 2x^3 - 1 = 0$ has atleast four imaginary roots. *(Cochin, B. Tech., 2005)*
4. Show that $x^6 - x^5 - 10x + 7 = 0$ has two positive and four imaginary roots.
5. Find the number and position of real roots of $x^4 + 4x^3 - 4x - 13 = 0$.
6. The equation $x^4 - 4x^3 + ax^2 + 4x + b = 0$ has two pairs of equal roots. Find the values of $a$ and $b$.

*Solve the equations (7—13) :*

7. $2x^4 - 3x^3 - 9x^2 + 15x - 5 = 0$, given that the sum of two of its roots is zero.
8. $x^3 - 8x^2 + 9x + 18 = 0$ given that two of its roots are in the ratio 1 : 2.
9. $x^3 - 4x^2 - 20x + 48 = 0$ given that the roots $\alpha$ and $\beta$ are connected by the relation $\alpha + 2\beta = 0$. *(S.V.T.U., B. Tech., 2007)*
10. $x^4 - 8x^3 + 21x^2 - 20x + 5 = 0$ given that the sum of two of the roots is equal to the sum of the other two.
11. $x^3 - 12x^2 + 39x - 28 = 0$, roots being in arithmetical progression. *(Madras B.E., 2001)*
12. $O, A, B, C$ are the four points on a straight line such that the distances of $A, B, C$ from $O$ are the roots of equation $ax^3 + 3bx^2 + 3cx + d = 0$. If $B$ is the middle point of $AC$, show that $a^2d - 3abc + 2b^3 = 0$. *(S.V.T.U., B. Tech., 2006)*
13. If $\alpha, \beta, \gamma$ are the roots of the equation $x^3 + 4x - 3 = 0$, find the value of $\alpha^{-1} + \beta^{-1} + \gamma^{-1}$.
14. If $\alpha, \beta, \gamma$ be the roots of the equation $x^3 - lx^2 + mx - n = 0$, find the value of (*i*) $\Sigma\alpha^2\beta^2$ (*ii*) $(\beta + \gamma)(\gamma + \alpha)(\alpha + \beta)$.

## 2.3. TRANSFORMATION OF EQUATIONS

**(1) To find an equation whose roots are m times the roots of the given equation,** *multiply the second term by m, third term by $m^2$ and so on* (all missing terms supplied with zero coefficients).

For instance, let the given equation be

$$3x^4 + 6x^3 + 4x^2 - 8x + 11 = 0 \qquad ...(i)$$

To multiply its roots by $m$, put $y = mx$ (or $x = y/m$) in (*i*). Then

$$3(y/m)^4 + 6(y/m)^3 + 4(y/m)^2 - 8(y/m) + 11 = 0$$

or multiplying by $m^4$, we get

$$3y^4 + m(6y^3) + m^2(4y^2) - m^3(y) + m^4(11) = 0$$

This is same as multiplying the second term by $m$, third term by $m^2$ and so on in (*i*).

**Cor. To find an equation whose roots are with opposite signs to those of the given equation,** *change the signs of every alternative term of the given equation beginning with the second.*

Changing the signs of roots of $(i)$ is same as multiplying its roots by $-1$.

$\therefore$ The required equation will be

$$3x^4 + (-1)6x^3 + (-1)^2 4x^2 - (-1)^3 8x + (-1)^4 11 = 0$$

or
$$3x^4 - 6x^3 + 4x^3 + 8x + 11 = 0$$

which is $(i)$ with signs of every alternate term changed beginning with the second.

**(2) To find an equation whose roots are reciprocal of the roots of the given equation,** *change $x$ to $1/x$.*

**Example 2.7.** *Solve $6x^3 - 11x^2 - 3x + 2 = 0$, given that its roots are in harmonic progression.*

**Sol.** Since the roots of the given equation are in H.P., the roots of the equation having reciprocal roots will be in A.P. $\therefore$ The equation with reciprocal roots is

$$6(1/x)^3 - 11(1/x)^2 - 3(1/x) + 2 = 0$$

or
$$2x^3 - 3x^2 - 11x + 6 = 0 \qquad ...(i)$$

Since the roots of the given equation are in H.P., therefore, the roots of $(i)$ are in A.P. Let the roots be $a - d$, $a$, $a + d$. Then $3a = 3/2$ and $a(a^2 - d^2) = -3$.

Solving these equations, we get $a = 1/2$, $d = 5/2$. Thus the roots of $(i)$ are $-2$, $1/2$, $3$.

Hence the roots of the given equation are $-1/2$, $2$, $1/3$.

**Example 2.8.** *If $\alpha, \beta, \gamma$ be the roots of the cubic $x^3 - px^2 + qx - r = 0$, form the equation whose roots are $\beta\gamma + 1/\alpha$, $\gamma\alpha + 1/\beta$, $\alpha\beta + 1/\gamma$.*

**Sol.** If $x$ is a root of the given equation and $y$, a root of the required equation, then

$$y = \beta\gamma + \frac{1}{\alpha} = \frac{\alpha\beta\gamma + 1}{\alpha} = \frac{r+1}{\alpha} = \frac{r+1}{x} \qquad (\because \ \alpha\beta\gamma = r)$$

Thus $x = (r + 1)/y$.

Substituting this value of $x$ in the given equation, we get

$$\left(\frac{r+1}{y}\right)^3 - p\left(\frac{r+1}{y}\right)^2 + q\left(\frac{r+1}{y}\right) - r = 0$$

or
$$ry^3 - q(r+1)y^2 + p(r+1)^2 y - (r+1)^3 = 0$$

which is the required equation.

**(3) Reciprocal equations.** *If an equation remains unaltered on changing $x$ to be $1/x$, it is called a* **reciprocal equation.**

Such equations are of the following types :

*(i) A reciprocal equation of an odd degree having coefficients of terms equidistant from the beginning and end equal.* It has a root $= -1$.

*(ii) A reciprocal equation of an odd degree having coefficients of terms equidistant from the beginning and end equal but opposite in sign.* It has a root $= 1$.

*(iii) A reciprocal equation of an even degree having coefficients of terms equidistant from the beginning and end equal but opposite in sign.* Such an equation has two roots $= 1$ and $-1$.

**The substitution $x + 1/x = y$ reduces the degree of the equation to half its former degree.**

**Example 2.9.** *Solve :* (*i*) $6x^5 - 41x^4 + 97x^3 - 97x^2 + 41x - 6 = 0$

(*Coimbatore, B. Tech., 2001*)

(*ii*) $6x^6 - 25x^5 + 31x^4 - 31x^2 + 25x - 6 = 0$. (*Madras B.E., 2003*)

**Sol.** (*i*) This is a reciprocal equation of odd degree with opposite signs.

$\therefore$ $x = 1$ is a root.

Dividing L.H.S. by $x - 1$, the equation reduces to

$$6x^4 - 35x^3 + 62x^2 - 35x + 6 = 0$$

Dividing by $x^2$, we have $6\left(x^2 + \frac{1}{x^2}\right) - 35\left(x + \frac{1}{x}\right) + 62 = 0$

Putting $x + \frac{1}{x} = y$ and $x^2 + \frac{1}{x^2} = y^2 - 2$, we get

$$6(y^2 - 2) - 35y + 62 = 0 \quad \text{or} \quad 6y^2 - 35y + 50 = 0$$

or $(3y - 10)(2y - 5) = 0$ $\quad\therefore\quad x + \frac{1}{x} = \frac{10}{3}$ or $\frac{5}{2}$

*i.e.* $3x^2 - 10x + 3 = 0$ or $2x^2 - 5x + 2 = 0$

*i.e.* $(3x - 1)(x - 3) = 0$ or $(2x - 1)(x - 2) = 0$

$\therefore$ $x = \frac{1}{3}, 3$ or $x = \frac{1}{2}, 2$

Hence the roots are $1, \frac{1}{3}, 3, \frac{1}{2}, 2$.

(*ii*) This is a reciprocal equation of even degree with opposite signs. $\therefore$ $x = 1, -1$ are its roots.

Dividing L.H.S. by $x - 1$ and $x + 1$, the given equation reduces to

$$6x^4 - 25x^3 + 37x^2 - 25x + 6 = 0.$$

Dividing by $x^2$, we get $6\left(x^2 + \frac{1}{x^2}\right) - 25\left(x + \frac{1}{x}\right) + 37 = 0$

Putting $x + \frac{1}{x} = y$ and $x^2 + \frac{1}{x^2} = y^2 - 2$, it becomes

$$6y^2 - 25y + 25 = 0 \quad \text{or} \quad (2y - 5)(3y - 5) = 0$$

$\therefore$ $x + \frac{1}{x} = y = \frac{5}{2}$ or $\frac{5}{3}$

*i.e.* $2x^2 - 5x + 2 = 0$ or $3x^2 - 5x + 3 = 0$

$\therefore$ $x = 2, \frac{1}{2}$ or $x = \frac{5 \pm \sqrt{(-11)}}{6}$.

Hence the roots are $1, -1, 2, \frac{1}{2}, \frac{5 \pm i\sqrt{11}}{6}$.

## 2.4. (1) SYNTHETIC DIVISION OF A POLYNOMIAL BY A LINEAR EXPRESSION

The division of the polynomial $f(x) = a_0x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots\ldots + a_{n-1}x + a_n$ by a binomial $x - \alpha$ is affected compactly by synthetic division as follows :

$$\begin{array}{cccccc|c} a_0 & a_1 & a_2 \ldots\ldots & a_{n-1} & & a_n & \alpha \\ & \alpha b_0 & \alpha b_1 \ldots\ldots & \alpha b_{n-2} & & \alpha b_{n-1} & \\ \hline a_0 & a_1 + \alpha b_0 & a_2 + \alpha b_1 \ldots\ldots & a_{n-1} + \alpha b_{n-2} & \vert & a_n + \alpha b_{n-1} & \\ (= b_0) & (= b_1) & (= b_2) & (= b_{n-1}) & \vert & (= R) & \end{array}$$

Hence quotient $= b_0 x^{n-1} + b_1 x^{n-2} + \ldots\ldots + b_{n-1}$ and remainder $= R$

**Explanation :** (*i*) Write down the coefficients of the powers of $x$(supplying missing powers of $x$ by zero coefficients) and write $\alpha$ on extreme right.

(*ii*) Put $a_0$ $(= b_0)$ as the first term of 3rd row and multiply it by a and write the product under $a_1$ and add, giving $a_1 + ab_0$ $(= b_1)$.

(*iii*) Multiply $b_1$ by $\alpha$ and write the product under $a_2$ and add, giving $a_2 + \alpha b_1$ $(= b_2)$ and so on.

(*iv*) Continue this process till we get $R$.

**(2) To diminish the roots of an equation f(x) = 0 by h,** *divide $f(x)$ by $x - h$ successively*. Then the *successive remainders, determine the coefficients of the required equation.*

Let the given equation be

$$a_0 x^n + a_1 x^{n-1} + \ldots\ldots + a_{n-1} x + a_n = 0 \qquad \ldots(i)$$

To diminish its roots $h$, put $y = x - h$ (or $x = y + h$) in (*i*) so that

$$a_0 (y + h)^n + a_1 (y + h)^{n-1} + \ldots.. + a_n = 0 \qquad \ldots(ii)$$

On simplification, it takes the form

$$A_0 y^n + A_1 y^{n-1} + \ldots.. + A_n = 0 \qquad \ldots(iii)$$

Its coefficients $A_0, A_1, \ldots.., A_n$ can easily be found with help of *synthetic division*. For this, we put $y = x - h$ in (*iii*) so that

$$A_0 (x - h)^n + A_1 (x - h)^{n-1} + \ldots\ldots + A_n = 0 \qquad \ldots(iv)$$

Clearly (*i*) and (*iv*) are identical. If we divide L.H.S. of (*iv*) by $x - h$, the remainder is $A_n$ and the quotient $Q = A_0 (x - h)^{n-1} + A_1 (x - h)^{n-2} + \ldots.. A_{n-1}$. Similarly if we divide $Q$ by $x - h$, the remainder is $A_{n-1}$ and the quotient is $Q_1$ (say). Again dividing $Q_1$ by $x - h$, $A_{n-2}$ will be obtained and so on.

**Obs.** To increase the roots by $h$, we take $h$ negative.

**Example 2.10.** *Transform the equation $x^3 - 6x^2 + 5x + 8 = 0$ into another in which the second term is missing.* (*Cochin, B. Tech., 2005*)

**Sol.** Sum of the roots of the given equation = 6.

In order that the second term in the transformed equation is missing, the sum of the roots is to be zero.

Since the equation has 3 roots, if we decrease each root by 2, the sum of the roots of the equation will become zero. To diminish the roots by 2, we divide $x^3 - 6x^2 + 5x + 8$ by $x - 2$ successively.

$$\begin{array}{cccc|c} 1 & -6 & 5 & 8 & 2 \\ & 2 & -8 & -6 & \\ \hline & -4 & -3 & \mathbf{2} & \\ & 2 & -4 & & \\ \hline & -2 & \mathbf{-7} & & \\ & 2 & & & \\ \hline \mathbf{1} & \mathbf{0} & & & \end{array}$$

Thus the transformed equation is $x^3 - 7x + 2 = 0$.

**(3) Synthetic division of a polynomial by a quadratic expression.** The division of the polynomial $f(x)$ by the quadratic $x^2 - \alpha x - \beta$ is carried out by the following synthetic scheme :

| | | | | | |
|---|---|---|---|---|---|
| $a_0$ | $a_1$ | $a_2$ | $a_3$ ………$a_{n-1}$ | $a_n$ | |
| | $\alpha\, b_0$ | $\alpha\, b_1$ | $\alpha\, b_2$ ……… $\alpha\, b_{n-2}$ | | $\alpha$ |
| | | $\beta b_0$ | $\beta b_1$ ………. $\beta b_{n-3}$ | $\beta b_{n-2}$ | $\beta$ |
| $a_0$ $(= b_0)$ | $a_1 + \alpha b_0$, $(= b_1)$ | $a_2 + \alpha b_1 + \beta b_0$, $(= b_2)$ | $a_3 + \alpha b_2 + \beta b_1, \ldots$ $(= b_3)$ ... | $a_n + \beta\, b_{n-2}$ $(= b_{n-1})$ $(= b_n)$ | |

Hence the quotient $= b_0\, x^{n-2} + b_1\, x^{n-3} + \ldots.. + b_{n-2}$ and the remainder $= b_{n-1}\, x + b_n$.

**Example 2.11.** *Divide $2x^5 - 3x^4 + 4x^3 - 5x^2 + 6x - 9$ by $x^2 - x + 2$ synthetically.*

**Sol.**

| | | | | | | |
|---|---|---|---|---|---|---|
| 2 | – 3 | 4 | – 5 | 6 | – 9 | |
| | 2 | – 1 | – 1 | – 4 | | 1 |
| | | – 4 | 2 | 2 | 8 | – 2 |
| 2 | – 1 | – 1 | – 4 | 4 | – 1 | |

Hence the quotient $= 2x^3 - x^2 - x - 4$ and the remainder $= 4x - 1$.

## PROBLEMS 2.2

1. Find the equation whose roots are 3 times the roots of $x^3 + 2x^2 - 4x + 1 = 0$.
2. Change the sign of the roots of the equation $x^7 + 3x^5 + x^3 - x^2 + 7x + 1 = 0$.
3. Find the equation whose roots are the negative reciprocals of the roots of $x^4 + 7x^3 + 8x^2 - 9x + 10 = 0$.
4. Solve the equation $81x^3 - 18x^2 - 36x + 8 = 0$, given that its roots are in H.P.
5. Solve : (*i*) $6x^5 + x^4 - 43x^3 - 43x^2 + x + 6 = 0$. (*S.V.T.U., B. Tech., 2006*)

   (*ii*) $3x^6 + x^5 - 27x^4 + 27x^2 - x - 3 = 0$. (*P.T.U., B. Tech., 2001*)

   (*iii*) $2x^5 + x^4 + x + 2 = 12x^3 + 12x^2$.

   (*iv*) $4x^4 - 20x^3 + 33x^2 - 20x + 4 = 0$. (*Madras B.E., 2003*)
6. Find the equation whose roots are the roots of :

   (*i*) $x^4 + x^3 - 3x^2 - x + 2 = 0$ each diminished by 3.

   (*ii*) $x^3 - 2x^2 + 1 = 0$ each increased by 1. (*Madras B.E., 2001*)
7. Show that the equation $x^4 - 10x^3 + 23x^2 - 6x - 15 = 0$ can be transformed into reciprocal equation by diminishing the roots by 2. Hence solve the equation.
8. Find the equation whose roots are the squares of the roots of $x^3 - x^2 + 8x - 6 = 0$.
9. If $\alpha, \beta, \gamma$ are the roots of the equation $2x^3 + 3x^2 - x - 1 = 0$, form the equation whose roots are $(1 - \alpha)^{-1}$, $(1 - \beta)^{-1}$ and $(1 - \gamma)^{-1}$.
10. If $\alpha, \beta, \gamma$ be the roots of $x^3 + px^2 + q = 0$, form the equation whose roots are $\alpha + \beta - \gamma$, $\beta + \gamma - \alpha$, $\gamma + \alpha - \beta$.
11. If $\alpha, \beta, \gamma$ be the roots of $x^3 - 7x + 6 = 0$, form an equation whose roots are $(\beta - \gamma)^2$, $(\gamma - \alpha)^2$, $(\alpha - \beta)^2$.
12. Divide $15x^7 - 16x^6 + 30x^5 - 3x^4 - 5x^3 - 2x^2 + 5x + 8$ by $x^2 - x + 1$ synthetically.

## 2.5. (1) INITIAL APPROXIMATION

Almost all the numerical methods of solution of an equation require an initial rough approximation to its root which is obtained with the help of *Intermediate value property* of equations (§ 2.2). Graphical method is also used to find an approximate value of the root which we explain below.

### (2) Graphical solution of equations

Let the equation be $f(x) = 0$.

(*i*) *Find the interval $(a, b)$ in which a root of $f(x) = 0$ lies.*

(*ii*) *Write the equation $f(x) = 0$ as $\phi(x) = \psi(x)$ where $\psi(x)$ contains only terms in $x$ and the constants.*

(*iii*) *Draw the graphs of $y = \phi(x)$ and $y = \psi(x)$ on the same scale and with respect to the same axes.*

(*iv*) *Read the abscissae of the points of intersection of the curves $y = \phi(x)$ and $y = \psi(x)$. These are the initial approximations to the roots of $f(x) = 0$.*

Sometimes it may not be convenient to write the given equation $f(x) = 0$ in the form $\phi(x) = \psi(x)$. In such cases, we proceed as follows :

(*i*) Form a table for the value of $x$ and $y = f(x)$ directly.

(*ii*) Plot these points and pass a smooth curve through them.

(*iii*) Read the abscissae of the points where this curve cuts the $x$-axis.

These are rough approximations to the roots of $f(x) = 0$.

**Example 2.12.** *Find graphically an approximate value of the root of the equation $3 - x = e^{x-1}$.*

**Sol.** Let $f(x) = e^{x-1} + x - 3 = 0$ ...(*i*)

$f(1) = 1 + 1 - 3 = -$ ve and $f(2) = e + 2 - 3 = 2.718 - 1 = +$ ve

$\therefore$ A root of (*i*) lies between $x = 1$ and $x = 2$.

Let us write (*i*) as $e^{x-1} = 3 - x$.

The abscissa of the point of intersection of the curves

$y = e^{x-1}$ ...(*ii*) and $y = 3 - x$ ...(*iii*)

will give the required root.

*To plot (ii)*, we form the following table of values :

| $x$ | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 1.6 | 1.7 | 1.8 | 1.9 | 2.0 |
|---|---|---|---|---|---|---|---|---|---|---|
| $y = e^{x-1}$ | 1.11 | 1.22 | 1.35 | 1.49 | 1.65 | 1.82 | 2.01 | 2.23 | 2.46 | 2.72 |

Taking the origin at (1, 1) and 1 small unit along either axis = 0.02, we plot these points and pass a smooth curve through them as shown in Fig. 2.2.

*To draw the line (iii)*, we join the points (1, 2) and (2, 1) on the same scale and with the same axes.

From the figure, we get the required root to be $x = 1.44$ nearly.

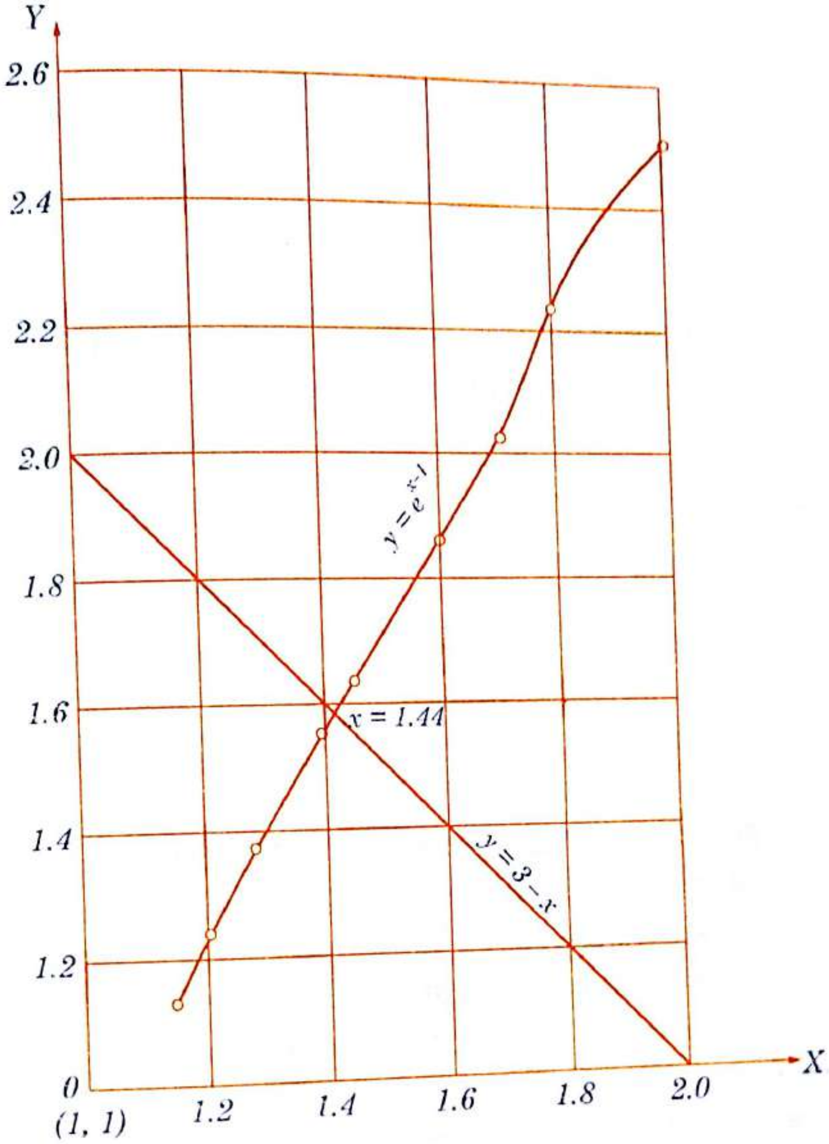


Fig. 2.2

**Example 2.13.** *Obtain graphically an approximate value of the root of $x = \sin x + \pi/2$.*

**Sol.** Let us write the given equation as $\sin x = x - \pi/2$.

The abscissa of the point of intersection of the curve $y = \sin x$ and the line $y = x - \pi/2$ will give a rough estimate of the root.

*To draw the curve $y = \sin x$*, we form the following table :

| $x$ | 0 | $\pi/4$ | $\pi/2$ | $3\pi/4$ | $\pi$ |
|---|---|---|---|---|---|
| $y$ | 0 | 0.71 | 1 | 0.71 | 0 |

Taking 1 unit along either axis $= \pi/4 = 0.8$ nearly, we plot the curve as shown in Fig. 2.3.

Also we draw the line $y = x - \pi/2$ to the same scale and with the same axes.

From the graph, we get $x = 2.3$ radians approximately.

*To draw (iii)*, we form the following table :

| $x$ | 1.57 | 2.36 | 3.14 |
|---|---|---|---|
| $\cosh x$ | 2.51 | 5.34 | 11.57 |
| $y = -\operatorname{sech} x$ | – 0.4 | – 0.19 | – 0.09 |

Then we plot the curve (*iii*) to the same scale with the same axes.

From the above figure, we get the lowest root to be approximately $x = 1.57 + 0.29 = 1.86$.

## PROBLEMS 2.3

*Find the approximate value of the root of the following equations graphically (1—4) :*

1. $x^3 - x - 1 = 0$ (*Madras B.E., 2000 S*)
2. $x^3 - 6x^2 + 9x - 3 = 0$
3. $\tan x = 1.2\, x$
4. $x = 3 \cos (x - \pi/4)$.
5. Find the approximate value of the smallest real root of the equation $e^{-x} - \sin x = 0$.

## 2.6. RATE OF CONVERGENCE

Let $x_0, x_1, x_2$, ...... be the values of a root ($\alpha$) of an equation at the 0th, 1st, 2nd ...... iterations while its actual value is 3.5567. The values of this root calculated by three different methods, are as given below :

| *Root* | *1st method* | *2nd method* | *3rd method* |
|---|---|---|---|
| $x_0$ | 5 | 5 | 5 |
| $x_1$ | 5.6 | 3.8527 | 3.8327 |
| $x_2$ | 6.4 | 3.5693 | 3.56834 |
| $x_3$ | 8.3 | 3.55798 | 3.55743 |
| $x_4$ | 9.7 | 3.55687 | 3.55672 |
| $x_5$ | 10.6 | 3.55676 | |
| $x_6$ | 11.9 | 3.55671 | |

The values in the 1st method do not converge towards the root 3.5567. In the 2nd and 3rd methods, the values converge to the root after 6th and 4th iterations respectively. Clearly 3rd method converges faster than the 2nd method. This *fastness of convergence in any method is represented by its* **rate of convergence.**

If $e$ be the error then $e_i = \alpha - x_i = x_{i+1} - x_i$.

*If* $e_{i+1}/e_i$ *is almost constant, convergence is said to be* **linear** *i.e. slow.*

*If* $e_{i+1}/e_i^p$ *is nearly constant, convergence is said to be of order p i.e. faster.*

## 2.7. BISECTION METHOD

This method is based on the repeated application of the *intermediate value property*. Let the function $f(x)$ be continuous between $a$ and $b$. For definiteness, let $f(a)$ be negative and $f(b)$ be positive. Then the first approximation to the root is $x_1 = \frac{1}{2}(a + b)$.

If $f(x_1) = 0$, then $x_1$ is a root of $f(x) = 0$. Otherwise, the root lies between $a$ and $x_1$ or $x_1$ and $b$ according as $f(x_1)$ is positive or negative. Then we bisect the interval as before and continue the process until the root is found to desired accuracy.

In the Fig. 2.4, $f(x_1)$ is + ve, so that the root lies between $a$ and $x_1$. Then the second approximation to the root is $x_2 = \frac{1}{2}(a + x_1)$. If $f(x_2)$ is – ve, the root lies between $x_1$ and $x_2$. Then the third approximation to the root is $x_3 = \frac{1}{2}(x_1 + x_2)$ and so on.

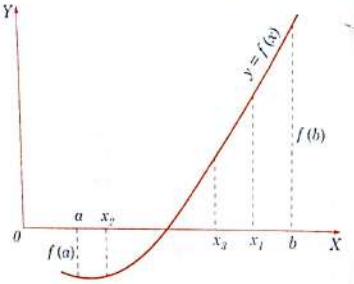


Fig. 2.5

Since the new interval containing the root, is exactly half the length of the previous one, the interval width is reduced by a factor of $\frac{1}{2}$ at each step. At the end of the $n$th step, the new interval will therefore, be of length $(b - a)/2^n$. If on repeating this process $n$ times, the latest interval is as small as given $\varepsilon$, then $(b - a)/2^n \leq \varepsilon$

or
$$n \geq [\log(b - a) - \log \varepsilon]/\log 2$$

*This gives the number of iterations required for achieving an accuracy* $\varepsilon$.

In particular, the minimum number of iterations required for converging to a root in the interval (0, 1) for a given $\varepsilon$ are as under :

| | | | |
|---|---|---|---|
| $\varepsilon$ : | $10^{-2}$ | $10^{-3}$ | $10^{-4}$ |
| $n$ : | 7 | 10 | 14 |

**Obs. 2.** As the error decreases with each step by a factor of $\frac{1}{2}$, (*i.e.* $e_{n+1}/e_n = \frac{1}{2}$), *the convergence in the bisection method is* **linear**.

**Example 2.15.** (*a*) *Find a root of the equation* $x^3 - 4x - 9 = 0$, *using the bisection method correct to three decimal places.* (*Mumbai, B. Tech., 2003*)

(*b*) *Using bisection method, find the negative root of the equation* $x^3 - 4x + 9 = 0$. (*J.N.T.U., B.Tech., 2009*)

**Sol.** (*a*) Let $\qquad f(x) = x^3 - 4x - 9$

Since $f(2)$ is – ve and $f(3)$ is + ve, a root lies between 2 and 3.

$\therefore$ First approximation to the root is

$$x_1 = \tfrac{1}{2}(2 + 3) = 2.5.$$

Thus $\qquad f(x_1) = (2.5)^3 - 4(2.5) - 9 = -3.375$ *i.e.* –ve.

$\therefore$ The root lies between $x_1$ and 3. Thus the second approximation to the root is

$$x_2 = \tfrac{1}{2}(x_1 + 3) = 2.75.$$

Then $\qquad f(x_2) = (2.75)^3 - 4(2.75) - 9 = 0.7969$ *i.e.* +ve.

$\therefore$ The root lies between $x_1$ and $x_2$. Thus the third approximation to the root is

$$x_3 = \tfrac{1}{2}(x_1 + x_2) = 2.625.$$

Then $\qquad f(x_3) = (2.625)^3 - 4(2.625) - 9 = -1.4121$ *i.e.* –ve.

The root lies between $x_2$ and $x_3$. Thus the fourth approximation to the root is

$$x_4 = \tfrac{1}{2}(x_2 + x_3) = 2.6875.$$

Repeating this process, the successive approximations are

$$x_5 = 2.71875,\quad x_6 = 2.70313,\quad x_7 = 2.71094$$
$$x_8 = 2.70703,\quad x_9 = 2.70508,\quad x_{10} = 2.70605$$



$$x_{11} = 2.70654,\quad x_{12} = 2.70642$$

Hence the root is 2.7064.

(*b*) If $\alpha, \beta, \gamma$ are the roots of the given equation, then $-\alpha, -\beta, -\gamma$ are the roots of $(-x)^3 - 4(-x) + 9 = 0$ $\therefore$ The negative root of the given equation is the positive root of $x^3 - 4x - 9 = 0$ which we have found above to be 2.7064.

Hence the negative root the given equation – 2.7064.

**Example 2.16.** *Using the bisection method, find an approximate root of the equation* $\sin x = 1/x$, *that lies between* $x = 1$ *and* $x = 1.5$ (*measured in radians*). *Carry out computations upto the 7th stage.* (*V.T.U., B.E., 2003 S*)

**Sol.** Let $f(x) = x \sin x - 1$. We know that $1^r = 57.3°$.

Since $\qquad f(1) = 1 \times \sin(1) - 1 = \sin(57.3°) - 1 = -0.15849$

and $\qquad f(1.5) = 1.5 \times \sin(1.5)^r - 1 = 1.5 \times \sin(85.95)° - 1 = 0.49625$ ;

a root lies between 1 and 1.5.

$\therefore$ First approximation to the root is $x_1 = \frac{1}{2}(1 + 1.5) = 1.25$.

Then $f(x_1) = (1.25)\sin(1.25) - 1 = 1.25 \sin(71.625°) - 1 = 0.18627$ and $f(1) < 0$.

$\therefore$ A root lies between 1 and $x_1 = 1.25$.

Thus the second approximation to the root is $x_2 = \frac{1}{2}(1 + 1.25) = 1.125$.

Then $f(x_2) = 1.125 \sin(1.125) - 1 = 1.125 \sin(64.46)° - 1 = 0.01509$ and $f(1) < 0$.

$\therefore$ A root lies between 1 and $x_2 = 1.125$.

Thus the third approximation to the root is $x_3 = \frac{1}{2}(1 + 1.125) = 1.0625$.

Then $f(x_3) = 1.0625 \sin(1.0625) - 1 = 1.0625 \sin(60.88) - 1 = -0.0718 < 0$ and $f(x_2) > 0$, *i.e.* now the root lies between $x_3 = 1.0625$ and $x_2 = 1.125$.

$\therefore$ Fourth approximation to the root is $x_4 = \frac{1}{2}(1.0625 + 1.125) = 1.09375$

Then $\qquad f(x_4) = -0.02836 < 0$ and $f(x_2) > 0$,

*i.e.* $\quad$ The root lies between $x_4 = 1.09375$ and $x_2 = 1.125$.

$\therefore$ Fifth approximation to the root is $x_5 = \frac{1}{2}(1.09375 + 1.125) = 1.10937$

Then $\qquad f(x_5) = -0.00664 < 0$ and $f(x_2) > 0$.

$\therefore$ The root lies between $x_5 = 1.10937$ and $x_2 = 1.125$.

Thus the sixth approximation to the root is

$$x_6 = \frac{1}{2}(1.10937 + 1.125) = 1.11719$$

Then $\qquad f(x_6) = 0.00421 > 0$. But $f(x_5) < 0$.

$\therefore$ The root lies between $x_5 = 1.10937$ and $x_6 = 1.11719$.

Thus the seventh approximation to the root is $x_7 = \frac{1}{2}(1.10937 + 1.11719) = 1.11328$

Hence the desired approximation to the root is 1.11328.

**Example 2.17.** *Find the root of the equation $\cos x = xe^x$ using the bisection method correct to four decimal places.* *(Mumbai, B. Tech., 2004)*

**Sol.** Let $f(x) = \cos x - xe^x$.

Since $f(0) = 1$ and $f(1) = -2.18$, so a root lies between 0 and 1.

$\therefore$ First approximation to the root is $x_1 = \frac{1}{2}(0 + 1) = 0.5$

Now $f(x_1) = 0.05$ and $f(1) = -2.18$, therefore the root lies between 1 and $x_1 = 0.5$.

$\therefore$ Second approximation to the root is $x_2 = \frac{1}{2}(0.5 + 1) = 0.75$

Now $f(x_2) = -0.86$ and $f(0.5) = 0.05$, therefore the root lies between 0.5 and 0.75.

$\therefore$ Third approximation to the root is $x_3 = \frac{1}{2}(0.5 + 0.75) = 0.625$

Now $f(x_3) = -0.36$ and $f(0.5) = 0.05$, therefore the root lies between 0.5 and 0.625.

$\therefore$ Fourth approximation to the root is $x_4 = \frac{1}{2}(0.5 + 0.625) = 0.5625$

Now $f(x_4) = -0.14$ and $0.5 = 0.05$, therefore the root lies between 0.5 and 0.5625

$\therefore$ Fifth approximation is $x_5 = \frac{1}{2}(0.5 + 0.5625) = 0.5312$

Now $f(x_5) = -0.04$ and $f(0.5) = 0.05$, therefore the root lies between 0.5 and 0.5312.

$\therefore$ Sixth approximation is $x_6 = \frac{1}{2}(0.5 + 0.5312) = 0.5156$

Hence the desired approximation to the root is 0.5156.

**Example 2.18.** *Find a positive real root of $x \log_{10} x = 1.2$ using the bisection method.*

**Sol.** Let $f(x) = x \log_{10} x - 1.2$.

Since $f(2) = -0.598$ and $f(3) = 0.231$, so a root lies between 2 and 3.

$\therefore$ First approximation to the root is $x_1 = \frac{1}{2}(2 + 3) = 2.5$.

Now $f(2.5) = -0.205$ and $f(3) = 0.231$, therefore a root lies between 2.5 and 3.

$\therefore$ Second approximation to the root is $x_2 = \frac{1}{2}(2.5 + 3) = 2.75$.

Now $f(2.75) = 0.008$ and $f(2.5) = -0.205$, therefore, a root lies between 2.5 and 2.75.

$\therefore$ Third approximation to the root is $x_3 = \frac{1}{2}(2.5 + 2.75) = 2.625$

Now $f(2.625) = -0.1$ and $f(2.75) = 0.008$, therefore a root lies between 2.625 and 2.75.

$\therefore$ Fourth approximation to the root is $x_4 = \frac{1}{2}(2.625 + 2.75) = 2.687$

Hence the desired root is 2.687.



## 2.8. METHOD OF FALSE POSITION or REGULA-FALSI METHOD

This is the oldest method of finding the real root of an equation $f(x) = 0$ and closely resembles the bisection method.

Here we choose two points $x_0$ and $x_1$ such that $f(x_0)$ and $f(x_1)$ are of opposite signs *i.e.* the graph of $y = f(x)$ crosses the $x$-axis between these points (Fig. 2.6). This indicates that a root lies between $x_0$ and $x_1$ and consequently $f(x_0) f(x_1) < 0$.

Equation of the chord joining the points $A[x_0, f(x_0)]$ and $B[x_1, f(x_1)]$ is

$$y - f(x_0) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0)$$

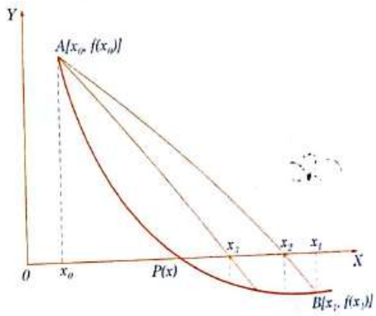


Fig. 2.6

The method consists in replacing the curve $AB$ by means of the chord $AB$ and taking the point of intersection of the chord with the $x$-axis as an approximation to the root. So the abscissa of the point where the chord cuts the $x$-axis ($y = 0$) is given by

$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) \qquad \ldots(1)$$

which is an approximation to the root.

If now $f(x_0)$ and $f(x_2)$ are of opposite signs, then the root lies between $x_0$ and $x_2$. So replacing $x_1$ by $x_2$ in (1), we obtain the next approximation $x_3$. (The root could as well lie between $x_1$ and $x_2$ and we would obtain $x_3$ accordingly). This procedure is repeated till the root is found to the desired accuracy. The iteration process based on (1) is known as the *method of false position* and its rate of convergence is faster than that of the bisection method.

**Example 2.19.** *Find a real root of the equation $x^3 - 2x - 5 = 0$ by the method of false position correct to three decimal places.* *(Manipal, B.E., 2005)*

**Sol.** Let $\qquad f(x) = x^3 - 2x - 5$

so that $\qquad f(2) = -1$ and $f(3) = 16$,

*i.e.* A root lies between 2 and 3.

$\therefore$ Taking $x_0 = 2, x_1 = 3, f(x_0) = -1, f(x_1) = 16$, in the method of false position, we get

$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 2 + \frac{1}{17} = 2.0588 \qquad ...(i)$$

Now $\qquad f(x_2) = f(2.0588) = -0.3908$

*i.e.*, the root lies between 2.0588 and 3.

$\therefore$ Taking $x_0 = 2.0588, x_1 = 3, f(x_0) = -0.3908, f(x_1) = 16$, in (*i*), we get

$$x_3 = 2.0588 - \frac{0.9412}{16.3908}(-0.3908) = 2.0813$$

Repeating this process, the successive approximations are

$$x_4 = 2.0862, \quad x_5 = 2.0915, x_6 = 2.0934,$$
$$x_7 = 2.0941, \quad x_8 = 2.0943 \text{ etc.}$$

Hence the root is 2.094 correct to 3 decimal places.

**Example 2.20.** *Find the root of the equation $\cos x = xe^x$ using the regula-falsi method correct to four decimal places.* (*Bhopal B. Tech., 2009*)

**Sol.** Let $\qquad f(x) = \cos x - xe^x = 0$

so that $\qquad f(0) = 1, f(1) = \cos 1 - e = -2.17798$

*i.e.*, the root lies between 0 and 1.

$\therefore$ Taking $x_0 = 0, x_1 = 1, f(x_0) = 1$ and $f(x_1) = -2.17798$ in the regula-falsi method, we get

$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 0 + \frac{1}{3.17798} \times 1 = 0.31467 \qquad ...(i)$$

Now $\qquad f(0.31467) = 0.51987$ *i.e.*, the root lies between 0.31467 and 1.

$\therefore$ Taking $x_0 = 0.31467, x_1 = 1, f(x_0) = 0.51987, f(x_1) = -2.17798$ in (*i*), we get

$$x_3 = 0.31467 + \frac{0.68533}{2.69785} \times 0.51987 = 0.44673$$

Now $\qquad f(0.44673) = 0.20356$ *i.e.*, the root lies between 0.44673 and 1.

$\therefore$ Taking $x_0 = 0.44673, x_1 = 1, f(x_0) = 0.20356, f(x_1) = -2.17798$ in (*i*), we get

$$x_4 = 0.44673 + \frac{0.55327}{2.38154} \times 0.20356 = 0.49402$$

Repeating this process, the successive approximations are

$$x_5 = 0.50995, x_6 = 0.51520, x_7 = 0.51692$$
$$x_8 = 0.51748, x_9 = 0.51767, x_{10} = 0.51775 \text{ etc.}$$

Hence the root is 0.5177 correct to 4 decimal places.

**Example 2.21.** *Find a real root of the equation $x \log_{10} x = 1.2$ by regula-falsi method correct to four decimal places.* (*J.N.T.U., B. Tech., 2008 ; V.T.U., B.E., 2008*)

**Sol.** Let $\qquad f(x) = x \log_{10} x - 1.2$

so that $\qquad f(1) = -$ ve, $f(2) = -$ ve and $f(3) = +$ ve.

$\therefore$ A root lies between 2 and 3.

Taking $x_0 = 2$ and $x_1 = 3, f(x_0) = -0.59794$ and $f(x_1) = 0.23136$, in the method of false position, we get



$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 2.72102 \qquad ...(i)$$

Now $\qquad f(x_2) = f(2.72102) = -0.01709$

*i.e.* the root lies between 2.72102 and 3.

$\therefore$ Taking $x_0 = 2.72102, x_1 = 3, f(x_0) = -0.01709$ and $f(x_1) = 0.23136$ in (*i*), we get

$$x_3 = 2.72102 + \frac{0.27898}{0.23136 + 0.01709} \times 0.01709 = 2.74021$$

Repeating this process, the successive approximations are

$$x_4 = 2.74024, x_5 = 2.74063 \text{ etc.}$$

Hence the root is 2.7406 correct to 4 decimal places.



**Example 2.22.** *Use the method of false position, to find the fourth root of 32 correct to three decimal places.*

**Sol.** Let $x = (32)^{1/4}$ so that $x^4 - 32 = 0$

Take $f(x) = x^4 - 32$. Then $f(2) = -16$ and $f(3) = 49$, *i.e.* a root lies between 2 and 3.

$\therefore$ Taking $x_0 = 2, x_1 = 3, f(x_0) = -16, f(x_1) = 49$ in the method of false position, we get

$$x_2 = x_0 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_0) = 2 + \frac{16}{65} = 2.2462 \qquad ...(i)$$

Now $f(x_2) = f(2.2462) = -6.5438$ *i.e.* the root lies between 2.2462 and 3.

$\therefore$ Taking $x_0 = 2.2462, x_1 = 3, f(x_0) = -6.5438, f(x_1) = 49$ in (*i*), we get

$$x_3 = 2.2462 - \frac{3 - 2.2462}{49 + 6.5438}(-6.5438) = 2.335$$

Now $f(x_3) = f(2.335) = -2.2732$ *i.e.* the root lies between 2.335 and 3.

$\therefore$ Taking $x_0 = 2.335$ and $x_1 = 3, f(x_0) = -2.2732$ and $f(x_1) = 49$ in (*i*), we obtain

$$x_4 = 2.335 - \frac{3 - 2.335}{49 + 2.2732}(-2.2732) = 2.3645$$

Repeating this process, the successive approximations are $x_5 = 2.3770, x_6 = 2.3779$ etc.

Since $x_5 = x_6$ upto 3 decimal places, we take $(32)^{1/4} = 2.378$.

## 2.9. SECANT METHOD

This method is an improvement over the method of false position as it does not require the condition $f(x_0) f(x_1) < 0$ of that method (Fig. 2.5). Here also the graph of the function $y = f(x)$ is approximated by a secant line but at each iteration, two most recent approximations to the root are used to find the next approximation. Also it is not necessary that the interval must contain the root.

Taking $x_0, x_1$ as the initial limits of the interval, we write the equation of the chord joining these as

$$y - f(x_1) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_1)$$

Then the abscissa of the point where it crosses the $x$-axis ($y = 0$) is given by

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

which is an approximation to the root. The general formula for successive approximations is, therefore, given by

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

**Obs.** If at any interation $f(x_n) = f(x_{n-1})$, this method fails and shows that it does not converge necessarily. *This is a drawback of secant method over the method of false position which always converges.* ***But if the secant method once converges, its rate of convergence is 1.6 which is faster than that of the method of false position.***

**Example 2.23.** *Find a root of the equation $x^3 - 2x - 5 = 0$ using secant method correct to three decimal places.*

**Sol.** Let $f(x) = x^3 - 2x - 5$ so that $f(2) = -1$ and $f(3) = 16$.

$\therefore$ Taking initial approximations $x_0 = 2$ and $x_1 = 3$, by secant method, we have

$$x_2 = x_1 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_1) = 3 - \frac{3-2}{16+1} 16 = 2.058823$$

Now $\quad f(x_2) = -0.390799$

$$\therefore \quad x_3 = x_2 - \frac{x_2 - x_1}{f(x_2) - f(x_1)} f(x_2) = 2.081263$$

and $\quad f(x_3) = -0.147204$

$$\therefore \quad x_4 = x_3 - \frac{x_3 - x_2}{f(x_3) - f(x_2)} f(x_3) = 2.094824$$

and $\quad f(x_4) = 0.003042$

$$\therefore \quad x_5 = x_4 - \frac{x_4 - x_3}{f(x_4) - f(x_3)} f(x_4) = 2.094549$$

Hence the root is 2.094 correct to 3 decimal places.

**Example 2.24.** *Find the root of the equation $xe^x = \cos x$ using the secant method correct to four decimal places.* (U.P.T.C., MCA., 2009)

**Sol.** Let $\quad f(x) = \cos x - xe^x = 0.$

Taking the initial approximations $x_0 = 0, x_1 = 1$

so that $f(x_0) = 1, f(x_1) = \cos 1 - e = -2.17798$

Then by secant method, we have

$$x_2 = x_1 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_1) = 1 + \frac{1}{3.17798}(-2.17798) = 0.31467$$

Now $\quad f(x_2) = 0.51987$

$$\therefore \quad x_3 = x_2 - \frac{x_2 - x_1}{f(x_2) - f(x_1)} f(x_2) = 0.44673 \quad \text{and} \quad f(x_3) = 0.20354$$



$$\therefore \quad x_4 = x_3 - \frac{x_3 - x_2}{f(x_3) - f(x_2)} f(x_3) = 0.53171.$$

Repeating this process, the successive approximations are $x_5 = 0.51690$, $x_6 = 0.51775$, $x_7 = 0.51776$ etc.

Hence the root is 0.5177 correct to 4 decimal places.

**Obs.** Comparing Examples 2.18 and 2.21, we notice that the rate of convergence in secant method is definitely faster than that of the method of false position.

## 2.10. (1) ITERATION METHOD

To find the roots of the equation $f(x) = 0$ ...(i)

by successive approximations, we rewrite (i) in the form $x = \phi(x)$ ...(ii)

The roots of (i) are the same as the points of intersection of the straight line $y = x$ and the curve representing $y = \phi(x)$. Fig. 2.6 illustrates the working of the iteration method which provides a spiral solution.

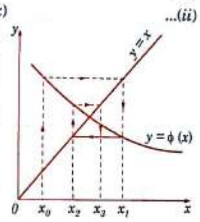

Fig. 2.6

Let $x = x_0$ be an initial approximation of the desired root $\alpha$. Then the first approximation $x_1$ is given by

$x_1 = \phi(x_0)$

Now treating $x_1$ as the initial value, the second approximation is $\quad x_2 = \phi(x_1)$

Proceeding in this way, the $n$th approximation is given by $\quad x_n = \phi(x_{n-1})$

**(2) Sufficient condition for convergence of iterations.** Now it is not sure whether the sequence of approximations $x_1, x_2, \ldots, x_n$ always converges to the same number which is a root of (1) or not. As such, we have to choose the initial approximation $x_0$ suitably so that the successive approximations $x_1, x_2, \ldots, x_n$ converge to the root $\alpha$. The following theorem helps in making the right choice of $x_0$:

**Theorem.** *If (i) $\alpha$ be a root of $f(x) = 0$ which is equivalent to $x = \phi(x)$,*

*(ii) I, be any interval containing the point $x = \alpha$,*

*(iii) $|\phi'(x)| < 1$ for all $x$ in I,*

*then the sequence of approximations $x_0, x_1, x_2, \ldots, x_n$ will converge to the root $\alpha$ provided the initial approximation $x_0$ is chosen in I.*

**Proof.** Since $\alpha$ is a root of $x = \phi(x)$, we have $\alpha = \phi(\alpha)$

If $x_{n-1}$ and $x_n$ be 2 successive approximations to $\alpha$, we have $x_n = \phi(x_{n-1})$

$$x_n - \alpha = \phi(x_{n-1}) - \phi(\alpha) \quad \text{...(1)}$$

By mean value theorem, $\dfrac{\phi(x_{n-1}) - \phi(\alpha)}{x_{n-1} - \alpha} = \phi'(\xi)$ where $x_{n-1} < \xi < \alpha$

Hence (1) becomes $x_n - \alpha = (x_{n-1} - \alpha)\,\phi'(\xi)$

If $|\phi'(x_i)| \leq k < 1$ for all $i$, then ...(2)

$$|x_n - \alpha| \leq k\,|x_{n-1} - \alpha|$$

Similarly $\quad |x_{n-1} - \alpha| \leq k\,|x_{n-2} - \alpha|$

*i.e.,* $\quad |x_n - \alpha| \leq k^2\,|x_{n-2} - \alpha|$

Proceeding in this way, $\quad | x_n - \alpha | \le k^n \, | x_0 - \alpha |$

As $n \to \infty$, the R.H.S. tends to zero, therefore, the sequence of approximations converges to the root $\alpha$.

**Obs. 1.** The smaller the value of $\phi'(x)$, the more rapid will be the convergence.

**2.** This method of iteration is particularly useful for finding the real roots of an equation given in the form of an infinite series.

**(3) Acceleration of convergence.** From (2), we have

$$| x_n - \alpha | \le k \; | x_{n-1} - \alpha | \,, k < 1.$$

It is clear from this relation that the iteration method is **linearly convergent**. *This slow rate of convergence can be improved by using the following method :*

**(4) Aitken's $\Delta^2$ method.** Let $x_{i-1}$, $x_i$, $x_{i+1}$ be three successive approximations to the desired root $\alpha$ of the equation $x = \phi(x)$. Then we know that

$$\alpha - x_i = k(\alpha - x_{i-1}), \; \alpha - x_{i+1} = k(\alpha - x_i)$$

Dividing, we get $\dfrac{\alpha - x_i}{\alpha - x_{i-1}} = \dfrac{\alpha - x_{i-1}}{\alpha - x_i}$

whence $$\alpha = x_{i+1} - \frac{(x_{i-1} - x_i)^2}{x_{i-1} - 2x_i + x_{i-1}} \qquad ...(3)$$

But in the sequence of successive approximations, we have

$$\Delta x_i = x_{i+1} - x_i$$
$$\Delta^2 x_i = \Delta(\Delta x_1) = \Delta(x_{i+1} - x_i) = \Delta x_{i+1} - \Delta x_i$$
$$= x_{i+2} - x_{i+1} - (x_{i+1} - x_i) = x_{i+2} - 2x_{i+1} + x_i$$
$$\therefore \quad \Delta^2 x_{i-1} = x_{i+1} - 2x_i + x_{i-1}$$

Hence (3) can be written as $$\alpha = x_{i+1} - \frac{(\Delta x_i)^2}{\Delta^2 x_{i-1}} \qquad ...(4)$$

which yields successive approximations to the root $\alpha$.

**Example 2.25.** *Find a real root of the equation $\cos x = 3x - 1$ correct to three decimal places using*

*(i) Iteration method*

*(ii) Aitken's $\Delta^2$ method.* *(J.N.T.U., B.Tech., 2009)*

**Sol.** *(i)* We have $\quad f(x) = \cos x - 3x + 1 = 0$

$f(0) = 2 = +$ ve and $f(\pi/2) = -3\pi/2 + 1 = -$ ve

$\therefore$ A root lies between 0 and $\pi/2$.

Rewriting the given equation as $x = \frac{1}{3}(\cos x + 1) = \phi(x)$, we have

$$\phi'(x) = \frac{\sin x}{3} \quad \text{and} \quad | \phi'(x) | = \tfrac{1}{3} \, | \sin x | < 1 \text{ in } (0, \pi/2).$$

Hence the iteration method can be applied and we start with $x_0 = 0$. Then the successive approximations are,

$$x_1 = \phi(x_0) = \tfrac{1}{3}(\cos 0 + 1) = 0.6667$$
$$x_2 = \phi(x_1) = \tfrac{1}{3}(\cos 0.6667 + 1) = 0.5953$$



$$x_3 = \phi(x_2) = \tfrac{1}{3}(\cos 0.5953 + 1) = 0.6093$$
$$x_4 = \phi(x_3) = \tfrac{1}{3}(\cos 0.6093 + 1) = 0.6067$$
$$x_5 = \phi(x_4) = \tfrac{1}{3}(\cos 0.6067 + 1) = 0.6072$$
$$x_6 = \phi(x_5) = \tfrac{1}{3}(\cos 0.6072 + 1) = 0.6071$$

Hence $x_5$ and $x_6$ being almost the same, the root is 0.607 correct to 3 decimal places.

*(ii)* We calculate $x_1, x_2, x_3$ as above. To use Aitken's method, we have

| $x$ | $\Delta x$ | $\Delta^2 x$ |
|---|---|---|
| $x_1 = 0.667$ | | |
| | $-0.0714$ | |
| $x_2 = 0.5953$ | | $\underline{0.0854}$ |
| | $\underline{0.014}$ | |
| $x_3 = \underline{0.6093}$ | | |

Hence $$x_4 = x_3 - \frac{(\Delta x_2)^2}{\Delta^2 x_1} = 0.6093 - \frac{(0.014)^2}{0.0854} = 0.607$$

which corresponds to six iterations in normal form.

Thus the required root is 0.607.

**Example 2.26.** *Using iteration method, find a root of the equation $x^3 + x^2 - 1 = 0$ correct to four decimal places.* *(U.P.T.U., B.Tech., 2006)*

**Sol.** We have $f(x) = x^3 + x^2 - 1 = 0$

Since $f(0) = -1$ and $f(1) = 1$, a root lies between 0 and 1.

Rewriting the given equation as $x = (x + 1)^{-1/2} = \phi(x)$, we have $\phi'(x) = -\dfrac{1}{2}(x + 1)^{-3/2}$ and $| \phi'(x) | < 1$ for $x < 1$. Hence the iteration method can be applied. Starting with $x_0 = 0.75$, the successive approximations are

$$x_1 = \phi(x_0) = \frac{1}{\sqrt{(x_0 + 1)}} = 0.7559$$

$$x_2 = \phi(x_1) = \frac{1}{\sqrt{(0.7559 + 1)}} = 0.75466$$

$$x_3 = 0.75492, \; x_4 = 0.75487, \; x_5 = 0.75488$$

Hence $x_4$ and $x_5$ being almost the same, the root is 0.7548 correct to 4 decimal places.

**Example 2.27.** *Apply iteration method to find the negative root of the equation $x^3 - 2x + 5 = 0$ correct to four decimal places.*

**Sol.** If $\alpha, \beta, \gamma$ are the roots of the given equation, then $-\alpha, -\beta, -\gamma$ are the roots of

$$(-x)^3 - 2(-x) + 5 = 0$$

$\therefore$ The negative root of the given equation is the positive root of

$$f(x) = x^3 - 2x - 5 = 0. \qquad ...(i)$$

Since $f(2) = -1$ and $f(3) = 16$, a root lies between 2 and 3.

Rewriting *(i)* as $x = (2x + 5)^{1/3} = \phi(x)$, we have $\phi'(x) = \frac{1}{3}(2x+5)^{-2/3}.\ 2$ and $|\phi'(x)| < 1$ for $x < 3$.

$\therefore$ The iteration method can be applied :

Starting with $x_0 = 2$. The successive approximations are

$$x_1 = \phi(x_0) = (2x_0 + 5)^{1/3} = 2.08008$$
$$x_2 = \phi(x_1) = 2.09235, \quad x_3 = 2.09422$$
$$x_4 = 2.09450, \quad x_5 = 2.09454$$

Since $x_4$ and $x_5$ being almost the same, the root of (*i*) is 2.0945 correct to 4 decimal places.

Hence the negative root of the given equation is – 2.0945.

**Example 2.28.** *Find a real root of $2x - \log_{10} x = 7$ correct to four decimal places using iteration method.* (*U.P.T.U., B. Tech., 2004*)

**Sol.** We have $f(x) = 2x - \log_{10} x - 7$

$$f(3) = 6 - \log_{10} 3 - 7 = 6 - 0.4771 - 7 = -1.4471$$
$$f(4) = 8 - \log_{10} 4 - 7 = 8 - 0.602 - 7 = 0.398$$

$\therefore$ A root lies between 3 and 4.

Rewriting the given equation as $x = \frac{1}{2}(\log_{10} x + 7) = \phi(x)$, we have

$$\phi'(x) = \frac{1}{2}\left(\frac{1}{x}\log_{10} e\right)$$

$\therefore$ $|\phi'(x)| < 1$ when $3 < x < 4$ $\qquad [\because \log_{10} e = 0.4343]$

Since $|f(4)| < |f(3)|$, the root is near to 4.

Hence the iteration method can be applied. Taking $x_0 = 3.6$, the successive approximations are

$$x_1 = \phi(x_0) = \frac{1}{2}(\log_{10} 3.6 + 7) = 3.77815$$
$$x_2 = \phi(x_1) = \frac{1}{2}(\log_{10} 3.77815 + 7) = 3.78863$$
$$x_3 = \phi(x_2) = \frac{1}{2}(\log 3.78863 + 7) = 3.78924$$
$$x_4 = \phi(x_3) = \frac{1}{2}(\log 3.78924 + 7) = 3.78927$$

Hence $x_3$ and $x_4$ being almost equal, the root is 3.7892 correct to 4 decimal places.

**Example 2.29.** *Find the smallest root of the equation*

$$1 - x + \frac{x^2}{(2!)^2} - \frac{x^3}{(3!)^2} + \frac{x^4}{(4!)^2} - \frac{x^5}{(5!)^2} + \ldots = 0$$

**Sol.** Writing the given equation as

$$x = 1 + \frac{x^2}{(2!)^2} - \frac{x^3}{(3!)^2} + \frac{x^4}{(4!)^2} - \frac{x^5}{(5!)^2} + \ldots = \phi(x)$$

Omitting $x^2$ and higher powers of $x$, we get $x = 1$ approximately.



Taking $x_0 = 1$, we obtain

$$x_1 = \phi(x_0) = 1 + \frac{1}{(2!)^2} - \frac{1}{(3!)^2} + \frac{1}{(4!)^2} - \frac{1}{(5!)^2} + \ldots = 1.2239$$

$$x_2 = \phi(x_1) = 1 + \frac{(1.2239)^2}{(2!)^2} - \frac{(1.2239)^3}{(3!)^2} + \frac{(1.2239)^4}{(4!)^2} - \frac{(1.2239)^5}{(5!)^2} + \ldots = 1.3263$$

Similarly $x_3 = 1.38$, $x_4 = 1.409$, $x_5 = 1.425$

$x_6 = 1.434$, $x_7 = 1.439$, $x_8 = 1.442$.

The values of $x_7$ and $x_8$ indicate that the root is 1.44 correct to 2 decimal places.

## PROBLEMS 2.4

1. Find a root of the following equations, using the bisection method correct to three decimal places :
 (*i*) $x^3 - x - 11 = 0$ (*Kerala, B. Tech., 2003*) (*ii*) $x^3 - x^2 - 1 = 0$ (*J.N.T.U., B. Tech., 2009*)
 (*iii*) $2x^3 + x^2 - 20x + 12 = 0$ (*iv*) $x^4 - x - 10 = 0$.
2. Evaluate a real root of the following equations by bisection method:
 (*i*) $x - \cos x = 0$ (*Mumbai, B.E., 2004*) (*ii*) $e^{-x} - x = 0$
 (*iii*) $e^x = 4 \sin x$.
3. Find a real root of the following equations correct to three decimal places, by the method of false position :
 (*i*) $x^3 + x - 1 = 0$ (*Ranchi, B. Tech., 2000*) (*ii*) $x^3 - 4x - 9 = 0$ (*V.T.U., B., Tech.,2007*)
 (*iii*) $x^6 - x^4 - x^3 - 1 = 0$. (*Nagarjuna, B.E., 2001*)
4. Using Regula falsi method, compute the real root of the following equations correct to three decimal places :
 (*i*) $e^x - 4x = 0$ (*ii*) $xe^x = 2$ (*S.V.T.U., B.E., 2007*)
 (*iii*) $\cos x = 3x - 1$ (*iv*) $x \tan x = -1$
 (*v*) $2x - \log x = 7$ (*J.N.T.U., B. Tech., 2006*)
 (*vi*) $3x + \sin x = e^x$.
5. Locate the root of $f(x) = x^{10} - 1 = 0$, between 0 and 1.3 using bisection method and method of false position. Comment on which method is preferable. (*Pune BVP., B. Tech., 2004*)
6. Find a root of the following equations correct to three decimal places by the secant method :
 (*i*) $x^3 + x^2 + x + 7 = 0$ (*ii*) $x - e^{-x} = 0$
 (*iii*) $x \log_{10} x = 1.9$.
7. Use the iteration method to find a root of the equations to four decimal places : (*Madras, B. Tech., 2006*)
 (*i*) $x^3 + x^2 - 100 = 0$ (*ii*) $x^3 - 9x + 1 = 0$
 (*iii*) $x = \frac{1}{2} + \sin x$ (*iv*) $\tan x = x$
 (*v*) $e^x = 5x$ (*vi*) $2^x - x - 3 = 0$ which lies between (– 3, – 2).
8. Evaluate $\sqrt{30}$ by (*i*) secant method (*ii*) iteration method correct to four decimal places.
9. Find the root of the equation $2x = \cos x + 3$ correct to three decimal places using (*i*) Iteration method, (*ii*) Aitken's $\Delta^2$ method.

**10.** Find the real root of the equation $x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} - \frac{x^{11}}{1320} + ... = 0.443$ correct to three decimal places using iteration method.

## 2.11. NEWTON-RAPHSON METHOD

Let $x_0$ be an approximate root of the equation $f(x) = 0$. If $x_1 = x_0 + h$ be the exact root, then $f(x_1) = 0$.

$\therefore$ Expanding $f(x_0 + h)$ by Taylor's series $f(x_0) + h f'(x_0) + \frac{h^2}{2!} f''(x_0) + ...... = 0$

Since $h$ is small, neglecting $h^2$ and higher powers of $h$, we get $f(x_0) + h f'(x_0) = 0$

or
$$h = -\frac{f(x_0)}{f'(x_0)} \qquad ...(1)$$

$\therefore$ A closer approximation to the root is given by

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

Similarly starting with $x_1$, a still better approximation $x_2$ is given by

$$x_2 = x_1 - \frac{f(x_0)}{f'(x_1)}.$$

In general,
$$\mathbf{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}} \quad (n = 0, 1, 2 ...) \qquad ...(2)$$

which is known as the *Newton-Raphson formula* or *Newton's iteration formula*.

**Obs. 1.** *Newton's method is useful in cases of large values of $f'(x)$ i.e.* when the graph of $f(x)$ while crossing the $x$-axis is nearly vertical.

For if $f'(x)$ is small in the vicinity of the root, then by (1), $h$ will be large and the computation of the root is slow or may not be possible. *Thus this method is not suitable in those cases where the graph of $f(x)$ is nearly horizontal while crossing the x-axis.*

**Obs. 2. Geometrical interpretation.** Let $x_0$ be a point near the root $\alpha$ of the equation $f(x) = 0$ (Fig. 2.7). Then the equation of the tangent at $A_0$ $[x_0, f(x_0)]$ is

$$y - f(x_0) = f'(x_0)(x - x_0).$$

It cuts the $x$-axis at
$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

which is a first approximation to the root $\alpha$. If $A_1$ is the point corresponding to $x_1$ on the curve, then the tangent at $A_1$ will cut the $x$-axis at $x_2$ which is nearer to $\alpha$ and is, therefore, a second approximation to the root. Repeating this process, we approach the root $\alpha$ quite rapidly. Hence the method consists in replacing the part of the curve between the point $A_0$ and the x-axis by means of the tangent to the curve at $A_0$.

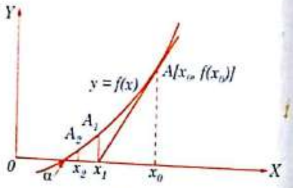

Fig. 2.7

**Obs. 3.** Newton's method is generally used to improve the result obtained by other methods. *It is applicable to the solution of both algebraic and transcendental equations.*



**Obs. 4.** *Newton's formula converges provided the initial approximation $x_0$ is chosen sufficiently close to the root.*

If it is not near the root, the procedure may lead to an endless cycle. A bad initial choice will lead one astray. *Thus a proper choice of the initial guess is very important for the success of the Newton's method.*

Comparing (2) with the relation $x_{n+1} = \phi(x_n)$ of the iteration method, we get

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

In general, $\phi(x) = x - \frac{f(x)}{f'(x)}$ which gives $\phi'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$.

Since the iteration method (§ 2.10) converges if $|\phi'(x)| < 1$

$\therefore$ Newton's formula will converge if $|f(x) f''(x)| < |f'(x)|^2$ in the interval considered.

Assuming $f(x)$, $f'(x)$ and $f''(x)$ to be continuous, we can select a small interval in the vicinity of the root $\alpha$, in which the above condition is satisfied. Hence the result.

Newtons method converges conditionally while Regula-falsi method always converges. However when once Newton-Raphson method converges, it converges faster and is preferred.

**Obs. 5. Newton's method has a quadratic convergence.**

Suppose $x_n$ differs from the root $\alpha$ by a small quantity $\varepsilon_n$ so that

$$x_0 = \alpha + \varepsilon_n \text{ and } x_{n+1} = \alpha + \varepsilon_{n+1}.$$

Then (2) becomes
$$\alpha + \varepsilon_{n+1} = \alpha + \varepsilon_n - \frac{f(\alpha + \varepsilon_n)}{f'(\alpha + \varepsilon_n)}$$

*i.e.*
$$\varepsilon_{n+1} = \varepsilon_n - \frac{f(\alpha + \varepsilon_n)}{f'(\alpha + \varepsilon_n)}$$

$$= \varepsilon_n - \frac{f(\alpha) + \varepsilon_n f'(\alpha) + \frac{1}{2!}\varepsilon_n^2 f''(\alpha) + ......}{f'(\alpha) + \varepsilon_n f''(\alpha) + ......}$$ by Taylor's expansion.

$$= \varepsilon_n - \frac{\varepsilon_n f'(\alpha) + \frac{1}{2}\varepsilon_n^2 f''(\alpha) + ......}{f'(\alpha) + \varepsilon_n f''(\alpha) + ......} = \frac{\varepsilon_n^2}{2}\frac{f''(\alpha)}{f'(\alpha)}. \qquad [\because f(\alpha) = 0]$$

This shows that the subsequent error at each step, is proportional to the square of the previous error and as such the convergence is quadratic. Thus Newton-Raphson method has **second order convergence.**

**Example 2.30.** *Find the positive root of $x^4 - x = 10$ correct to three decimal places, using Newton-Raphson method.* (J.N.T.U., B. Tech., 2008)

**Sol.** Let $f(x) = x^4 - x - 10$

so that $f(1) = -10 = -\text{ve}$, $f(2) = 16 - 2 - 10 = 4 = +\text{ ve}$.

$\therefore$ A root of $f(x) = 0$ lies between 1 and 2.

Let us take $x_0 = 2$

Also $f'(x) = 4x^3 - 1$

Newton-Raphson's formula is

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

Putting $n = 0$, the first approximation $x_1$ is given by

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 2 - \frac{f(2)}{f'(2)}$$

$$= 2 - \frac{4}{4 \times 2^3 - 1} = 2 - \frac{4}{31} = 1.871$$

Putting $n = 1$ in $(i)$, the second approximation is

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 1.871 - \frac{f(1.871)}{f'(1.871)}$$

$$= 1.871 - \frac{(1.871)^4 - (1.871) - 10}{4(1.871)^3 - 1}$$

$$= 1.871 - \frac{0.3835}{25.199} = 1.856$$

Putting $n = 2$ in $(ii)$, the third approximation is

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 1.856 - \frac{(1.856)^4 - (1.856) - 10}{4(1.856)^3 - 1}$$

$$= 1.856 - \frac{0.010}{24.574} = 1.856$$

Here $x_2 = x_3$. Hence the desired root is 1.856 correct to three decimal places.

**Example 2.31.** *Find by Newton's method, the real root of the equation $3x = \cos x + 1$, correct to four decimal places.* (V.T.U., B.E., 2009)

**Sol.** Let $f(x) = 3x - \cos x - 1$

$f(0) = -2 = -$ ve, $f(1) = 3 - 0.5403 - 1 = 1.4597 = +$ ve.

So a root of $f(x) = 0$ lies between 0 and 1. It is nearer to 1. Let us take $x_0 = 0.6$.

Also $f'(x) = 3 + \sin x$

$\therefore$ Newton's iteration formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{3x_n - \cos x_n - 1}{3 + \sin x_n}$$

$$= \frac{x_n \sin x_n + \cos x_n + 1}{3 + \sin x_n} \quad \ldots(i)$$

Putting $n = 0$, the first approximation $x_1$ is given by

$$x_1 = \frac{x_0 \sin x_0 + \cos x_0 + 1}{3 + \sin x_0} = \frac{(0.6) \sin (0.6) + \cos (0.6) + 1}{3 + \sin (0.6)}$$

$$= \frac{0.6 \times 0.5729 + 0.82533 + 1}{3 + 0.5729} = 0.6071$$

Putting $n = 1$ in $(i)$, the second approximation is

$$x_2 = \frac{x_1 \sin x_1 + \cos x_1 + 1}{3 + \sin x_1} = \frac{0.6071 \sin (0.6071) + \cos (0.6071) + 1}{3 + \sin (0.6071)}$$



$$= \frac{0.6071 \times 0.57049 + 0.8213 + 1}{3 + 0.57049} = 0.6071$$

Here $x_1 = x_2$. Hence the desired root is 0.6071 correct to four decimal places.

**Example 2.32.** *Using Newton's iterative method, find the real root of $x \log_{10} x = 1.2$ correct to five decimal places.* (Mumbai, B. Tech., 2004)

**Sol.** Let $f(x) = x \log_{10} x - 1.2$

$f(1) = -1.2 = -$ ve, $f(2) = 2 \log_{10} 2 - 1.2 = 0.59794 = -$ ve

and $f(3) = 3 \log_{10} 3 - 1.2 = 1.4314 - 1.2 = 0.23136 = +$ ve.

So a root of $f(x) = 0$ lies between 2 and 3. Let us take $x_0 = 2$.

Also $f'(x) = \log_{10} x + x \cdot \frac{1}{x} \log_{10} e = \log_{10} x + 0.43429.$

$\therefore$ Newton's iteration formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = \frac{0.43429 x_n + 1.2}{\log_{10} x_n + 0.43429} \quad \ldots(i)$$

Putting $n = 0$, the first approximation is

$$x_1 = \frac{0.43429 \times x_0 + 1.2}{\log_{10} x_0 + 0.43429} = \frac{0.43429 \times 2 + 1.2}{\log_{10} 2 + 0.43429}$$

$$= \frac{0.86858 \times 1.2}{0.30103 + 0.43429} = 2.81$$

Similarly putting $n = 1, 2, 3, 4$ in $(i)$, we get

$$x_2 = \frac{0.43429 \times 2.81 + 1.2}{\log_{10} 2.81 + 0.43429} = 2.741$$

$$x_3 = \frac{0.43429 \times 2.741 + 1.2}{\log_{10} 2.741 + 0.43429} = 2.74064$$

$$x_4 = \frac{0.43429 \times 2.741 + 1.2}{\log_{10} 2.74064 + 0.43429} = 2.74065$$

$$x_5 = \frac{0.43429 \times 2.7465 + 1.2}{\log_{10} 2.74065 + 0.43429} = 2.74065$$

Here $x_4 = x_5$. Hence the required root is 2.74065 correct to five decimal places.

## 2.12. SOME DEDUCTIONS FROM NEWTON-RAPHSON FORMULA

We can derive the following useful results from the Newton's iteration formula :

(1) *Iterative formula to find $1/N$ is* $\mathbf{x_{n+1} = x_n(2 - Nx_n)}$

(2) *Iterative formula to find $\sqrt{N}$ is* $\mathbf{x_{n+1} = \frac{1}{2}(x_n + N/x_n)}$

(3) *Iterative formula to find $1/\sqrt{N}$ is* $\mathbf{x_{n+1} = \frac{1}{2}(x_n + 1/Nx_n)}$

(4) *Iterative formula to find $\sqrt[k]{N}$ is* $\mathbf{x_{n+1} = \frac{1}{k}[(k-1)\, x_n + N/x_n^{k-1}]}$

*Proofs.* (1) Let $x = 1/N$ or $1/x - N = 0$

Taking $f(x) = 1/x - N$, we have $f'(x) = -x^{-2}$.

Then Newton's formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{(1/x_n - N)}{-x_n^{-2}} = x_n + \left(\frac{1}{x_n} - N\right)x_n^2.$$

$$= x_n + x_n - Nx_n^2 = x_n(2 - Nx_n)$$

(2) Let $x = \sqrt{N}$ or $x^2 - N = 0$.

Taking $f(x) = x^2 - N$, we have $f'(x) = 2x$.

Then Newton's formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - N}{2x_n} = \frac{1}{2}(x_n + N/x_n)$$

(3) Let $x = \frac{1}{\sqrt{N}}$ or $x^2 - \frac{1}{N} = 0$

Taking $f(x) = x^2 - 1/N$, we have $f'(x) = 2x$.

Then Newton's formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 1/N}{2x_n} = \frac{1}{2}\left(x_n + \frac{1}{Nx_n}\right)$$

(4) Let $x = \sqrt[k]{N}$ or $x^k - N = 0$

Taking $f(x) = x^k - N$, we have $f'(x) = kx^{k-1}$

Then Newton's formula gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^k - N}{kx_n^{k-1}} = \frac{1}{k}\left[(k-1)x_n + \frac{N}{x_n^{k-1}}\right].$$

**Example 2.33.** *Evaluate the following (correct to four decimal places) by Newton's iteration method :*

*(i)* $1/31$

*(ii)* $\sqrt{5}$

*(iii)* $1/\sqrt{14}$

*(iv)* $\sqrt[3]{24}$ *(Anna, B.E., 2007)*

*(v)* $(30)^{-1/5}$. *(Madras, B.E., 2003)*

**Sol.** *(i)* Taking $N = 31$, the above formula (1) becomes

$$x_{n+1} = x_n(2 - 31x_n)$$

Since an approximate value of $1/31 = 0.03$, we take $x_0 = 0.03$.

Then $x_1 = x_0(2 - 31x_0) = 0.03(2 - 31 \times 0.03) = 0.0321$

$x_2 = x_1(2 - 31x_1) = 0.0321(2 - 31 \times 0.0321) = 0.032257$

$x_3 = x_2(2 - 31x_2) = 0.032257(2 - 31 \times 0.032257) = 0.03226$

Since $x_2 = x_3$ upto 4 decimal places, we have $1/31 = 0.0323$.

*(ii)* Taking $N = 5$, the above formula (2), becomes $x_{n+1} = \frac{1}{2}(x_n + 5/x_n)$.

Since an approximate value of $\sqrt{5} = 2$, we take $x_0 = 2$.



Then $x_1 = \frac{1}{2}(x_0 + 5/x_0) = \frac{1}{2}(2 + 5/2) = 2.25$

$x_2 = \frac{1}{2}(x_1 + 5/x_1) = 2.2361$

$x_3 = \frac{1}{2}(x_2 + 5/x_2) = 2.2361$

Since $x_2 = x_3$ upto 4 decimal places, we have $\sqrt{5} = 2.2361$.

*(iii)* Taking $N = 14$, the above formula (3), becomes $x_{n+1} = \frac{1}{2}[x_n + 1/(14x_n)]$

Since an approximate value of $1/\sqrt{14} = 1/\sqrt{16} = \frac{1}{4} = 0.25$, we take $x_0 = 0.25$.

Then $x_1 = \frac{1}{2}[x_0 + (14x_0)^{-1}] = \frac{1}{2}[0.25 + (14 \times 0.25)^{-1}] = 0.26785$

$x_2 = \frac{1}{2}[x_1 + (14x_1)^{-1}] = \frac{1}{2}[0.26785 + (14 \times 0.26785)^{-1}] = 0.2672618$

$x_3 = \frac{1}{2}[x_2 + (14x_2)^{-1}] = \frac{1}{2}[0.2672618 + (14 \times 0.2672618)^{-1}] = 0.2672612$

Since $x_2 = x_3$ upto 4 decimal places, we take $1/\sqrt{14} = 0.2673$.

*(iv)* Taking $N = 24$ and $k = 3$, the above formula (4) becomes $x_{n+1} = \frac{1}{3}[2x_n + 24/x_n^2]$

Since an approximate value of $(24)^{1/3} = (27)^{1/3} = 3$, we take $x_0 = 3$.

Then $x_1 = \frac{1}{3}(2x_0 + 24/x_0^2) = \frac{1}{3}(6 + 24/9) = 2.88889$

$x_2 = \frac{1}{3}(2x_1 + 24/x_1^2) = \frac{1}{3}[(2 \times 2.88889) + 24/(2.88889)^2] = 2.88451$

$x_3 = \frac{1}{3}(2x_2 + 24/x_2^2) = \frac{1}{3}[2 \times 2.88451 + 24/(2.88451)^2] = 2.8845$

Since $x_2 = x_3$ upto 4 decimal places, we take $(24)^{1/3} = 2.8845$.

*(v)* Taking $N = 30$ and $k = -5$, the above formula (4) becomes

$$x_{n+1} = \frac{1}{-5}(6x_n + 30/x_n^{-6}) = \frac{x_n}{5}(6 - 30x_n^5)$$

Since an approximate value of $(30)^{-1/5} = (32)^{-1/5} = 1/2$, we take $x_0 = 1/2$

Then $x_1 = \frac{x_0}{5}(6 - 30x_0^5) = \frac{1}{10}(6 - 30/2^5) = 0.50625$

$x_2 = \frac{x_1}{5}(6 - 30x_1^5) = \frac{0.50625}{5}[6 - 30(0.50625)^5] = 0.506495$

$x_3 = \frac{x_2}{5}(6 - 30x_2^5) = \frac{0.506495}{5}[6 - 30(0.506495)^5] = 0.506496$

Since $x_2 = x_3$ upto 4 decimal places, we take $(30)^{-1/5} = 0.5065$.

## PROBLEMS 2.5

1. Find by Newton-Raphson method, a root of the following equations correct to 3 decimal places :
   (i) $x^3 - 3x + 1 = 0$ *(Bhopal, B.E., 2009)* (ii) $x^3 - 2x - 5 = 0$ *(P.T.U., B. Tech., 2005)*
   (iii) $x^3 - 5x + 3 = 0$ *(Mumbai, B.E., 2004)* (iv) $3x^3 - 9x^2 + 8 = 0$. *(Madras, B.E., 2003)*
2. Using Newton's iterative method, find a root of the following equations correct to 4 decimal places :
   (i) $x^4 + x^3 - 7x^2 - x + 5 = 0$ which lies between 2 and 3.
   (ii) $x^5 - 5x^2 + 3 = 0$.
3. Find the negative root of the equation $x^3 - 21x + 3500 = 0$ correct to 2 decimal places by Newton's method.
4. Using Newton-Raphson method, find a root of the following equations correct to 3 decimal places :
   (i) $x^2 + 4 \sin x = 0$ *(Hazaribagh, B.E., 2009)*
   (ii) $x \tan x + 1 = 0$ *(J.N.T.U., B. Tech., 2006)*
   (iii) $e^x = x^3 + \cos 25x$ which is near 4.5. *(V.T.U., B. Tech., 2007)*
   (iv) $x \log_{10} x = 12.34$, start with $x_0 = 10$. *(Anna, B. Tech., 2004)*
   (v) $\cos x = xe^x$ *(J.N.T.U., B. Tech., 2009)* (vi) $e^x \sin x = 1$.
5. The bacteria concentration in a reservoir varies as $C = 4e^{-2t} + e^{-0.1t}$. Using Newton Raphson method, calculate the time required for the bacteria concentration to be 0.5.
6. Use Newton's method to find the smallest root of the equation $e^x \sin x = 1$ to four places of decimal.
7. The current $i$ in an electric circuit is given by $i = 10e^{-t} \sin 2\pi t$ where $t$ is in seconds. Using Newton's method, find the value of $t$ correct to 3 decimal places for $i = 2$ amp.
8. Find the iterative formulae for finding $\sqrt{N}$, $\sqrt[3]{N}$ where $N$ is a real number, using Newton-Raphson formula.
   Hence evaluate : (a) $\sqrt{10}$. *(J.N.T.U., B. Tech., 2008)*
   (b) $\sqrt{21}$ *(U.P.T.U., MCA, 2009)*
   (c) the cube-root of 41 to three places of decimal. *(Madras B.E., 2003)*
9. Develop an algorithm using N.R. method, to find the fourth root of a positive number $N$ and hence find $\sqrt[4]{32}$.
10. Evaluate the following (correct to 3 decimal places) by using the Newton-Raphson method. *(W.B.T.U., B. Tech., 2005)*
    (i) 1/18 *(J.N.T.U., B. Tech., 2004)* (ii) $1/\sqrt{15}$ (iii) $(28)^{-1/4}$.
11. *Obtain Newton-Raphson extended formula*

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} - \frac{1}{2}\frac{[f(x_0)]^2 f''(x_0)}{\{f'(x_0)\}^3}$$

*for the root of the equation $f(x) = 0$.*

Hence find the root of the equation $\cos x = xe^x$ correct to five decimal places.

[**Sol.** Expanding $f(x)$ in the neighbourhood of $x_0$ by Taylor's series ; we have

$$0 = f(x) = f(x_0 + \overline{x - x_0}) = f(x_0) + (x - x_0) f'(x_0) \text{ to first approx.}$$

Hence the first approximation to the root is given by

$$x_1 - x_0 = -f(x_0)/f'(x_0) \qquad \ldots(i)$$



Again by Taylor's series to the second approximation, we get

$$f(x_1) = f(x_0) + (x_1 - x_0) f'(x_0) + \frac{1}{2!}(x_1 - x_0)^2 f''(x_0)$$

Since $x_1$ is an approximation to the root, $f(x_1) = 0$

$$\therefore \quad f(x_0) + (x_1 - x_0) f'(x_0) + \tfrac{1}{2}(x_1 - x_0)^2 f''(x_0) = 0$$

or

$$x_1 - x_0 = -\frac{f(x_0)}{f'(x_0)} - \frac{1}{2}\left\{\frac{-f(x_0)}{f'(x_0)}\right\}^2 f''(x_0) \qquad \text{[by } (i)]$$

whence follows the desired formula. This is known as **Chebyshev formula** of third order.]

## 2.13. MULLER'S METHOD

**(1)** This method is a generalisation of the secant method as it doesn't require the derivative of the function. It is an iterative method that requires three starting points. Here, $y = f(x)$ is approximated by a second degree parabola passing through these three points $(x_{i-2}, y_{i-2})$, $(x_{i-1}, y_{i-1})$ and $(x_i, y_i)$ in the vicinity of the root. Then a root of this quadratic is taken as the next approximation $x_{i+1}$ to the root of $f(x) = 0$.

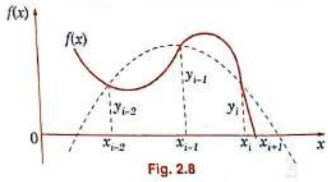

Fig. 2.8

**(2)** Let $x_{i-2}, x_{i-1}, x_i$ be three approximations to the root $\alpha$ of the equation $f(x) = 0$ and $y_{i-2}, y_{i-1}, y_i$ be the corresponding values of $f(x)$.

Assuming the equation of the parabola through the points $(x_{i-2}, y_{i-2})$, $(x_{i-1}, y_{i-1})$ and $(x_i, y_i)$ to be

$$y = ax^2 + bx + c, \qquad \ldots(1)$$

we get

$$\left.\begin{aligned} y_{i-2} &= ax_{i-2}^2 + bx_{i-2} + c \\ y_{i-1} &= ax_{i-1}^2 + bx_{i-1} + c \\ y_i &= ax_i^2 + bx_i + c \end{aligned}\right\} \qquad \ldots(2)$$

and

Eliminating $a$, $b$, $c$ from (1) and (2), we obtain

$$\begin{vmatrix} y & x^2 & x & 1 \\ y_{i-2} & x_{i-2}^2 & x_{i-2} & 1 \\ y_{i-1} & x_{i-1}^2 & x_{i-1} & 1 \\ y_i & x_i^2 & x_i & 1 \end{vmatrix} = 0$$

which can be written as

$$y = \frac{(x - x_{i-1})(x - x_i)}{(x_{i-2} - x_{i-1})(x_{i-2} - x_i)} y_{i-2} + \frac{(x - x_{i-2})(x - x_i)}{(x_{i-1} - x_{i-2})(x_{i-1} - x_i)} y_{i-1} + \frac{(x - x_{i-2})(x - x_{i-1})}{(x_i - x_{i-2})(x_i - x_{i-1})} y_i \qquad \ldots(3)$$

We now define

$$\lambda = \frac{x - x_i}{x_i - x_{i-1}}, \quad \lambda_i = \frac{x_i - x_{i-1}}{x_{i-1} - x_{i-2}} \quad \text{and} \quad \delta_i = \frac{x_i - x_{i-2}}{x_{i-1} - x_{i-2}} \qquad \ldots(4)$$

Then (3) simplifies to

$$y = \frac{(y_{i-2}\lambda_i + y_{i-1}\delta_i + y_i)\lambda_i\lambda^2}{\delta_i} + \frac{y_{i-2}\lambda_i^2 - y_{i-1}\delta_i^2 + y_i(\lambda_i + \delta_i)}{\delta_i}\lambda + y_i \quad ...(5)$$

From (4), we get $x = x_i + \lambda(x_i - x_{i-1})$ ...(6)

Now to find a better approximation to the root, we need the unknown quantity $\lambda$. To determine $\lambda$, we put $y = 0$ in (5) giving

$$(y_{i-2}\lambda_i - y_{i-1}\delta_i + y_i)\lambda_i\lambda^2 + \mu_i\lambda + \delta_i y_i = 0 \quad ...(7)$$

where $\mu_i = y_{i-2}\lambda_i^2 - y_{i-1}\delta_i^2 + y_i(\lambda_i + \delta_i)$.

Dividing throughout by $\lambda_i\lambda^2$ and solving for $1/\lambda$*, we get

$$\frac{1}{\lambda} = \frac{-\mu_i \pm \sqrt{[\mu_i^2 - 4y_i\delta_i\lambda_i(y_{i-2}\lambda_i - y_{i-1}\delta_i + y_i)]}}{2y_i\delta_i}$$

Since $x$ is close to $x_i$, $\lambda$ should be small in magnitude. Therefore the sign should be so chosen to make the numerator largest in magnitude. Then (6) gives a better approximation to the root.

**Obs.** *This method is iterative and converges for almost all initial approximations quadratically. In case no better approximations are known, we take,* $x_{i-2} = -1$, $x_{i-1} = 0$ *and* $x_i = 1$.

**Example 2.34.** *Apply Muller's method to find the root of the equation* $\cos x = xe^x$ *which lies between 0 and 1.*

**Sol.** Let $y = \cos x - xe^x$

Taking the initial approximations as

$$x_{i-2} = -1, x_{i-1} = 0, x_i = 1$$

We obtain

$$y_{i-2} = \cos 1 + e^{-1}, y_{i-1} = 1, y_i = \cos 1 - e$$

and

$$\lambda = x - 1, \lambda_i = 1, \delta_i = 2$$

$$\mu_i = (\cos 1 + e^{-1}) - 4 + 3(\cos 1 - e).$$

$\therefore$ From (7), we get two values of $\lambda^{-1}$. ...(i)

We choose the –ve sign so that the numerator in (i) is largest in magnitude and obtain $\lambda = -0.5585$.

$\therefore$ The next approximation to the root is given by (6) as

$$x_{i+1} = x_i + \lambda(x_i - x_{i-1}) = 1 - 0.5585 = 0.4415.$$

Repeating the above process, we get

$$x_{i+2} = 0.5125, x_{i+3} = 0.5177, x_{i+4} = 0.5177$$

Hence the root is 0.518 correct to 3 decimal places.

## PROBLEMS 2.6

*Using Muller's method, find a root of the following equations, correct to three decimal places :*

1. $x^3 - 2x - 1 = 0$ (*U.P.T.U., B. Tech., 2006*) 2. $x^3 - x^2 - x - 1 = 0$
3. $x^3 + 2x^2 + 10x - 20 = 0$ taking $x_0 = 0, x_1 = 1$ and $x_2 = 2$.
4. $\log x = x - 3$ taking $x_0 = 0.25, x_1 = 0.5$ and $x_2 = 1$.

*As a direct solution of (7) usually leads to inaccurate results, we solve it for $1/\lambda$.



## 2.14. ROOTS OF POLYNOMIALS EQUATIONS

The methods so far discussed for finding the roots of equations can also be applied to polynomials. These methods, however, do not work well when the polynomial equations contain multiple or complex roots. We now discuss methods for finding all the real and complex roots of polynomials. These methods are specially designed for polynomials and cannot be applied to transcendental equations. We begin with Horner's method which is the best for finding the approximate values of real roots of a numerical polynomial equation.

## 2.15. APPROXIMATE SOLUTION OF POLYNOMIAL EQUATIONS—HORNER'S METHOD

This method consists in diminution of the roots of an equation by successive digits occurring in the roots.

If the root of an equation lies between $a$ and $a + 1$, then the value of this root will be $a.bcd$......, where $b, c, d$...... are digits in its decimal part. To obtain these, we proceed as follows :

(*i*) Diminish the roots of the given equation by $a$ so that the root of the new equation is $o.\ bcd$......

(*ii*) Then multiply the roots of the transformed equation by 10 so that the root of the new equation is $b\ .\ cd$......

(*iii*) Now diminish the root by $b$ and multiply the roots of the resulting equation by 10 so that the root is $c.d$......

(*iv*) Next diminish the root by $c$ and so on. By continuing this process, the root may be evaluated to any desired degree of accuracy digit by digit. The method will be clear from the following example :

**Example 2.35.** *Find by Horner's method, the positive root of the equation* $x^3 + x^2 + x - 100 = 0$ *correct to three decimal places.*

**Sol.** *Step I.* Let $f(x) = x^3 + x^2 + x - 100$

By Descartes' rule of signs, there is only one positive root. Also $f(4) = -$ ve and $f(5) =$ +ve, therefore, the root lies between 4 and 5.

*Step II.* Diminish the roots of given equation by 4 so that the transformed equation is

$$x^3 + 13x^2 + 57x - 16 = 0. \quad ...(i)$$

Its root lies between 0 and 1. (We draw a zig-zag line above the set of figures 13, 57, – 16 which are the coefficients of the terms in (*i*) as shown below.) Now multiply the roots of (*i*) by 10 for which attach one zero to the second term, two zeros to the third term and three zeros to the fourth term. Then we get the equation

$$f_1(x) = x^3 + 130x^2 + 5700x - 16000 = 0 \quad ...(ii)$$

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 1 | − 100 | (4.264 |
| | 4 | 20 | 84 | |
| | 5 | 21 | − 16000 | |
| | 4 | 36 | 11928 | |
| | 9 | 5700 | − 4072000 | |
| | 4 | 264 | 3788376 | |
| | 130 | 5964 | −283624000 | |
| | 2 | 268 | | |
| | 132 | 623200 | | |
| | 2 | 8196 | | |
| | 134 | 631396 | | |
| | 2 | 8232 | | |
| | 1360 | 63962800 | | |
| | 6 | | | |
| | 1366 | | | |
| | 6 | | | |
| | 1372 | | | |
| | 6 | | | |
| | 13780 | | | |

Its root lies between 0 and 10.

Clearly $f_1(2) = -\text{ve}, f_1(3) = +\text{ve}$.

$\therefore$ The root of (*ii*) lies between 2 and 3 *i.e.* first figure after decimal is 2.

*Step III.* Diminish the roots of $f_1(x) = 0$ by 2 so that the next transformed equation is

$$x^3 + 136x^2 + 6232x - 4072 = 0. \qquad \ldots(iii)$$

Its root lies between 0 and 1. (We draw the second zig-zag line above the set of figures 136, 6232, – 4072). Multiply the roots of (*iii*) by 10, *i.e.* attach one zero to second term, two zeros to third term and three zeros to the fourth term. Then the new equation is

$$f_2(x) = x^3 + 1360x^2 + 623200x - 4072000 = 0$$

Its root lies between 0 and 10, which is nearly $= \dfrac{4072000}{623200} = 6.$

Hence second figure after decimal place is 6.

*Step IV.* Diminish the roots of $f_2(x) = 0$ by 6, so that the transformed equation is

$$x^3 + 1378x^2 + 639628x - 283624 = 0.$$

Its root lies between 0 and 1. (We draw the third zig-zag line above the set of figures 1378, 639628, – 283624.) As before multiply its roots by 10, *i.e.* attach one zero to the second term, two zeros to the third term and three zeros to the fourth term. Then the equation becomes

$$f_3(x) = x^3 + 13780x^2 + 63962800x - 283624000 = 0$$

Its root lies between 0 and 10, which is nearly $= \dfrac{283624000}{63962800} = 4$. Thus the roots of $f_3(x) = 0$ are to be diminished by 4 *i.e.* the third figure after decimal place is 4. But there is no need to proceed further as the root is required correct to three decimal places only.

Hence the root is 4.264.



**Obs. 1.** *After two steps of diminishing, we apply the* **principle of trial divisor** *in which we divide the last coefficient by last but one coefficient to get the next integer by which the roots are to be diminished.* These last two coefficients should have opposite signs.

**Obs. 2.** *At any stage if the trial divisor suggests the next integer to be zero, then we should again multiply the roots by 10 and write zero in decimal place of the root.*

**Example 2.36.** *Find the cube root of 30 correct to 3 decimal places, using Horner's method.*

**Sol.** *Step I.* Let $x = \sqrt[3]{30}$ *i.e.* $f(x) = x^3 - 30 = 0$

Now $f(3) = -3\ (-\text{ve}), \quad f(4) = 34\ (+\text{ve})$

$\therefore$ the root lies between 3 and 4.

*Step II.* Diminish the roots of the given equation by 3 so that the transformed equation is

$$x^3 + 9x^2 + 27x - 3 = 0 \qquad \ldots(i)$$

Its roots lies between 0 and 1. (We draw a zig-zag line above the set of numbers 9, 27, – 3 which are the coefficients of the terms in (*i*)). Now multiply the roots of (*i*) by 10 for which attach one zero to the second term, two zeros to the third term and three zeros to the fourth term. Then we get the equation

$$f_1(x) = x^3 + 90x^2 + 2700x - 3000 = 0 \qquad \ldots(ii)$$

Its roots lies between 0 and 10.

Clearly $f_1(1) = -\text{ve}, \quad f_2(2) = +\text{ve}$

$\therefore$ The root of (*ii*) lies between 1 and 2 *i.e.* first figure after decimal place is 1.

*Step III.* Diminish the roots of $f_1(x) = 0$ by 1, so that the next transformed equation is

$$x^3 + 93x^2 + 2883x - 209 = 0 \qquad \ldots(iii)$$

Its root lies between 0 and 1. (We draw a second zig-zag line above the set of figures 93, 2883, – 209). Multiply the roots of (*iii*) by 10 *i.e.* attach one zero to second term, two zeros to third term and three zeros to the fourth term. Then the new equation is

$$f_2(x) = x^3 + 930x^2 + 288300x - 209000 = 0$$

Its root lies between 0 and 10, which is nearly = 209000/288300 = 0.724 > 0 and < 1.

Hence second figure after decimal place is 0.

| | | | | |
|---|---|---|---|---|
| 1 | 0 | 0 | − 30 | (3.107 |
| | 3 | 9 | 27 | |
| | 3 | 9 | − 30000 | |
| | 3 | 18 | 2791 | |
| | 6 | 2700 | − 209000000 | |
| | 3 | 91 | | |
| | 90 | 2791 | | |
| | 1 | 92 | | |
| | 91 | 28830000 | | |
| | 1 | | | |
| | 92 | | | |
| | 1 | | | |
| | 9300 | | | |

*Step IV.* Diminish the root of $f_2(x) = 0$ by 0 and then multiply its roots by 10 so that

$$f_3(x) = x^3 + 9300x^2 + 28830000x - 209000000 = 0$$

Its root lies between 0 and 10, which is nearly

$$= 209000000/28830000 = 7.2 > 7 \text{ and } < 8.$$

Thus the roots of $f_3(x) = 0$ are to be diminished by 7 *i.e.* the third figure after decimal is 7.

Hence the required root is 3.107.

## PROBLEMS 2.7

1. Find by Horner's method, the root (correct to three decimal places) of the equations
(*i*) $x^3 - 3x + 1 = 0$ which lies between 1 and 2. (*ii*) $x^3 + x - 1 = 0$
(*iii*) $x^3 - 3x^2 + 2.5 = 0$ which lies between 1 and 2. (*Madras B.E., 2000*)
2. Using Horner's method, find the largest real root of $x^3 - 4x + 2 = 0$ correct to three decimal places.
3. Show that a root of the equation $x^4 + x^3 - 4x^2 - 16 = 0$ lies between 2 and 3. Find its value correct upto two decimal places by Horner's method.
4. Find the negative root of the equation $x^3 - 9x^2 + 18 = 0$ correct to two decimal places by Horner's method.
5. Find the cube root of 25, correct to four decimal places, using Horner's method.

## 2.16. (1) MULTIPLE ROOTS

If $\alpha$ is a root of $f(x) = 0$ of order $m$, then $f(\alpha) = 0, f'(\alpha) = 0, \ldots\ldots, f^{m-1}(\alpha) = 0$ and $f^m(\alpha) \neq 0$. Such on equation can be written as $f(x) = (x - \alpha)^m \phi(x) = 0$. In other words, if $\alpha$ is a root of $f(x) = 0$ repeated $m$ times, then it is also a root of $f'(x) = 0$ repeated $(m - 1)$ times, of $f''(x) = 0$ repeated $(m - 2)$ times and so on.

**(2) Multiple roots by Newton's method.** Let $\alpha$ be a root of the polynomial equation $f(x) = 0$ which is repeated $m$ times, If $x_0, x_1, x_2, \ldots\ldots, x_{n+1}$, be its successive approximations then on the lines of Newton's iterative method, we have $\mathbf{x_{n+1} = x_n - m\dfrac{f(x_n)}{f'(x_n)}}$

which is called the *generalised Newton's formula*. It reduces to Newton-Raphson formula for $m = 1$.

**Obs. 1.** If initial approximation $x_0$ is sufficiently close to the root $\alpha$, then the expressions

$$x_0 - m\frac{f(x_0)}{f'(x_0)},\ x_0 - (m-1)\frac{f'(x_0)}{f''(x_0)},\ x_0 - (m-2)\frac{f''(x_0)}{f'''(x_0)},\ \ldots\ldots$$

will have the same value.

**Obs. 2.** *Generalised Newton's formula has a* **second order convergence** *for determining a multiple root.* (see Example 2.38).

**Example 2.37.** *Find the double root of the equation $x^3 - x^2 - x + 1 = 0$.*

**Sol.** Let $f(x) = x^3 - x^2 - x + 1$

so that $f'(x) = 3x^2 - 2x - 1,\ f''(x) = 6x - 2$



Starting with $x_0 = 0.9$, we have

$$x_0 - 2\frac{f(x_0)}{f'(x_0)} = 0.9 - \frac{2 \times 0.019}{-0.37} = 1.003$$

and

$$x_0 - (2-1)\frac{f'(x_0)}{f''(x_0)} = 0.9 - \frac{(-0.37)}{3.4} = 1.009$$

The closeness of these values implies that there is a double root near $x = 1$.

$\therefore$ Choosing $x_1 = 1.01$ for the next approximation, we get

$$x_1 - 2\frac{f(x_1)}{f'(x_1)} = 1.01 - \frac{2 \times 0.0002}{0.0403} = 1.0001$$

$$x_1 - (2-1)\frac{f'(x_1)}{f''(x_1)} = 1.01 - \frac{0.0403}{4.06} = 1.0001$$

This shows that there is a double root at $x = 1.0001$ which is quite near the actual root $x = 1$.

**Example 2.38.** *Show that the generalised Newton's formula $x_{n+1} = x_n - 2f(x_n)/f'(x_n)$ gives a quadratic convergence when the equation $f(x) = 0$ has a pair of double roots in the neighbourhood of $x = x_n$.*

**Sol.** Suppose $x = \alpha$ is a double root near $x = x_n$.

Then $$f(\alpha) = 0, f'(\alpha) = 0 \quad \ldots(i)$$

We have $$\varepsilon_{n+1} = \varepsilon_n - \frac{2f(\alpha + \varepsilon_n)}{f'(\alpha + \varepsilon_n)}$$

Expanding $f(\alpha + \varepsilon)$ and $f'(\alpha + \varepsilon)$ in powers of $\varepsilon_n$ and using (*i*), we get

$$\varepsilon_{n+1} = \varepsilon_n - \frac{2\left[\frac{\varepsilon_n^2}{2!}f''(\alpha) + \frac{\varepsilon_n^2}{3!}f'''(\alpha) + \ldots\ldots\right]}{\left[\varepsilon_n f''(\alpha) + \frac{\varepsilon_n^2}{2!}f'''(\alpha) + \ldots\ldots\right]}$$

$$= \varepsilon_n - \frac{\varepsilon_n\left[f''(\alpha) + \frac{1}{3}\varepsilon_n f'''(\alpha)\right]}{f''(\alpha) + \frac{\varepsilon_n}{2}f'''(\alpha)} \text{ approx.}$$

$$= \frac{1}{6}\varepsilon_n^2 \frac{f'''(\alpha)}{f''(\alpha) + \frac{\varepsilon_n}{2}f'''(\alpha)} = \frac{1}{6}\varepsilon_n^2 \frac{f'''(\alpha)}{f''(\alpha)}$$

which shows that $\varepsilon_{n+1} \propto \varepsilon_n^2$ and so the convergence is of second order.

## 2.17. COMPLEX ROOTS

We know that the complex roots of an equation occur in conjugate pairs *i.e.* if $\alpha + i\beta$ is a root of $f(x) = 0$, $\alpha - i\beta$ is also its root. In other words, $[x - (\alpha + i\beta)]$ and $[x - (\alpha - i\beta)]$ are factors of $f(x)$ or $(x - \alpha - i\beta)(x - \alpha + i\beta) = x^2 - 2x\alpha + \alpha^2 + \beta^2$ is a factor of $f(x)$. This implies that we should try to isolate complex roots by finding the appropriate quadratic factors of the original polynomial. A method which is often used for finding such quadratic factors of polynomials is Lin-Bairstow's method. However Newton's method can also be used to find the complex roots of a polynomial equation which we illustrate below :

**Example 2.39.** *Solve $x^4 - 5x^3 + 20x^2 - 40x + 60 = 0$, by Newton's method given that all the roots of the given equation are complex.*

**Sol.** Let $f(x) = x^4 - 5x^3 + 20x^2 - 40x + 60 = 0$ ...(i)

so that $f'(x) = 4x^3 - 15x^2 + 40x - 40$

$\therefore$ Newton-Raphson method gives

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^4 - 5x_n^3 + 20x_n^2 - 40x_n + 60}{4x_n^3 - 15x_n^2 + 40x_n - 40}$$

$$= \frac{3x_n^4 - 10x_n^3 + 20x_n^2 - 60}{4x_n^3 - 15x_n^2 + 40x_n - 40}$$

Putting $n = 0$ and taking $x_0 = 2(1 + i)$ by trial, we get

$$x_1 = \frac{3(2+2i)^4 - 10(2+2i)^3 + 20(2+2i)^2 - 60}{4(2+2i)^3 - 15(2+2i)^2 + 40(2+2i) - 40} = 1.92\,(1 + i)$$

Similarly

$$x_2 = \frac{3(1.92+1.92i)^4 - 10(1.92+1.92i)^3 + 20(1.92+1.92i) - 60}{4(1.92+1.92i)^3 - 15(1.92+1.92i)^2 + 40(1.92+1.92i) - 40} = 1.915 + 1.908i$$

Since complex roots occur in conjugate pairs so the roots of (i) are $1.915 \pm 1.908i$ up to 3 places of decimals. Assuming that the other pair of roots of (i) is $\alpha \pm i\beta$, we have

Sum of the roots $= (\alpha + i\beta) + (\alpha - i\beta) + (1.915 + 1.908i) + (1.915 - 1.908i) = 5$

*i.e.* $2\alpha + 3.83 = 5$ or $\alpha = 0.585$

Also the product of roots $= (\alpha^2 + \beta^2)\{(1.915)^2 + (1.908)^2\} = 60$

which gives $\beta = 2.805$. Hence the other two roots are $0.585 \pm 2.805i$.

## 2.18 LIN-BAIRSTOW'S METHOD

This method is often used for finding the complex roots of a polynomial equation with real coefficients, such as

$$f(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots\ldots + a_{n-1} x + a_n = 0 \quad \ldots(1)$$

Since complex roots occur in pairs as $\alpha \pm i\beta$, each pair corresponds to a quadratic factor

$$\{x - (\alpha + i\beta)\}\{x - (\alpha - i\beta)\} = x^2 - 2\alpha x + \alpha^2 + \beta^2$$

which is of the form $x^2 + px + q$

If we divide $f(x)$ by $x^2 + px + q$, we obtain the quotient $Q_{n-2} = x^{n-2} + b_1 x^{n-3} + \ldots\ldots + b_{n-2}$ and the remainder $R_n = Rx + S$.

Thus $f(x) = (x^2 + px + q)(x^{n-2} + b_1 x^{n-3} + \ldots\ldots + b_{n-2}) + Rx + S$ ...(2)

If $x^2 + px + q$ divides $f(x)$ completely, the remainder $Rx + S = 0$ *i.e.* $R = 0$, $S = 0$. Obviously $R$ and $S$ both depend upon $p$ and $q$. So our problem is to find $p$ and $q$ such that

$$R\,(p, q) = 0,\; S\,(p, q) = 0. \quad \ldots(3)$$

Let $p + \Delta p$, $q + \Delta q$ be the actual values of $p$ and $q$ which satisfy (3). Then

$$R\,(p + \Delta p, q + \Delta q) = 0,\; S\,(p + \Delta p, q + \Delta q) = 0 \quad \ldots(4)$$



To find the corrections $\Delta p$, $\Delta q$, we expand these by Taylor's series and neglect second and higher order terms.

$$\therefore \left.\begin{aligned} R\,(p, q) + \frac{\partial R}{\partial p}\Delta p + \frac{\partial R}{\partial q}\Delta q = 0 \\ S\,(p, q) + \frac{\partial S}{\partial p}\Delta p + \frac{\partial S}{\partial q}\Delta q = 0 \end{aligned}\right\} \quad \ldots(5)$$

We solve these simultaneous equations for $\Delta p$ and $\Delta q$ and then the procedure is repeated with the corrected values for $p$ and $q$. Now to compute the coefficients $b_i$, $R$ and $S$, we compare the coefficients of like powers of $x$ in (2) giving

$$b_1 = a_1 - p$$
$$b_2 = a_2 - pb_1 - q$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$b_i = a_i - pb_{i-1} - qb_{i-2} \quad \ldots(6)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$R = a_{n-1} - pb_{n-2} - qb_{n-3}, \quad S = a_n - q\,b_{n-2}$$

We now introduce $b_{n-1}$ and $b_n$ and define

$$b_i = a_i - p\,b_{i-1} - q\,b_{i-2},\ i = 1, 2, \ldots\ldots n \quad \ldots(7)$$

where $b_0 = 1$, $b_{-1} = 0 = b_{-2}$

Comparing the last two equations with those of (6), we get

$$b_{n-1} = a_{n-1} - p\,b_{n-2} - q\,b_{n-3} = R$$
$$b_n = a_n - p\,b_{n-1} - q\,b_{n-2} = S - p\,b_{n-1}$$

giving $R = b_{n-1}$ and $S = b_n + p\,b_{n-1}$ ...(8)

Substituting these values in (5), we get

$$b_{n-1} + \frac{\partial b_{n-1}}{\partial p}\Delta p + \frac{\partial b_{n-1}}{\partial p}\Delta q = 0$$

$$b_n + pb_{n-1} + \left(\frac{\partial b_n}{\partial p} + p\frac{\partial b_{n-1}}{\partial p} + b_{n-1}\right)\Delta p + \left(\frac{\partial b_n}{\partial q} + p\frac{\partial b_{n-1}}{\partial q}\right)\Delta q = 0$$

Multiplying the first of these equations by $p$ and subtracting from the second, we get

$$\left.\begin{aligned} \frac{\partial b_{n-1}}{\partial p}\Delta p + \frac{\partial b_{n-1}}{\partial q}\Delta q + b_{n-1} = 0 \\ \left(\frac{\partial b_n}{\partial p} + b_{n-1}\right)\Delta p + \frac{\partial b_n}{\partial q}\Delta q + b_n = 0 \end{aligned}\right\} \quad \ldots(9)$$

Now differentiating (7) w.r.t. $p$ and $q$ partially and noting that all $a_i$'s are constants and all $b_i$'s are functions of $p$ and $q$, we have

$$\left.\begin{aligned} \frac{\partial b_i}{\partial p} = -b_{i-1} - p\frac{\partial b_{i-1}}{\partial p} - q\frac{\partial b_{i-2}}{\partial p}\,;\ \frac{\partial b_{-1}}{\partial p} = 0 = \frac{\partial b_{-2}}{\partial p} \\ \frac{\partial b_i}{\partial q} = -b_{i-2} - p\frac{\partial b_{i-1}}{\partial q} - q\frac{\partial b_{i-2}}{\partial q}\,;\ \frac{\partial b_1}{\partial q} = 0 = \frac{\partial b_{-2}}{\partial q} \end{aligned}\right\} \quad \ldots(10)$$

Also from (6), we get

$$\frac{\partial b_0}{\partial p} = 0 = \frac{\partial b_1}{\partial q}\ ;\ \frac{\partial b_1}{\partial p} = -b_0,\ \frac{\partial b_2}{\partial q} = -b_0 - p\,\frac{\partial b_1}{\partial q} = -b_0$$

$$\frac{\partial b_2}{\partial p} = -b_1 - p\,\frac{\partial b_1}{\partial p} = -b_1 + pb_0$$

$$\frac{\partial b_3}{\partial q} = -b_1 - p\,\frac{\partial b_2}{\partial q} - p\,\frac{\partial b_1}{\partial q} = -b_1 + pb_0$$

Thus we have $\dfrac{\partial b_2}{\partial q} = \dfrac{\partial b_1}{\partial p}$ and $\dfrac{\partial b_3}{\partial q} = \dfrac{\partial b_2}{\partial p}$

By mathematical induction, we shall prove that $\dfrac{\partial b_{i+1}}{\partial q} = \dfrac{\partial b_i}{\partial p}$, for all $i$

Let the result be true for $i = r$, then $\quad \dfrac{\partial b_{r+1}}{\partial q} = \dfrac{\partial b_r}{\partial p}$ ...(11)

But using (10)

$$\frac{\partial b_{r+2}}{\partial q} = -b_r - p\,\frac{\partial b_{r+1}}{\partial q} - q\,\frac{\partial b_r}{\partial q}$$

and

$$\frac{\partial b_{r+1}}{\partial p} = -b_r - p\,\frac{\partial b_r}{\partial p} - q\,\frac{\partial b_{r-1}}{\partial p} = -b_r - p\,\frac{\partial b_{r+1}}{\partial q} - q\,\frac{\partial b_r}{\partial q} \qquad \text{[by (11)]}$$

This shows that $\dfrac{\partial b_{r+2}}{\partial q} = \dfrac{\partial b_{r+1}}{\partial q}$ i.e. the result is true for $i = r + 1$. But it is for $i = 1$ and 2. Hence by induction, it is true for all values of $i$.

Now writing $\dfrac{\partial b_{i+1}}{\partial q} = \dfrac{\partial b_i}{\partial p} = -c_{i-1}$, $i = 0, 1, 2, \ldots, n-1$ ...(12)

the equations (10) can be expressed as

$$c_{i-1} = b_{i-1} - p\,c_{i-2} - q\,c_{i-3},\ c_{i-2} = b_{i-2} - p\,c_{i-3} - q\,c_{i-4}$$

These can be compressed into a single equation

$$c_i = b_i - p\,c_{i-1} - q\,c_{i-2}$$

with

$$c_0 = 0,\ c_{-1} = 0,\quad i = 1, 2, \ldots, (n-1)$$

Thus $c_i$ is computed from $b_i$ in exactly the same way as $b_i$ from $a_i$ in (7). ...(13)

Differentiating the relations in (8) and using (12), we get

$$\frac{\partial R}{\partial p} = \frac{\partial b_{n-1}}{\partial p} = -c_{n-2},\ \frac{\partial R}{\partial q} = \frac{\partial b_{n-1}}{\partial q} = -c_{n-3}$$

and

$$\frac{\partial S}{\partial p} = \frac{\partial b_n}{\partial p} + b_{n-1} + p\,\frac{\partial b_{n-1}}{\partial p} = -c_{n-1} - pc_{n-2} + b_{n-1},$$

$$\frac{\partial S}{\partial q} = \frac{\partial b_n}{\partial q} + p\,\frac{\partial b_{n-1}}{\partial q} = -c_{n-2} - pc_{n-3}$$



Substituting these in (5), we get

$$b_{n-1} - c_{n-2}\,\Delta p - c_{n-3}\,\Delta q = 0$$

and

$$b_n + pb_{n-1} + (-c_{n-1} - pc_{n-2} + b_{n-1})\,\Delta p + (-c_{n-2} - pc_{n-3})\,\Delta q = 0$$

or

$$\left.\begin{aligned} c_{n-2}\,\Delta p + c_{n-3}\,\Delta q &= b_{n-1} \\ (c_{n-1} - b_{n-1})\,\Delta p + c_{n-2}\,\Delta q &= b_n \end{aligned}\right\} \qquad \text{...(14)}$$

After finding the values of $b_i$'s and $c_i$'s from (7) and (13) and putting in (14), we obtain the approximate values of $\Delta p$ and $\Delta q$ say $\Delta p_0$ and $\Delta q_0$. If $p_0$, $q_0$ be the initial approximations then their improved values are $p_1 = p_0 + \Delta p_0$, $q_1 = q_0 + \Delta q_0$. Now taking $p_1$ and $q_1$ as the initial values and repeating the process, we can get better values of $p$ and $q$.

**Obs.** The values of $b_i$'s and $c_i$'s are found by the following (synthetic division) scheme :

| | | | | | | |
|---|---|---|---|---|---|---|
| $a_0(=1)$ | $a_1$ | $a_2$ | $a_3 \ldots\ldots a_{n-2}$ | $a_{n-1}$ | $a_n$ | |
| | $-pb_0$ | $-pb_1$ | $-pb_2 \ldots\ldots -pb_{n-3}$ | $-pb_{n-2}$ | $-pb_{n-1}$ | $-p$ |
| | | $-qb_0$ | $-qb_1 \ldots\ldots -qb_{n-4}$ | $-qb_{n-3}$ | $-qb_{n-2}$ | $-q$ |
| $b_0(=1)$ | $b_1$ | $b_2$ | $b_3 \ldots\ldots b_{n-2}$ | $b_{n-1}$ | $b_n$ | |
| | $-pc_0$ | $-pc_1$ | $-pc_2 \ldots\ldots -pc_{n-3}$ | $-pc_{n-2}$ | | $-p$ |
| | | $-qc_0$ | $-qc_1 \ldots\ldots -qc_{n-4}$ | $-qc_{n-3}$ | | $-q$ |
| $c_0(=1)$ | $c_1$ | $c_2$ | $c_3 \ldots\ldots$ | $c_{n-2}$ | $c_{n-1}$ | |

**Example 2.40.** *Solve $x^4 - 5x^3 + 20x^2 - 40x + 60 = 0$, given that all the roots of $f(x) = 0$ are complex, by using Lin-Bairstow method.*

**Sol.** Starting with the values $p_0 = -4$, $q_0 = 8$, we have

| | | | | | |
|---|---|---|---|---|---|
| 1 | $-5$ | 20 | $-40$ | 60 | |
| — | 4 | $-4$ | 32 | 0 | 4 |
| — | — | $-8$ | 8 | $-64$ | $-8$ |
| 1 | $-1$ | 8 | $0(=b_{n-1})$ | $-4(=b_n)$ | |
| | 4 | 12 | 48 | | 4 |
| | | $-8$ | $-24$ | | $-8$ |
| 1 | $3(=c_{n-3})$ | $12(=c_{n-2})$ | $24(=c_{n-1})$ | | |

$$\therefore \quad c_{n-1} - b_{n-1} = 24 - 0 = 24$$

Corrections $\Delta p_0$ and $\Delta q_0$ are given by

$$c_{n-2}\,\Delta p_0 + c_{n-3}\,\Delta q_0 = b_{n-1} \quad \text{i.e.} \quad 12\,\Delta p_0 + 3\,\Delta q_0 = 0$$

$$(c_{n-1} - b_{n-1})\,\Delta p_0 + c_{n-2}\,\Delta q_0 = b_n \quad \text{i.e.} \quad 24\,\Delta p_0 + 12\,\Delta q_0 = -4$$

Solving, we get $\Delta p_0 = 0.1667$, $\Delta q_0 = -0.6667$

$$\therefore \quad p_1 = p_0 + \Delta p_0 = -3.8333$$

$$q_1 = q_0 + \Delta q_0 = 7.3333$$

Now repeating the same process *i.e.* dividing $f(x)$ by $x^2 - 3.8333x + 7.3333$, we get

| | | | | | |
|---|---|---|---|---|---|
| 1 | $-5$ | 20 | $-40$ | 60 | 3.8333 |
| | 3.8333 | $-4.4723$ | 31.4116 | $-0.125$ | $-7.3333$ |
| | | $-7.3333$ | 8.5558 | $-60.092$ | |
| 1 | $-1.1667$ | 8.1944 | $-0.0326$ $(= b_{n-1})$ | $-0.217$ $(= b_n)$ | |
| | 3.8333 | 10.2219 | 42.4845 | | 3.8333 |
| | | $-7.3333$ | $-19.555$ | | $-7.3333$ |
| 1 | 2.6666 $(= c_{n-3})$ | 11.083 $(= c_{n-2})$ | 22.8969 $(= c_{n-1})$ | | |

$\therefore \quad c_{n-1} - b_{n-1} = 22.8969 + 0.0326 = 22.9295$

Corrections $\Delta p_1$ and $\Delta q_1$ are given by

Lin-Bairstow

$$11.083\, \Delta p_1 + 2.6666\, \Delta q_1 = -0.0326$$
$$22.9295\, \Delta p_1 + 11.083\, \Delta q_1 = -0.217$$

Solving, we get $\Delta p_1 = 0.0033$ and $\Delta q_1 = -0.0269$

$$p_2 = p_1 + \Delta p_1 = -3.83,\ q_2 = q_1 + \Delta q_1 = 7.3064.$$

So one of the quadratic factors of $f(x)$ is

$$x^2 - 3.83\,x + 7.3064. \qquad ...(i)$$

If $\alpha \pm i\beta$ be its roots, then $2\alpha = 3.83$, $\alpha^2 + \beta^2 = 7.3064$ giving $\alpha = 1.9149$ and $\beta = 1.9077$.

Hence a pair of roots is $1.9149 \pm 1.9077\, i$.

To find the remaining two roots of $f(x) = 0$, we divide $f(x)$ by $(i)$ as follows [by § 2.5 (3)] :

| | | | | | |
|---|---|---|---|---|---|
| 1 | $-5$ | 20 | $-40$ | 60 | 3.83 |
| | 3.83 | $-4.4811$ | 31.4539 | | $-7.3064$ |
| | | $-7.3064$ | 8.5485 | $-60.0038$ | |
| 1 | $-1.17$ | 8.2125 | 0.0024 $\approx 0$ | $-0.0038$ $\approx 0$ | |

$\therefore$ The other quadratic factor is $x^2 - 1.17x + 8.2125$

If $\gamma \pm i\delta$ be its roots, then $2\delta = 1.17$, $\gamma^2 + \delta^2 = 8.2125$ giving $\gamma = 0.585$ and $\delta = 2.8054$.

Hence the other pair of roots is $0.585 \pm 2.8054\, i$.

## 2.19. GRAEFFE'S ROOT SQUARING METHOD

This method has an advantage over the other methods that it does not require any prior information about the roots. But it is applicable to polynomial equations only and is capable of giving all the roots. Consider the polynomial equation

$$x^n + a_1 x^{n-1} + a_2 x^{n-2} + \ldots\ldots + a_{n-1} x + a_n = 0 \qquad ...(1)$$

Separating the even and odd powers of $x$ and squaring, we get

$$(x^n + a_2 x^{n-2} + a_4 x^{n-4} + \ldots\ldots)^2 = (a_1 x^{n-1} + a_3 x^{n-3} + \ldots\ldots)^2$$



Putting $x^2 = y$ and simplifying, the new equation becomes

$$y^n + b_1 y^{n-1} + \ldots\ldots + b_{n-1} y + b_n = 0 \qquad ...(2)$$

where

$$\left.\begin{aligned} b_1 &= -a_1^2 + 2a_2 \\ b_2 &= a_2^2 - 2a_1 a_3 + 2a_4 \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ b_n &= (-1)^n a_n^2 \end{aligned}\right\} \qquad ...(3)$$

If $\alpha_1, \alpha_2, \ldots\ldots \alpha_n$ be the roots of (1) then the roots of (2) are $\alpha_1^2, \alpha_2^2, \ldots\ldots \alpha_n^2$.

After $m$ squarings, let the new transformed equation be

$$z^n + c_1 z^{n-1} + \ldots\ldots + c_{n-1} z + c_n = 0$$

whose roots $\gamma_1, \gamma_2, \ldots\ldots, \gamma_n$ are such that $\gamma_i = \alpha_i^{2m}$, $i = 1, 2, \ldots\ldots n$. ...(4)

Assuming that $|\alpha_1| > |\alpha_2| > \ldots\ldots > |\alpha_n|$, then $|\gamma_1| >> |\gamma_2| >> \ldots\ldots >> |\gamma_n|$

where >> stands for 'much greater than'.

Thus $\dfrac{|\gamma_2|}{|\gamma_1|} = \dfrac{\gamma_2}{\gamma_1}, \ldots\ldots, \dfrac{|\gamma_n|}{|\gamma_{n-1}|} = \dfrac{\gamma_n}{\gamma_{n-1}}$ are negligible as compared to unity. ...(5)

Also $\gamma_i$ being an even power of $\alpha_i$ is always positive.

$\therefore$ From (4), we have

$$\Sigma\gamma_1 = -c_1 \quad i.e. \quad c_1 = -\gamma_1\left(1 + \frac{\gamma_2}{\gamma_1} + \frac{\gamma_3}{\gamma_1} + \ldots\right)$$

$$\Sigma\gamma_1\gamma_2 = c_2 \quad i.e. \quad c_2 = \gamma_1\gamma_2\left(1 + \frac{\gamma_3}{\gamma_1} + \ldots\right)$$

$$\Sigma\gamma_1\gamma_2\gamma_3 = -c_3 \quad i.e. \quad c_3 = -\gamma_1\gamma_2\gamma_3\left(1 + \frac{\gamma_4}{\gamma_1} + \ldots\right)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\gamma_1\gamma_2 \ldots \gamma_n = (-1)^n c_n \ i.e.\ c_n = (-1)^n \gamma_1\gamma_2 \ldots \gamma_n.$$

Hence by (5), we get $c_1 \approx -\gamma_1$, $c_2 \approx \gamma_1\gamma_2$, $c_3 \approx \gamma_1\gamma_2\gamma_3, \ldots\ldots$

*i.e.* $\qquad \gamma_1 \approx -c_1, \quad \gamma_2 \approx -c_2/c_1, \quad \gamma_3 \approx -c_3/c_2, \ldots\ldots, \gamma_n \approx -c_n/c_{n-1}$

Now since $\gamma_1 = \alpha_i^{2m}$, $\quad \therefore \quad \alpha_i = (\gamma_i)^{1/2m} = |c_i/c_{i-1}|$ ...(6)

Thus we can determine $\alpha_1, \alpha_2, \ldots\ldots \alpha_n$, the roots of (1).

**Obs. 1.** *Double root.* If the magnitude of $c_i$ is half the square of the magnitude of the corresponding coefficient in the previous equation after a few squarings, then it shows that $\alpha_i$ is a double root of (1). We find this double root as follows :

$$\gamma_k \approx \frac{c_k}{c_{k-1}} \quad \text{and} \quad \gamma_{k+1} \approx -\frac{c_{k+1}}{c_k}$$

$$\therefore \quad \gamma_k\gamma_{k+1} \approx \gamma_k^2 \approx \left|\frac{c_{k+1}}{c_{k-1}}\right| \quad i.e. \quad \alpha_k^{2m} = \gamma_k^2 = \left|\frac{c_{k+1}}{c_{k-1}}\right| \qquad ...(7)$$

This gives the magnitude of the double root and substituting in (1), we can find its sign.

**Obs. 2.** *Complex roots.* If $\alpha_r$ and $\alpha_{r+1}$ form a complex pair $\rho_r e^{\pm i\phi_r}$, then the coefficients of $x^{n-r}$ in successive squarings would fluctuate both in magnitude and sign by an amount $2\rho_r^m \cos m\phi_r$.

For $m$ sufficiently large $\rho_r$ and $\phi_r$ can be determined by

$$\rho_r^{2(2^m)} = \left|\frac{c_{r+1}}{c_{r-1}}\right|, 2\rho_r^m \cos m\phi_r = \frac{c_{r+1}}{c_{r-1}} \quad \ldots(8)$$

If (1) has only one pair of complex roots say : $\rho_r e^{\pm i\phi_r} = \xi + i\eta$, then we can find all the real roots. Thereafter $\xi$ is given by

$$\alpha_1 + \alpha_2 + \ldots + \alpha_{r-1} + 2\xi + \alpha_{r+2} + \ldots + \alpha_n = -a_1 \quad \ldots(9)$$

and $\eta$ is given by

$$\rho_r^2 = \xi^2 + \eta^2 \quad \text{or} \quad \eta = \sqrt{(\rho_r^2 - \xi^2)} \quad \ldots(10)$$

**Example 2.41.** *Find all roots of the equation $x^3 - 2x^2 - 5x + 6 = 0$ by Graeffe's method, squaring thrice.* (*Madras, B.E., 2000*)

**Sol.** Let

$$f(x) = x^3 - 2x^2 - 5x + 6 = 0 \quad \ldots(i)$$

$+ \quad - \quad - \quad +$

By Descartes rule of signs, there being two changes of sign, (*i*) has two positive roots.

Also

$$f(-x) = -x^3 - 2x^2 + 5x + 6$$

$- \quad - \quad + \quad +$

*i.e.* one change in sign, there is one negative root.

Rewriting (*i*) as $x^3 - 5x = 2x^2 - 6$ and squaring, we get $y(y-5)^2 = (2y-6)^2$ where $y = x^2$

or
$$y(y^2 + 49) = 14y^2 + 36 \quad \ldots(ii)$$

Squaring again and putting $y^2 = z$, we obtain $z(z+49)^2 = (14z + 36)^2$

or
$$z(z^2 + 1393) = 98z^2 + 1296 \quad \ldots(iii)$$

Squaring once again and putting $z^2 = u$, we get $u(u + 1393)^2 = (98u + 1296)^2$

or
$$u^3 - 6818u^2 + 1686433u - 1679616 = 0 \quad \ldots(iv)$$

If the roots of (*iv*) are $\gamma_1, \gamma_2, \gamma_3$, then $\gamma_1 = -c_1 = 6818$,

$$\gamma_2 = -\frac{c_2}{c_1} = \frac{1686433}{6818} = 247.3501.$$

$$\gamma_3 = -\frac{c_3}{c_2} = \frac{1679616}{1686433} = 0.996$$

If $\alpha_1, \alpha_2, \alpha_3$ be the roots of (*i*), then

$$|\alpha_1| = (\gamma_1)^{1/8} = 3.014443 \approx 3$$
$$|\alpha_2| = (\gamma_2)^{1/8} = 1.991425 \approx 2$$
$$|\alpha_3| = (\gamma_3)^{1/8} = 0.999499 \approx 1$$

The sign of a root is found by substituting the root in $f(x) = 0$. We find $f(3) = 0$, $f(-2) = 0$, $f(1) = 0$.

Hence the roots are 3, – 2, 1.

**Example 2.42.** *Apply Graeffe's method to find all the roots of the equation $x^4 - 3x + 1 = 0$.* (*Madras, B.E., 1995*)

**Sol.** We have

$$f(x) = x^4 - 3x + 1 = 0 \quad \ldots(i)$$

$+ \quad - \quad +$

$\therefore$ There being two changes in sign, (*i*) has two positive real roots and no negative real root.

Thus the remaining two roots are complex.



1.2

Rewriting (*i*) as $x^4 + 1 = 3x$, and squaring, we get $(y^2 + 1)^2 = 9y$ where $y = x^2$.

Squaring again and putting $y^2 = z$, we obtain

$$(z+1)^4 = 81z \text{ or, } z^4 + 4z^3 + 6z^2 - 77z + 1 = 0$$

or
$$z^4 + 6z^2 + 1 = -z(4z^2 - 77) \quad \ldots(ii)$$

real root 1

Squaring once again and putting $z^2 = u$, we get $(u^2 + 6u + 1)^2 = u(4u - 77)^2$

or
$$u^4 - 4u^3 + 654u^2 - 5917u + 1 = 0 \quad \ldots(iii)$$

If $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ be the roots of (*i*), then the roots of (*iii*) are $\alpha_1^8, \alpha_2^8, \alpha_3^8, \alpha_4^8$. Thus (*iii*) gives

$$\alpha_1^8 = 4 \qquad i.e.\ \alpha_1 = 1.1892$$

$$\alpha_2^8 = \frac{654}{4} \doteq 163.5 \qquad i.e.\ \alpha_2 = 1.891$$

$$\alpha_3^8 = \frac{5917}{654} = 9.0474 \qquad i.e.\ \alpha_3 = 1.3169$$

$$\alpha_4^8 = \frac{1}{5917} = 0.00017 \qquad i.e.\ \alpha_4 = 0.3379$$

From (*ii*) and (*iii*), we observe that the magnitudes of the coefficients $c_1$ and $c_4$ have become constant. This indicates that $\alpha_1$ and $\alpha_4$ are the real roots whereas $\alpha_2$ and $\alpha_3$ are a pair of complex roots.

$\therefore$ The real roots $\alpha_1 = 1.1892$ and $\alpha_4 = 0.3379$.

Now let us find the complex roots $\rho_2 e^{\pm i\phi_2} = \xi + i\eta$.

From (*iii*), its magnitude is given by

$$\rho_2^{2(2^3)} = \frac{c_{2+1}}{c_{2-1}} \quad \text{or} \quad \rho_2^{16} = \frac{5917}{4} = 1479.25$$

whence $\rho_2 = 1.5781$

Also from (*i*), $\alpha_1 + 2\xi + \alpha_4 = 0$

This gives $\quad \xi = -\frac{1}{2}(\alpha_1 + \alpha_4) = -0.7636$ and $\eta = \sqrt{(\rho_2^2 - \xi^2)} = 1.381$

Hence the complex roots are $-0.7636 \pm 1.381\,i$.

## PROBLEMS 2.8

1. Find a double root of the equation $x^3 - 5x^2 + 8x - 4 = 0$ which is near 1.8.
2. Find the multiplicity and the multiple root of the equation $x^4 - 11x^3 + 36x^2 - 16x - 64 = 0$ which is near 3.9.
3. Apply Newton's method to find a pair of complex roots of the equation $x^4 + x^3 + 5x^2 + 4x + 4 = 0$ starting with $x_0 = i$.
4. Apply Lin-Bairstow method to find a quadratic factor of the equation $x^4 + 5x^3 + 3x^2 - 5x - 9$ close to $x^2 + 3x - 5$. (*Madras MCA, 1997 S*)
5. Find the roots of the equation $x^4 + 9x^3 + 36x^2 + 51x + 27 = 0$ to three decimal places using Bairstow iterative method.
6. Find the quadratic factors of the equation $x^4 - 8x^3 + 39x^2 - 62x + 50 = 0$ by using Lin-Bairstow method (upto third iteration) starting with $p_0 = 0, q_0 = 0$.

7. Solve $x^3 - 8x^2 + 17x - 10 = 0$ by Graeffe's method. (*Madras, B.E., 1996*)
8. Apply Graeffe's method to find all the roots of the equation $x^3 - 6x^2 + 11x - 6 = 0$ (*Madras, B.E., 1997*)
9. Solve the equation $x^3 - 5x^2 - 17x + 20 = 0$ by Graeffe's method, squaring three times.
10. Find all the roots of the equation $x^3 - 4x^2 + 5x - 2 = 0$ by Graeffe's method, squaring thrice. (*Madras, B.E., 1998*)
11. Determine all roots of the equation $x^3 - 9x^2 + 18x - 6 = 0$ by Graeffe's method. (*Madras B.E., 2001*)

## 2.20. COMPARISON OF ITERATIVE METHODS

1. Convergence in the case of *the bisection method* is slow but steady. It is, however, the simplest method and it never fails.

2. *The method of false position* is slow and it is first order convergent. Convergence however, is guaranteed. Most often, it is found superior to the bisection method.

3. *The secant method* is not guaranteed to converge. But its order of convergence being 1.62, it converges faster than the method of false position. This method is considered most economical giving reasonably rapid convergence at a low cost.

4. Of all the above methods, *Newton-Raphson method* has the fastest rate of convergence. The method is quite sensitive to the starting value. Also it may diverge if $f'(x)$ is near zero during the iterative cycle.

5. For locating the complex roots, *Newton's method* can be used. *Muller's method* is also effective for finding complex roots.

6. If all the roots of the given equation are required then the *Lin-Bairstow method* is recommended. After a quadratic factor has been found, then Lin-Bairstow method must be applied on the reduced polynomial. If the location of some roots is known, first find these roots to a desired accuracy and then apply the Lin-Bairstow method on the reduced polynomial.

7. If the roots of the given polynomial are real and distinct then *Graeffe's root squaring method* is quite useful.

## 2.21. OBJECTIVE TYPE OF QUESTIONS

### PROBLEMS 2.9

*Select the correct answer or fill up the blanks in the following questions :*

1. The order of convergence in Newton-Raphson method is
   (*a*) 2 (*b*) 3
   (*c*) 0 (*d*) none.
2. The Newton-Raphson algorithm for finding the cube root of $N$ is ............ .
3. The bisection method for finding the roots of an equation $f(x) = 0$ is ............ .
4. In Regula-falsi method, the first approximation is given by ............ .
5. If $f(x) = 0$ is an algebraic equation, the Newton-Raphson method is given by $x_{n+1} = x_n - f(x_n)/?$
   (*a*) $f(x_{n-1})$ (*b*) $f'(x_{n-1})$
   (*c*) $f'(x_n)$ (*d*) $f''(x_n)$.



6. In the Regula-falsi method of finding the real root of an equation, the curve AB is replaced by ...... .
7. Newton's iterative formula to find the value of $\sqrt{N}$ is ........ .
8. A root of $x^3 - x + 4 = 0$ obtained using bisection method correct to two places, is ........ .
9. Newton-Raphson formula converges when ........ .
10. In the case of bisection method, the convergence is
    (*a*) linear (*b*) quadratic (*c*) very slow
11. Out of method of false position and Newton-Raphson method, the rate of convergence is faster for ............ .
12. Using Newton's method, the root of $x^3 = 5x - 3$ between 0 and 1 correct to two decimal places, is ........ .
13. The Newton-Raphson method fails when
    (*a*) $f'(x)$ is negative (*b*) $f'(x)$ is too large
    (*c*) $f'(x)$ is zero (*d*) Never fails.
14. The condition for the convergence of the iteration method for solving $x = \phi(x)$ is ........ .
15. While finding a root of an equation by Regula-falsi method, the number of iterations can be reduced ........ .
16. Newton's method is useful when the graph of the function while crossing the $x$-axis is nearly vertical. (True or False)
17. Difference between Transcendental equation and polynomial equation is ........ .
18. The interval in which a real root of the equation $x^3 - 2x - 5 = 0$ lies is ........ .
19. The iterative formula for finding the reciprocal of $N$ is $x_{n+1}$ = ........ .
20. While finding the root of an equation by the method of false position, the number of iterations can be reduced ...... .

# 3

# SOLUTION OF SIMULTANEOUS ALGEBRAIC EQUATIONS



## 3.1. INTRODUCTION TO DETERMINANTS

**(1) Definition.** The expression $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}$ is called a *determinant of the second order* and stands for '$a_1b_2 - a_2b_1$'. It contains 4 numbers $a_1, b_1, a_2, b_2$ (called *elements*) which are arranged along two horizontal lines (called *rows*) and two vertical lines (called *columns*).

Similarly
$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \qquad ...(i)$$
is called a *determinant of the third order*. It consists of 9 elements which are arranged in 3 rows and 3 columns.

In general, a *determinant of the nth order* is of the form

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \cdots\cdots a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots\cdots a_{2n} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & a_{n3} \cdots\cdots a_{nn} \end{vmatrix}$$

which is a block of $n^2$ *elements* in the form of a square along $n$ *rows* and $n$ *columns*. The diagonal through the left hand top corner which contains the elements $a_{11}, a_{22}, a_{33}, \ldots\ldots, a_{nn}$ is called the *leading diagonal*.





**(2) Expansion of a determinant.** *The* **cofactor** *of an element in a determinant is the determinant obtained by deleting the row and the column which intersect at that element, with the proper sign.* The sign of an element in the $i$th row and $j$th column is $(-1)^{i+j}$. The cofactor of an element is usually denoted by the corresponding capital letter.

For instance, the cofactor of $b_3$ in (i) is $B_3 = (-1)^{3+2} \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$

*A determinant can be expanded in terms of any row or column as follows :*

*Multiply each element of the row (or column) in terms of which we intend expanding the determinant, by its cofactor and then add up all these products.*

$\therefore$ Expanding (i) by $R_1$ (i.e. 1st row),

$$\Delta = a_1A_1 + b_1B_1 + c_1C_1$$
$$= a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}$$
$$= a_1 (b_2c_3 - b_3c_2) - b_1 (a_2c_3 - a_3c_2) + c_1 (a_2b_3 - a_3b_2)$$

Similarly expanding by $C_2$ (i.e. 2nd column),

$$\Delta = b_1B_1 + b_2B_2 + b_3B_3$$
$$= -b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + b_2 \begin{vmatrix} a_1 & c_1 \\ a_3 & c_3 \end{vmatrix} - b_3 \begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}$$
$$= b_1(a_2c_3 - a_3c_2) + b_2(a_1c_3 - a_3c_1) - b_3(a_1c_2 - a_2c_1).$$

**Example 3.1.** *Find the value of* $\Delta = \begin{vmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \end{vmatrix}$

**Sol.** Since there are two zeros in the second row, therefore, expanding by $R_2$, we get

$$\Delta = -\begin{vmatrix} 1 & 2 & 3 \\ 3 & 0 & 1 \\ 0 & 1 & 2 \end{vmatrix} + 0 - 3 \begin{vmatrix} 0 & 1 & 3 \\ 2 & 3 & 1 \\ 3 & 0 & 2 \end{vmatrix} + 0$$

(Expand By $C_1$) (Expand by $R_1$)

$$= -[1(0 \times 2 - 1 \times 1) - 3(2 \times 2 - 1 \times 3) + 0]$$
$$- 3[0 - (2 \times 2 - 3 \times 1) + 3(2 \times 0 - 3 \times 3)]$$
$$= -(-1 - 3) - 3(-1 - 27) = 4 + 84 = 88.$$

**(3) Basic properties.** The following properties enable us to simplify and evaluate a given determinant without expanding it :

I. *A determinant remains unaltered by changing its rows into columns and columns into rows.*

II. *If two parallel lines of a determinant are interchanged, the determinant retains its numerical value but changes in sign.*

III. *A determinant vanishes if two of its parallel lines are identical.*

IV. *If each element of a line be multiplied by the same factor, the whole determinant is multiplied by that factor.*

V. *If each element of a line consists of m terms, the determinant can be expressed as the sum of m determinants.*

*VI. If to each element of a line be added equi-multiples of the corresponding elements of one or more parallel lines, the determinant remains unaltered.*

For instance $\begin{vmatrix} a_1 + pb_1 - qc_1 & b_1 & c_1 \\ a_2 + pb_2 - qc_2 & b_2 & c_2 \\ a_3 + pb_3 - qc_3 & b_3 & c_3 \end{vmatrix}$

$$= \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} + p \begin{vmatrix} b_1 & b_1 & c_1 \\ b_2 & b_2 & c_2 \\ b_3 & b_3 & c_3 \end{vmatrix} - q \begin{vmatrix} c_1 & b_1 & c_1 \\ c_2 & b_2 & c_2 \\ c_3 & b_3 & c_3 \end{vmatrix}$$

$$= \Delta + 0 + 0 = \Delta \qquad \text{[From } (iv)]$$

**(4) Rule for multiplication of determinants :**

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \times \begin{vmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{vmatrix}$$

$$= \begin{vmatrix} a_1l_1 + b_1m_1 + c_1n_1 & a_2l_1 + b_2m_1 + c_2n_1 & a_3l_1 + b_3m_1 + c_3n_1 \\ a_1l_2 + b_1m_2 + c_1n_2 & a_2l_2 + b_2m_2 + c_2n_2 & a_3l_2 + b_3m_2 + c_3n_2 \\ a_1l_3 + b_1m_3 + c_1n_3 & a_2l_3 + b_2m_3 + c_2n_3 & a_3l_3 + b_3m_3 + c_3n_3 \end{vmatrix}$$

*i.e. the product of two determinants of the same order is itself a determinant of that order.*

**Example 3.2.** *If* $\begin{vmatrix} a & a^2 & a^3 - 1 \\ b & b^2 & b^3 - 1 \\ c & c^2 & c^3 - 1 \end{vmatrix} = 0$ *in which a, b, c are different, show that abc = 1.*

**Sol.** As each term of $C_3$ in the given determinant consists of two terms, we express it as a sum of two determinants.

$$\begin{vmatrix} a & a^2 & a^3 - 1 \\ b & b^2 & b^3 - 1 \\ c & c^2 & c^3 - 1 \end{vmatrix} = \begin{vmatrix} a & a^2 & a^3 \\ b & b^2 & b^3 \\ c & c^2 & c^3 \end{vmatrix} + \begin{vmatrix} a & a^2 & -1 \\ b & b^2 & -1 \\ c & c^2 & -1 \end{vmatrix}$$

$$= abc \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} - \begin{vmatrix} a & a^2 & 1 \\ b & b^2 & 1 \\ c & c^2 & 1 \end{vmatrix}$$

[Taking common $a$, $b$, $c$ from $R_1$, $R_2$, $R_3$ respectively of the first determinant and $-1$ om $C_3$ of the second determinant.]

$$= abc \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} - \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix}$$

[Passing $C_3$ over $C_2$ and $C_1$ in the second determinant]

$$\therefore \quad \begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} (abc - 1) = 0$$



Hence $abc = 1$, since $\begin{vmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{vmatrix} \neq 0$ as $a, b, c$ are all different.

**Example 3.3.** *Solve the equation* $\begin{vmatrix} x + 2 & 2x + 3 & 3x + 4 \\ 2x + 3 & 3x + 4 & 4x + 5 \\ 3x + 5 & 5x + 8 & 10x + 17 \end{vmatrix} = 0$

**Sol.** Operating $R_3 - (R_1 + R_2)$, we get

$$\begin{vmatrix} x + 2 & 2x + 3 & 3x + 4 \\ 2x + 3 & 3x + 4 & 4x + 5 \\ 0 & 1 & 3x + 8 \end{vmatrix} = 0 \qquad \text{(Operate } R_2 - R_1 \text{ and } R_1 + R_2)$$

or $$\begin{vmatrix} x + 2 & 2x + 4 & 6x + 12 \\ x + 1 & x + 1 & x + 1 \\ 0 & 1 & 3x + 8 \end{vmatrix} = 0$$

or $$(x + 1)(x + 2) \begin{vmatrix} 1 & 2 & 6 \\ 1 & 1 & 1 \\ 0 & 1 & 3x + 8 \end{vmatrix} = 0$$

To bring one more zero in $C_1$, operate $R_1 - R_2$.

$$\therefore \quad (x + 1)(x + 2) \begin{vmatrix} 0 & 1 & 5 \\ 1 & 1 & 1 \\ 0 & 1 & 3x + 8 \end{vmatrix} = 0$$

Now expand by $C_1$.

$\therefore \quad -(x + 1)(x + 2)(3x + 8 - 5) = 0$ or $-3(x + 1)(x + 2)(x + 1) = 0$.

Thus $\quad x = -1, -1, -2.$

**Example 3.4.** *Prove that* $\begin{vmatrix} 1 + a & 1 & 1 & 1 \\ 1 & 1 + b & 1 & 1 \\ 1 & 1 & 1 + c & 1 \\ 1 & 1 & 1 & 1 + d \end{vmatrix} = abc\left(1 + \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right)$

**Sol.** Let $\Delta$ be the given determinant.

Taking $a$, $b$, $c$, $d$ common from $R_1$, $R_2$, $R_3$, $R_4$ respectively, we get

$$\Delta = abcd \begin{vmatrix} a^{-1} + 1 & a^{-1} & a^{-1} & a^{-1} \\ b^{-1} & b^{-1} + 1 & b^{-1} & b^{-1} \\ c^{-1} & c^{-1} & c^{-1} + 1 & c^{-1} \\ d^{-1} & d^{-1} & d^{-1} & d^{-1} + 1 \end{vmatrix}$$

[Operate $R_1 + (R_2 + R_3 + R_4)$ and take out the common factor from $R_1$]

$$= abcd\,(1 + a^{-1} + b^{-1} + c^{-1} + d^{-1}) \begin{vmatrix} 1 & 1 & 1 & 1 \\ b^{-1} & b^{-1} + 1 & b^{-1} & b^{-1} \\ c^{-1} & c^{-1} & c^{-1} + 1 & c^{-1} \\ d^{-1} & d^{-1} & d^{-1} & d^{-1} + 1 \end{vmatrix}$$

[Operate $C_2 - C_1$, $C_3 - C_1$, $C_4 - C_1$]

$$= abcd\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right)\begin{vmatrix} 1 & 0 & 0 & 0 \\ b^{-1} & 1 & 0 & 0 \\ c^{-1} & 0 & 1 & 0 \\ d^{-1} & 0 & 0 & 1 \end{vmatrix}$$

$$= abcd\left(1+\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right).$$

■ **Example 3.5.** *Evaluate* $\begin{vmatrix} a^2+\lambda^2 & ab+c\lambda & ca-b\lambda \\ ab-c\lambda & b^2+\lambda^2 & bc+a\lambda \\ ca+b\lambda & bc-a\lambda & c^2+\lambda^2 \end{vmatrix} \times \begin{vmatrix} \lambda & c & -b \\ -c & \lambda & a \\ b & -a & \lambda \end{vmatrix}$

**Sol.** By the rule of multiplication of determinants, the resulting determinant

$$\Delta = \begin{vmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{vmatrix}$$

where $d_{11} = (a^2+\lambda^2)\lambda + (ab+c\lambda)c + (ca-b\lambda)(-b) = \lambda(a^2+b^2+c^2+\lambda^2)$

$d_{12} = (a^2+\lambda^2)(-c) + (ab+c\lambda)\lambda + (ca-b\lambda)a = 0,\ d_{13} = 0,$

$d_{21} = 0,\ d_{22} = \lambda(a^2+b^2+c^2+\lambda^2),\ d_{23} = 0.$

$d_{31} = 0,\ d_{32} = 0,\ d_{33} = \lambda(a^2+b^2+c^2+\lambda^2).$

Hence
$$\Delta = \begin{vmatrix} \lambda(a^2+b^2+c^2+\lambda^2) & 0 & 0 \\ 0 & \lambda(a^2+b^2+c^2+\lambda^2) & 0 \\ 0 & 0 & \lambda(a^2+b^2+c^2+\lambda^2) \end{vmatrix}$$

$$= \lambda^3(a^2+b^2+c^2+\lambda^2)^3.$$

## PROBLEMS 3.1

1. If $\begin{vmatrix} x & x^2 & 1+x^3 \\ y & y^2 & 1+y^3 \\ z & z^2 & 1+z^3 \end{vmatrix} = 0$, then prove, without expansion, that $xyz = -1$ where $x, y, z$ are unequal.

2. Evaluate $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 3 & 4 \\ 1 & 2 & 4 & 4 \\ 1 & 2 & 3 & 5 \end{vmatrix}$

*Prove the following results : (3 to 7)*

3. $\begin{vmatrix} a+b & b+c & c+a \\ l+m & m+n & n+l \\ p+q & q+r & r+p \end{vmatrix} = 2$

4. $\begin{vmatrix} a-b-c & 2b & 2c \\ 2a & b-c-a & 2c \\ 2a & 2b & c-a-b \end{vmatrix} = (a+b+c)^3$

5. $\begin{vmatrix} a+b & a & b \\ a & a+c & c \\ b & c & b+c \end{vmatrix} = 4abc$

6. $\begin{vmatrix} 4 & 5 & 6 & x \\ 5 & 6 & 7 & y \\ 6 & 7 & 8 & z \\ x & y & z & 0 \end{vmatrix}$ is a perfect square.



7. $\begin{vmatrix} 1 & a & a^2 & a^3+bcd \\ 1 & b & b^2 & b^3+cda \\ 1 & c & c^2 & c^3+dab \\ 1 & d & d^2 & d^3+abc \end{vmatrix}$ vanishes.

8. Solve the equation $\begin{vmatrix} x+1 & 2x+1 & 3x+1 \\ 2x & 4x+3 & 6x+3 \\ 4x+1 & 6x+4 & 8x+4 \end{vmatrix} = 0.$

9. Find the value of the determinant $(M)$ if $M = 3A^2 + AB + B^2$

where $A = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & -1 & 0 \end{vmatrix},\ B = \begin{vmatrix} 5 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 2 & 3 \end{vmatrix}$

without evaluating $A$ and $B$ independently.

10. Express $\begin{vmatrix} 2bc-a^2 & c^2 & b^2 \\ c^2 & 2ca-b^2 & a^2 \\ b^2 & a^2 & 2ab-c^2 \end{vmatrix}$

as the square of a determinant and hence find its value.

## 3.2. INTRODUCTION TO MATRICES

**(1) Def.** *A system of mn numbers arranged in a rectangular array of m rows and n olumns is called an* $m \times n$ **matrix.** Such a matrix is denoted by

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

**(2) Special matrices**

**1.** *Row and column matrices.* A matrix having a single row is called a *row matrix* while a matrix having a single column is called a *column matrix.*

**2.** *Square matrix.* A matrix having $n$ rows and $n$ columns is called a *square matrix.*

A square matrix is said to be *singular* if its determinant is zero otherwise it is called *non-singular.*

The elements $a_{ii}$ in a square matrix form the *leading diagonal* and their sum $\Sigma a_{ii}$ is called the *trace* of the matrix.

**3.** *Unit matrix.* A diagonal matrix of order $n$ which has unity for all its diagonal elements is called a *unit matrix of order n* and is denoted by $I_n$.

**4.** *Null matrix.* If all the elements of a matrix are zero, it is called a *null matrix.*

**5.** *Symmetric and skew-symmetric matrices.* A square matrix $[a_{ij}]$ is said to be *symmetric* when $a_{ij} = a_{ji}$ for all $i$ and $j$.

If $a_{ij} = -a_{ji}$ for all $i$ and $j$ so that all the leading diagonal elements are zero, then the matrix is called *skew-symmetric.* Examples of symmetric and skew-symmetric matrices are respectively

$$\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \text{ and } \begin{bmatrix} 0 & h & -g \\ -h & 0 & f \\ g & -f & 0 \end{bmatrix}$$

**6.** *Triangular matrix.* A square matrix all of whose elements below the leading diagonal are zero is called an *upper triangular matrix*. A square matrix all of whose elements above the leading diagonal are zero is called a *lower triangular matrix*.

**(3) Operations on matrices**

**1.** *Equality of matrices.* Two matrices $A$ and $B$ are said to be equal if and only if

(*i*) they are of the same order,

and (*ii*) each element of $A$ is equal to the corresponding element of $B$.

**2.** *Addition and subtraction of matrices.* If $A$ and $B$ be two matrices of the same order, then their sum $A + B$ is defined as the matrix each element of which is the sum of the corresponding elements of $A$ and $B$.

Similarly $A - B$ is defined as the matrix whose elements are obtained by subtracting the elements of $B$ from the corresponding elements of $A$.

**3.** *Multiplication of a matrix by a scalar.* The product of a matrix $A$ by a scalar $k$ is a matrix whose each element is $k$ times the corresponding elements of $A$.

**4.** *Multiplication of matrices.* Two matrices can be multiplied only when the number of columns in the first is equal to the number of rows in the second. Such matrices are said to be *conformable*. Thus if $A$ and $B$ be $(m \times n)$ and $(n \times p)$ matrices, then their product $C = AB$ is defined and will be a $(m \times p)$ matrix. *The elements of C are obtained by the following rule :*

*Element $c_{ij}$ of C = sum of the products of corresponding elements of the ith row of A with those of the jth column of B.*

For example, if $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$ and $B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$

then
$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} \\ a_{41}b_{11} + a_{42}b_{21} + a_{43}b_{31} & a_{41}b_{12} + a_{42}b_{22} + a_{43}b_{32} \end{bmatrix}$$

**Obs. 1.** In general $AB \neq BA$ even if both exist.

**2.** If $A$ be a square matrix, then the product $AA$ is defined as $A^2$. Similarly $A.A^2 = A^3$ etc.

■ **Example 3.6.** *Evaluate* $3A - 4B$, *where* $A = \begin{bmatrix} 3 & -4 & 6 \\ 5 & 1 & 7 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 3 \end{bmatrix}$.

**Sol.** We have $3A = \begin{bmatrix} 9 & -12 & 18 \\ 15 & 3 & 21 \end{bmatrix}$ and $4B = \begin{bmatrix} 4 & 0 & 4 \\ 8 & 0 & 12 \end{bmatrix}$

$$\therefore \quad 3A - 4B = \begin{bmatrix} 9-4 & -12-0 & 18-4 \\ 15-8 & 3-0 & 21-12 \end{bmatrix} = \begin{bmatrix} 5 & -12 & 14 \\ 7 & 3 & 9 \end{bmatrix}$$

■ **Example 3.7.** *If* $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{bmatrix}$ *and* $B = \begin{bmatrix} 1 & -2 \\ -1 & 0 \\ 2 & -1 \end{bmatrix}$, *form the product AB. Is BA defined ?*

**Sol.** Since the number of columns of $A$ = the number of rows of $B$ (each being = 3). The product $AB$ is defined and



$$= \begin{bmatrix} 0.1 + 1.-1 + 2.2 & 0.-2 + 1.0 + 2.-1 \\ 1.1 + 2.-1 + 3.2 & 1.-2 - 2.0 + 3-1 \\ 2.1 + 3.-1 + 4.2 & 2.-2 + 3.0 + 4.-1 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ 5 & -5 \\ 7 & -8 \end{bmatrix}$$

Again since the number of columns of $B \neq$ the number of rows of $A$.

$\therefore$ The product $BA$ is not defined.

■ **Example 3.8.** *If* $A = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 3 & 1 \\ 5 & 3 & 4 \end{bmatrix}$, *find the matrix B, such that* $AB = \begin{bmatrix} 3 & 4 & 2 \\ 1 & 6 & 1 \\ 5 & 6 & 4 \end{bmatrix}$.

(*Mumbai, B. Tech., 2005*)

**Sol.** Let
$$AB = \begin{bmatrix} 3 & 2 & 2 \\ 1 & 3 & 1 \\ 5 & 3 & 4 \end{bmatrix} \begin{bmatrix} l & m & n \\ p & q & r \\ u & v & w \end{bmatrix}$$
$$= \begin{bmatrix} 3l + 2p + 2u & 3m + 2q + 2v & 3n + 2r + 2w \\ 1 + 3p + u & m + 3q + v & n + 3r + w \\ 5l + 3p + 4u & 5m + 3q + 4v & 5n + 3r + 4w \end{bmatrix}$$
$$= \begin{bmatrix} 3 & 4 & 2 \\ 1 & 6 & 1 \\ 5 & 6 & 4 \end{bmatrix} \quad \text{(given)}$$

Equating corresponding elements, we get

$3l + 2p + 2u = 3,\quad l + 3p + u = 1,\quad 5l + 3p + 4u = 5$ ...(*i*)

$3m + 2q + 2v = 4,\quad m + 3q + v = 6,\quad 5m + 3q + 4v = 6$ ...(*ii*)

$3n + 2r + 2w = 2,\quad n + 3r + w = 1,\quad 5n + 3r + 4x = 4$ ...(*iii*)

Solving the equations (*i*), we get $l = 1, p = 0, u = 0$

Similarly equations (*ii*) give $m = 0, q = 2, v = 0$

and equations (*iii*) give $n = 0, r = 0, w = 1$

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

**(4) Related matrices**

***I.* Transpose of a matrix.** *The matrix obtained from a given matrix A, by interchanging rows and columns, is called the* **transpose** of $A$ and is denoted by $A'$.

**Obs. 1.** For a *symmetric matrix*, $A' = A$

and for a *skew-symmetric matrix*, $A' = -A$.

**2.** *The transpose of the product of two matrices is the product of their transposes taken in the reverse order*

*i.e.*
$$(AB)' = B'A'.$$

**3.** Any square matrix $A$ can be written as
$$A = \tfrac{1}{2}(A + A') + \tfrac{1}{2}(A - A') = B + C \text{ (say)}$$
such that
$$B' = \tfrac{1}{2}(A + A')' = \tfrac{1}{2}(A' + A) = B$$
*i.e.* $B$ is a symmetric matrix

and $\qquad C' = \frac{1}{2}(A - A')' = \frac{1}{2}(A' - A) = -C$

*i.e.* $\quad$ $C$ is a skew-symmetric matrix.

*Thus every square matrix can be expressed as the sum of a symmetric and a skew-symmetric matrix.*

■ **Example 3.9.** *Express the matrix A as the sum of a symmetric and a skew-symmetric matrix where* $A = \begin{bmatrix} 3 & -2 & 6 \\ 2 & 7 & -1 \\ 5 & 4 & 0 \end{bmatrix}$

**Sol.** We have $\quad A' = \begin{bmatrix} 3 & 2 & 5 \\ -2 & 7 & 4 \\ 6 & -1 & 0 \end{bmatrix}$

Then $\quad A + A' = \begin{bmatrix} 6 & 0 & 11 \\ 0 & 14 & 3 \\ 11 & 3 & 0 \end{bmatrix}, \quad A - A' = \begin{bmatrix} 0 & -4 & 1 \\ 4 & 0 & -5 \\ -1 & 5 & 0 \end{bmatrix}$

$$\therefore \quad A = \frac{1}{2}(A + A') + \frac{1}{2}(A - A')$$

$$= \begin{bmatrix} 3 & 0 & 5.5 \\ 0 & 7 & 1.5 \\ 5.5 & 1.5 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -2 & .5 \\ 2 & 0 & -2.5 \\ -.5 & 2.5 & 0 \end{bmatrix}$$

*II.* **Adjoint of a square matrix A** *is the transposed matrix of cofactors of A* and is written as adj $A$. Thus the adjoint of the matrix $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$ is $\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix}$.

*III.* **Inverse of a matrix.** *If A be a non-singular square matrix of order n, then a square matrix B of the same order such that AB = BA = I, is the called the* **inverse** *of A, I* being a unit matrix.

The inverse of A is written as $A^{-1}$ so that $A\,A^{-1} = A^{-1}A = I$

Also $\qquad A^{-1} = \dfrac{\text{Adj } A}{|A|}$

**Obs. 1.** Inverse of a matrix, when it exists, is unique.

**2.** $(A^{-1})^{-1} = A$

**3.** $(AB)^{-1} = B^{-1}A^{-1}$.

■ **Example 3.10.** *Find the inverse of* $A = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & -3 \\ -2 & -4 & -4 \end{bmatrix}$.

**Sol.** Here $\qquad |A| = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 3 & -3 \\ -2 & -4 & -4 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$ (say)



and $\qquad \text{adj } A = \begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} = \begin{bmatrix} -24 & -8 & -12 \\ 10 & 2 & 6 \\ 2 & 2 & 2 \end{bmatrix}$

Hence $\qquad A^{-1} = \dfrac{\text{adj } A}{|A|} = \dfrac{1}{8}\begin{bmatrix} -24 & -8 & -12 \\ 10 & 2 & 6 \\ 2 & 2 & 2 \end{bmatrix}$

**Note.** For other methods of finding the inverse of a matrix refer to chapter 4.

**(5) Rank of a matrix.** If we select any $r$ rows and $r$ columns from any matrix $A$, deleting all other rows and columns, then the determinant formed by these $r \times r$ elements is called the *minor of A of order r*. Clearly there will be a number of different minors of the same order, got by deleting different rows and columns from the same matrix.

**Def.** *A matrix is said to be of rank r when*

*(i) it has at least one non-zero minor of order r,*

*and (ii) every minor of order higher than r vanishes.*

**(6) Elementary transformations of a matrix.** The following operations, three of which refer to rows and three to columns are known as *elementary transformations*

*I. The interchange of any two rows (columns).*

*II. The multiplication of any row (column) by a non-zero number.*

*III. The addition of a constant multiple of the elements of any row (column) to the corresponding elements of any other row (column).*

**Notation.** The *elementary row transformations* will be denoted by the following symbols

(i) $R_{ij}$ for the interchange of the $i$th and $j$th rows.

(ii) $kR_i$ for multiplication of the $i$th row by $k$.

(iii) $R_i + pR_j$ for addition to the $i$th row, $p$ times the $j$th row.

The corresponding column transformation will be denoted by writing $C$ in place of $R$. These transformations being precisely those performed on the rows (columns) of a determinant, need no explanation.

**Obs. 1.** *Elementary transformations do not change either the order or rank of a matrix.* While the value of the minors may get changed by the transformations I and II, their zero or non-zero character remains unaffected.

**(7) Equivalent matrix.** *Two matrices A and B are said to be equivalent if one can be obtained from the other by a sequence of elementary transformations.* Two equivalent matrices have the same order and the same rank. The symbol ~ is used for equivalence.

**Elementary matrix.** *An elementary matrix is that, which is obtained from a unit matrix, by subjecting it to any of the elementary transformations.*

**Normal form of a matrix.** *Every non-zero matrix A of rank r, can be reduced by a sequence of elementary transformations, to the form* $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ *which is called the* **normal form** *of A.*

**Example 3.11.** *Determine the rank of the following matrices :*

$$(i)\begin{bmatrix}1&2&3\\1&4&2\\2&6&5\end{bmatrix}\qquad (ii)\begin{bmatrix}0&1&-3&-1\\0&0&1&1\\3&1&0&2\\1&1&-2&0\end{bmatrix}.$$

**Sol.** (*i*) Operate $R_2 - R_1$ and $R_3 - 2R_1$ so that the given matrix

$$\sim\begin{bmatrix}1&2&3\\0&2&-1\\0&2&-1\end{bmatrix} = A \quad \text{(say)}$$

Obviously, the 3rd order minor of $A$ vanishes. Also its 2nd order minors formed by its 2nd and 3rd rows are all zero. But another 2nd order minor is

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

Hence $R(A)$, the rank of the given matrix, is 2.

(*ii*) Given matrix

$$\sim\begin{bmatrix}0&1&-3&-1\\1&0&0&0\\3&1&-3&-1\\1&1&-3&-1\end{bmatrix}\quad [\text{Operating } C_3 - C_1, C_4 - C_1]$$

$$\sim\begin{bmatrix}0&1&-3&-1\\1&0&0&0\\3&0&0&0\\1&0&0&0\end{bmatrix}\quad [\text{Operating } R_3 - R_1, R_4 - R_1]$$

$$\sim\begin{bmatrix}0&1&-3&-1\\1&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix}\quad [\text{Operating } R_3 - 3R_2, R_4 - R_2]$$

$$\sim\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix} = A \text{ (say)}\quad [\text{Operating } C_3 + 3C_2, C_4 + C_2]$$

Obviously, the 4th order minor of $A$ is zero. Also every 3rd order minor of $A$ is zero. But, of all the 2nd order minors, only $\begin{vmatrix}0&1\\1&0\end{vmatrix} = -1 \neq 0.$

Hence $R(A)$, the rank of the given matrix, is 2.

**(8) Consistency of a system of linear equations.** Consider the system of $m$ linear equations in $n$ unknowns

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n &= k_1\\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n &= k_2\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots &\\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n &= k_m \end{aligned}\right\} \qquad \ldots(i)$$

To determine whether these equations are consistent or not, we find the ranks of the matrices

$$A = \begin{bmatrix}a_{11} & a_{12}\ldots\ldots a_{1n}\\ a_{21} & a_{22}\ldots\ldots a_{2n}\\ \ldots\ldots & \ldots\ldots\ldots\\ a_{m1} & a_{m2}\ldots\ldots a_{m3}\end{bmatrix} \text{ and } K = \begin{bmatrix}a_{11} & a_{12}\ldots\ldots a_{1n} & k_1\\ a_{21} & a_{22}\ldots\ldots a_{2n} & k_2\\ \ldots\ldots & \ldots\ldots\ldots & \ldots\\ a_{m1} & a_{m2}\ldots\ldots a_{mn} & k_m\end{bmatrix}$$



$A$ is the *coefficient matrix* and $K$ is called the *augmented matrix*.

*If* $R(A) \neq R(K)$, *the equations (i) are inconsistent i.e. have no solution.*

*If* $R(A) = R(K) = n$, *the equations (i) are consistent and have a unique solution.*

*If* $R(A) = R(K) < n$, *the equations are consistent but have an infinite number of solutions.*

**(9) System of linear homogeneous equations.** Consider the homogeneous linear equations

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n &= 0\\ a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n &= 0\\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots &\\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n &= 0 \end{aligned}\right\} \qquad \ldots(i)$$

Find the rank $r$ of the coefficient matrix $A$ by reducing it to the triangular form by elementary row operations.

*I. If* $r = n$, *the equations (i) have only a trivial solution* $x_1 = x_2 = \ldots\ldots = x_n = 0$.

*If* $r < n$, *the equations have* $(n - r)$ *independent solutions.* ($r$ cannot be $> n$)

The number of linearly independent solutions of (*i*) is $(n - r)$ means, if arbitrary values are assigned to $(n - r)$ of the variables, the values of the remaining variables can be uniquely found.

*II. When* $m < n$ (*i.e.* the number of equations is less than the number of variables) *the solution is always other than* $x_1 = x_2 = \ldots\ldots = x_n = 0$.

*III. When* $m = n$ (*i.e.* the number of equations = the number of variables) *the necessary and sufficient condition for solutions other than* $x_1 = x_2 = \ldots\ldots = x_n = 0$ *is that* $|A| = 0$ (*i.e.* the determinant of the coefficient matrix is zero).

**Example 3.12.** *Test for consistency and solve*

$5x + 3y + 7z = 4,\ 3x + 26y + 2z = 9,\ 7x + 2y + 10z = 5.$ (*P.T.U., B. Tech., 2005*)

**Sol.** We have
$$\begin{bmatrix}5&3&7\\3&26&2\\7&2&10\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}4\\9\\5\end{bmatrix}$$

Operate $3R_1, 5R_2$,
$$\begin{bmatrix}15&9&21\\15&130&10\\7&2&10\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}12\\45\\5\end{bmatrix}$$

Operate $R_2 - R_1$,
$$\begin{bmatrix}15&9&21\\0&121&-11\\7&2&10\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}12\\33\\5\end{bmatrix}$$

Operate $\frac{7}{8}R_1, 5R_3, \frac{1}{11}R_2$,
$$\begin{bmatrix}35&21&49\\0&11&-1\\35&10&50\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}28\\3\\25\end{bmatrix}$$

Operate $R_3 - R_1 + R_2, \frac{1}{7}R_1$,
$$\begin{bmatrix}5&3&7\\0&11&-1\\0&0&0\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}4\\3\\0\end{bmatrix}$$

In the last set of equations, the number of non-zero rows in the coefficient matrix being two, its rank is two. Also the number of non-zero rows in the augmented matrix being, its rank is two.

Now, the ranks of coefficient matrix and augmented matrix being equal, the equations are consistent. Also the given system is equivalent to

$$5x + 3y + 7z = 4,\ 11y - z = 3.$$

$$\therefore \qquad y = \frac{3}{11} + \frac{z}{11} \quad \text{and} \quad x = \frac{7}{11} - \frac{16}{11}z$$

where $z$ is a parameter.

Hence $\qquad x = \frac{7}{11}, y = \frac{3}{11}$ and $z = 0$ is a particular solution.

**Example 3.13.** *Examine the system of equations $3x + 3y + 2z = 1$, $x + 2y = 4$, $10y + 3z = -2$, $2x - 3y - z = 5$ for consistency and hence solve it.* (J.N.T.U., B. Tech., 2009)

**Sol.** We have

$$\begin{bmatrix} 3 & 3 & 2 \\ 1 & 2 & 0 \\ 0 & 10 & 3 \\ 2 & -3 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ -2 \\ 5 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 2 & 0 \\ 3 & 3 & 2 \\ 0 & 10 & 3 \\ 2 & -3 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ -2 \\ 5 \end{bmatrix} \qquad [\text{Interchanging } R_1 \ \& \ R_2]$$

$$\text{or} \qquad \begin{bmatrix} 1 & 2 & 0 \\ 0 & -3 & 2 \\ 0 & 10 & 3 \\ 0 & -7 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ -11 \\ -2 \\ -3 \end{bmatrix} \qquad [\text{Operating } R_2 - 3R_1, R_4 - 2R_1]$$

$$\text{or} \qquad \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -2/3 \\ 0 & 0 & 29/3 \\ 0 & 0 & -17/3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 11/3 \\ -116/3 \\ 68/3 \end{bmatrix} \qquad \left[\text{Operating } R_3 + \frac{10}{3}R_2, R_4 - \frac{7}{3}R_2\right]$$

$$\text{or} \qquad \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -2/3 \\ 0 & 0 & 29/3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 11/3 \\ -116/3 \\ 0 \end{bmatrix} \qquad \left[\text{Operating } R_4 + \frac{17}{29}R_3\right]$$

Now in the last set of equations, number of non-zero rows in the coefficient matrix being three, its rank is three.

Also the number of non-zero rows in the augmented matrix being three, its rank is three.

Since the ranks of the coefficient and the augmented matrices are equal, the given equations are consistent.

Also number of unknowns = Rank of the coefficient matrix.

Hence the given equations have a unique solution given by

$$x + 2y = 4,\ y - \frac{2}{3}z = \frac{11}{3},\ \frac{29}{3}z = -\frac{116}{3},$$

...e equations give $z = -4, y = 1, x = 2$.



**Example 3.14.** *Investigate the values of $\lambda$ and $\mu$ so that the equations*

$$2x + 3y + 5z = 9,\ 7x + 3y - 2z = 8,\ 2x + 3y + \lambda z = \mu,$$

*have (i) no solution, (ii) a unique solution and (iii) an infinite number of solutions.* (V.T.U., B.Tech., 2007)

**Sol.** We have $\qquad \begin{bmatrix} 2 & 3 & 5 \\ 7 & 3 & -2 \\ 2 & 3 & \lambda \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 8 \\ \mu \end{bmatrix}$

The system admits of unique solution if and only if, the coefficient matrix is of rank 3. This requires that

$$\begin{bmatrix} 2 & 3 & 5 \\ 7 & 3 & -2 \\ 2 & 3 & \lambda \end{bmatrix} = 15(5 - \lambda) \neq 0$$

Thus for a unique solution $\lambda \neq 5$ and $\mu$ may have any value. If $\lambda = 5$, the system will have no solution for those values of $\mu$ for which the matrices

$$A = \begin{bmatrix} 2 & 3 & 5 \\ 7 & 3 & -2 \\ 2 & 3 & \lambda \end{bmatrix} \quad \text{and} \quad K = \begin{bmatrix} 2 & 3 & 5 & 9 \\ 7 & 3 & -2 & 8 \\ 2 & 3 & 5 & \mu \end{bmatrix}$$

are not of the same rank. But $A$ is of rank 2 and $K$ is not of rank 2 unless $\mu = 9$. Thus if $\lambda = 5$ and $\mu \neq 9$, the system will have no solution.

If $\lambda = 5$ and $\mu = 9$, the system will have an infinite number of solutions.

**Example 3.15.** *Solve the equations*

$$4x + 2y + z + 3w = 0,\ 6x + 3y + 4z + 7w = 0,\ 2x + y + w = 0.$$

**Sol.** Rank of the coefficient matrix

$$\begin{bmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 4 & 2 & 1 & 3 \\ 0 & 0 & 5/2 & 5/2 \\ 0 & 0 & -1/2 & -1/2 \end{bmatrix} \qquad \left[\text{Operating } R_2 - \tfrac{3}{2}R_1, R_3 - \tfrac{1}{2}R_1\right]$$

$$\sim \begin{bmatrix} 4 & 2 & 1 & 3 \\ 0 & 0 & 5/2 & 5/2 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \left[\text{Operating } R_3 + \tfrac{1}{5}R_2\right]$$

is 2 which is less than the number of variables.

$\therefore$ The number of independent solutions $= 4 - 2 = 2$.

Also the given system is equivalent to

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

$$\therefore \qquad z = -w,\ x = -\tfrac{1}{2}(y + w).$$

Choosing $w = k_1$ and $x = k_2$, we have $y = -2k_2 - k_1$ and $z = -k_1$.

**Example 3.16.** *Find the values of $k$ for which the system of equations $(3k - 8)x + 3y + 3z = 0$, $3x + (3k - 8)y + 3z = 0$, $3x + 3y + (3k - 8)z = 0$ has a non-trivial solution.* (U.P.T.U., B. Tech., 2006)

**Sol.** For the given system of equations to have a non-trivial solution, the determinant of coefficient matrix should be zero.

*i.e.* $\begin{vmatrix} 3k-8 & 3 & 3 \\ 3 & 3k-8 & 3 \\ 3 & 3 & 3k-8 \end{vmatrix} = 0$ or $\begin{vmatrix} 3k-2 & 3 & 3 \\ 3k-2 & 3k-8 & 3 \\ 3k-2 & 3 & 3k-8 \end{vmatrix} = 0$

or $(3k-2)\begin{vmatrix} 1 & 3 & 3 \\ 1 & 3k-8 & 3 \\ 1 & 3 & 3k-8 \end{vmatrix} = 0$ [Operating $C_1 + (C_2 + C_3)$

or $(3k-2)\begin{vmatrix} 1 & 3 & 3 \\ 0 & 3k-11 & 0 \\ 0 & 0 & 3k-11 \end{vmatrix} = 0$ [Operating $R_2 - R_1, R_3 - R_1$

or $(3k-2)(3k-11)^2 = 0$ whence $k = 2/3$ or $11/3$.

## PROBLEMS 3.2

1. Find $x, y, z$ and $w$ given that $3\begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} x & 6 \\ -1 & 2w \end{bmatrix} + \begin{bmatrix} 4 & x+y \\ z+w & 3 \end{bmatrix}$

2. If $A = \begin{bmatrix} 1 & 3 & 0 \\ -1 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 3 \\ -1 & 1 & 2 \end{bmatrix}$, compute $AB$, $BA$ and show that $AB \neq BA$.

3. If $A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 3 & -1 \\ -3 & 1 & 2 \end{bmatrix}$ and $I$ is the unit matrix of order 3, evaluate $A^2 - 3A + 9I$.

4. Express each of the following matrices as the sum of a symmetric and a skew-symmetric matrices :

   (*i*) $\begin{bmatrix} 0 & 5 & -3 \\ 1 & 1 & 1 \\ 4 & 5 & 9 \end{bmatrix}$ (*ii*) $\begin{bmatrix} a & a & b \\ c & b & b \\ c & a & c \end{bmatrix}$

5. If $A = \begin{bmatrix} 2 & 5 & 3 \\ 3 & 1 & 2 \\ 1 & 2 & 1 \end{bmatrix}$, find adj $A$ and $A^{-1}$.

6. Find the inverse of the matrix $\begin{bmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{bmatrix}$

7. If $A = \frac{1}{3}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{bmatrix}$, prove that $A^{-1} = A'$.

8. If $A = \begin{bmatrix} 3 & -3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{bmatrix}$, show that $A^3 = A^{-1}$.

9. Factorize the matrix $\begin{bmatrix} 5 & -2 & 1 \\ 7 & 1 & -5 \\ 3 & 7 & 4 \end{bmatrix}$ into the form $LU$, where $L$ is lower triangular and $U$ is upper triangular matrix.



10. Determine the ranks of the following matrices :

    (*i*) $\begin{bmatrix} 1 & 3 & 4 & 3 \\ 3 & 9 & 12 & 3 \\ 1 & 3 & 4 & 1 \end{bmatrix}$ (*Madras B.E., 2000*) (*ii*) $\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 4 & 3 & 2 \\ 3 & 2 & 1 & 3 \\ 6 & 8 & 7 & 5 \end{bmatrix}$ (*Osmania, B.E., 2003*)

11. Examine for consistency the following equations and hence solve these :

    (*i*) $x + 2y = 1$, $7x + 14y = 12$.

    (*ii*) $2x - 3y + 7z = 5$, $3x + y - 3z = 13$, $2x + 19y - 47z = 32$. (*R.T.U., B. Tech., 2007*)

    (*iii*) $x + 2y + z = 3$, $2x + 3y + 2z = 5$, $3x - 5y + 5z = 2$, $3x + 9y - z = 4$. (*Bhillai, B. Tech., 2005*)

12. Show that if $\lambda \neq -5$, the system of equations $3x - y + 4z = 3$, $x + 2y - 3z = -2$, $6x + 5y + \lambda z = -3$ has a unique solution. If $\lambda = -5$, show that the equations are consistent.

13. Investigate for what values of $\lambda$ and $\mu$ the simultaneous equations $x + y + z = 6$, $x + 2y + 3z = 10$, $x + 2y + \lambda z = \mu$, have (*i*) no solution, (*ii*) a unique solution, (*iii*) an infinite number of solutions. (*U.P.T.U., B. Tech., 2006*)

14. Find the values of $a$ and $b$ for which the equations $x + ay + z = 3$, $x + 2y + 2z = b$, $x + 5y + 3z = 9$ are consistent. When will these equations have a unique solution ? (*Madras, B.E., 2003*)

15. Determine the values of $\lambda$ for which the following set of equations may possess non-trivial solutions :

    $$3x_1 + x_2 - \lambda x_3 = 0,\ 4x_1 - 2x_2 - 3x_3 = 0,\ 2\lambda x_1 + 4x_2 + \lambda x_3 = 0.$$

    For each permissible value of $\lambda$, determine the general solution. (*Kurukshetra, B. Tech., 2006*)

16. Solve completely the system of equations :

    $$3x + 4y - z - 6w = 0;\quad 2x + 3y + 2z - 3w = 0;$$
    $$2x + y - 14z - 9w = 0;\quad x + 3y + 13z + 3w = 0.$$

    (*J.N.T.U., B. Tech., 2002*)

## 3.3. SOLUTION OF LINEAR SIMULTANEOUS EQUATIONS

Simultaneous linear equations occur quite often in engineering and science. The analysis of electronic circuits consisting of invariant elements, analysis of a network under sinusoidal steady-state conditions, determination of the output of a chemical plant, finding the cost of chemical reactions are some of the problems which depend on the solution of simultaneous linear algebraic equations. The solution of such equations can be obtained by *Direct* or *Iterative methods*. We describe below some such methods of solution.

## 3.4. DIRECT METHODS OF SOLUTION

**(1) Method of determinants—Cramer's rule.** Consider the equations

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \quad ...(1)$$

If the determinant of coefficients be

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

then

$$x\Delta = \begin{vmatrix} xa_1 & b_1 & c_1 \\ xa_2 & b_2 & c_2 \\ xa_3 & b_3 & c_3 \end{vmatrix} \qquad [\text{Operate } C_1 + yC_2 + zC_3]$$

$$= \begin{vmatrix} a_1x + b_1y + c_1z & b_1 & c_1 \\ a_2x + b_2y + c_2z & b_2 & c_2 \\ a_3x + b_3y + c_3z & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix}$$

Thus

$$x = \begin{vmatrix} d_1 & b_1 & c_1 \\ d_2 & b_2 & c_2 \\ d_3 & b_3 & c_3 \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \quad \text{provided } \Delta \neq 0. \qquad ...(2)$$

Similarly

$$y = \begin{vmatrix} a_1 & d_1 & c_1 \\ a_2 & d_2 & c_2 \\ a_3 & d_3 & c_3 \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \qquad ...(3)$$

and

$$z = \begin{vmatrix} a_1 & b_1 & d_1 \\ a_2 & b_2 & d_2 \\ a_3 & b_3 & d_3 \end{vmatrix} \div \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \qquad ...(4)$$

The equations (2), (3) and (4) giving the values of $x, y, z$ constitute the *Cramer's rule** which reduces the solution of the linear system (1) to a problem in evaluation of determinants.

**Obs.** 1. Cramer's rule fails for $\Delta = 0$.

2. This method is quite general but involves a lot of labour when the number of equations exceeds four. For a $10 \times 10$ system, Cramer's rule requires about 70,000,000 multiplications. We shall explain another method which requires only 333 multiplications, for the same $10 \times 10$ system. As such, the Cramer's rule is not at all suitable for large systems.

**Example 3.17.** *Apply Cramer's rule to solve the equations*
*$3x + y + 2z = 3$, $2x - 3y - z = -3$, $x + 2y + z = 4$.*

**Sol.** Here

$$\Delta = \begin{vmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & 2 & 1 \end{vmatrix} = 8$$

$\therefore$

$$x = \frac{1}{\Delta}\begin{vmatrix} 3 & 1 & 2 \\ -3 & -3 & -1 \\ 4 & 2 & 1 \end{vmatrix} = \frac{1}{8}(8) = 1,$$

$$y = \frac{1}{\Delta}\begin{vmatrix} 3 & 3 & 2 \\ 2 & -3 & -1 \\ 1 & 4 & 1 \end{vmatrix} = \frac{1}{8}(16) = 2$$

*Gabriel Cramer (1704–1752), a Swiss mathematician.



and

$$z = \frac{1}{\Delta}\begin{vmatrix} 3 & 1 & 3 \\ 2 & -3 & -3 \\ 1 & 2 & 4 \end{vmatrix} = \frac{1}{8}(-8) = -1$$

Hence $x = 1$, $y = 2$ and $z = -1$.

**(2) Matrix inversion method.** Consider the equations

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \qquad ...(1)$$

If

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}, \quad X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix},$$

then the equations (1) are equivalent to the matrix equation $AX = D$ ...(2)

Multiplying both sides of (2) by the inverse matrix $A^{-1}$, we get

$$A^{-1}AX = A^{-1}D \quad \text{or} \quad IX = A^{-1}D \qquad [\because A^{-1}A = I]$$

or

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

i.e.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} \times \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \qquad ...(3)$$

where $A_1, B_1$, etc. are the cofactors of $a_1, b_1$, etc. in the determinant $|A|$.

Hence equating the values of $x, y, z$ to the corresponding elements in the product on the right side of (3) we get the desired solution.

**Obs.** This method fails when $A$ is a singular matrix *i.e.* $|A| = 0$. Although this method is quite general, yet it is not suitable for large systems since the evaluation of $A^{-1}$ by cofactors becomes very cumbersome. We shall now explain some methods which can be applied to any number of equations.

**Example 3.18.** *Solve the equations of Ex. 3.17 by matrix inversion method.*

**Sol.** Here

$$A = \begin{bmatrix} 3 & 1 & 2 \\ 2 & -3 & -1 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \qquad \text{(say)}$$

$\therefore$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{bmatrix} \times \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \frac{1}{8}\begin{bmatrix} -1 & 3 & 5 \\ -3 & 1 & 7 \\ 7 & -5 & -11 \end{bmatrix} \times \begin{bmatrix} 3 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Hence $x = 1$, $y = 2$, $z = -1$.

**(3) Gauss elimination method.** In this method, the unknowns are eliminated successively and the system is reduced to an upper triangular system from which the unknowns are found by back substitution. The method is quite general and is well-adapted for computer operations. Here we shall explain it by considering a system of three equations for sake of clarity.

Consider the equations

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \qquad ...(1)$$

*Step I. To eliminate x from second and third equations.*

Assuming $a_1 \neq 0$, we eliminate $x$ from the second equation by subtracting $(a_2/a_1)$ times the first equation from the second equation. Similarly we eliminate $x$ from the third equation by eliminating $(a_3/a_1)$ times the first equation from the third equation. We thus, get the new system

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ b_2'y + c_2'z &= d_2' \\ b_3'y + c_3'z &= d_3' \end{aligned}\right\} \quad ...(2)$$

Here the first equation is called the *pivotal equation* and $a_1$ is called the *first pivot*.

*Step II. To eliminate y from third equation in (2).*

Assuming $b_2' \neq 0$, we eliminate $y$ from the third equation of (2), by subtracting $(b_3'/b_2')$ times the second equation from the third equation. We thus, get the new system

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ b_2'y + c_2'z &= d_2' \\ c_3''z &= d_3'' \end{aligned}\right\} \quad ...(3)$$

Here the second equation is the *pivotal equation* and $b_2'$ is the *new pivot*.

*Step III. To evaluate the unknowns.*

The values of $x, y, z$ are found from the reduced system (3) by back substitution.

**Obs. 1.** On writing the given equations as

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} \quad i.e., \quad AX = D,$$

this *method consists in* ***transforming the coefficient matrix A to upper triangular matrix*** *by elementary row transformations only.*

**Obs. 2.** Clearly the method will fail if any one of the pivots $a_1$, $b_2'$ or $c_3''$ becomes zero. In such cases, we rewrite the equations in a different order so that the pivots are non-zero.

**Obs. 3.** *Partial and complete pivoting.* In the first step, the numerically largest coefficient of $x$ is chosen from all the equations and brought as the first pivot by interchanging the first equation with the equation having the largest coefficient of $x$. In the second step, the numerically largest coefficient of $y$ is chosen from the remaining equations (leaving the first equation) and brought as the *second pivot* by interchanging the second equation with the equation having the largest coefficient of $y$. This process is continued till we arrive at the equation with the single variable. This modified procedure is called *partial pivoting*.

If we are not keen about the elimination of $x, y, z$ in a specified order, then we can choose at each stage the numerically largest coefficient of the entire matrix of coefficients. This requires not only an interchange of equations but also an interchange of the position of the variables. This method of elimination is called *complete pivoting*. It is more complicated and does not appreciably improve the accuracy.

**Example 3.19.** *Apply Gauss elimination method to solve the equations $x + 4y - z = -5$; $x + y - 6z = -12$; $3x - y - z = 4$.* *(Mumbai, B. Tech., 2005)*

**Sol.** We have

| | Check sum | |
|---|---|---|
| $x + 4y - z = -5$ | $-1$ | ...(i) |
| $x + y - 6z = -12$ | $-16$ | ...(ii) |
| $3x - y - z = 4$ | $5$ | ...(iii) |



*Step I.* To eliminate $x$, operate $(ii) - (i)$ and $(iii) - 3(i)$ :

| | Check sum | |
|---|---|---|
| $-3y - 5z = -7$ | $-15$ | ...(iv) |
| $-13y + 2z = 19$ | $8$ | ...(v) |

*Step II.* To eliminate y, operate $(v) - \frac{13}{3}(iv)$ :

| | Check sum | |
|---|---|---|
| $\frac{71}{3}z = \frac{148}{3}$ | $73$ | ...(vi) |

*Step III.* By back-substitution, we get

From $(vi)$ : $\quad z = \frac{148}{71} = 2.0845$

From $(iv)$ : $\quad y = \frac{7}{3} - \frac{5}{3}\left(\frac{148}{71}\right) = -\frac{81}{71} = -1.1408$

From $(i)$ : $\quad x = -5 - 4\left(-\frac{81}{71}\right) + \left(\frac{148}{71}\right) = \frac{117}{71} = 1.6479$

Hence, $x = 1.6479$, $y = -1.1408$, $z = 2.0845$.

**Note.** A useful check is provided by noting the sum of the coefficients and terms on the right, operating on those numbers as on the equations and checking that the derived equations have the correct sum.

**Otherwise :** We have $\begin{bmatrix} 1 & 4 & -1 \\ 1 & 1 & -6 \\ 3 & -1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -5 \\ -12 \\ 4 \end{bmatrix}$

Operate $R_2 - R_1$ and $R_3 - 3R_1$, $\begin{bmatrix} 1 & 4 & -1 \\ 0 & -3 & -5 \\ 0 & -13 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -5 \\ -7 \\ 19 \end{bmatrix}$

Operate $R_3 - \frac{13}{3}R_2$, $\begin{bmatrix} 1 & 4 & -1 \\ 0 & -3 & -5 \\ 0 & 0 & 71/3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -5 \\ -7 \\ 148/3 \end{bmatrix}$

Thus, we have $z = 148/71 = 2.0845$,

$$3y = 7 - 5z = 7 - 10.4225 = -3.4225 \; i.e., \; y = -1.1408$$

and $$x = -5 - 4y + z = -5 + 4(1.1408) + 2.0845 = 1.6479$$

Hence $x = 1.6479$, $y = -1.1408$, $z = 2.0845$.

**Example 3.20.** *Solve $10x - 7y + 3z + 5u = 6$, $-6x + 8y - z - 4u = 5$, $3x + y + 4y + 11u = 2$, $5x - 9y - 2z + 4u = 7$ by Gauss elimination method.* *(S.V.T.U., B. Tech., 2007)*

**Sol.** We have

| | Check sum | |
|---|---|---|
| $10x - 7y + 3z + 5u = 6$ | $17$ | ...(i) |
| $-6x + 8y - z - 4u = 5$ | $2$ | ...(ii) |
| $3x + y + 4z + 11u = 2$ | $21$ | ...(iii) |
| $5x - 9y - 2z + 4u = 7$ | $5$ | ...(iv) |

*Step I.* To eliminate $x$, operate $\left[(ii) - \left(\frac{-6}{10}\right)(i)\right], \left[(iii) - \frac{3}{10}(i)\right], \left[(iv) - \frac{5}{10}(i)\right]$ :

| | | *Check sum* | |
|---|---|---|---|
| | $3.8y + 0.8z - u = 8.6$ | 12.2 | ...(*v*) |
| | $3.1y + 3.1z + 9.5u = 0.2$ | 15.9 | ...(*vi*) |
| | $-5.5y - 3.5z + 1.5u = 4$ | $-3.5$ | ...(*vii*) |

*Step II.* To eliminate $y$, operate $\left[(vi) - \frac{3.1}{3.8}(v)\right], \left[(vii) - \left(-\frac{5.5}{3.8}\right)(v)\right]$:

$$2.4473684z + 10.315789u = -6.8157895 \quad ...(viii)$$
$$-2.3421053z + 0.0526315u = 16.447368 \quad ...(ix)$$

*Step III.* To eliminate $z$, operate $\left[(ix) - \left(\frac{-2.3421053}{2.4473684}\right)(viii)\right]$:

$$9.9249319u = 9.9245977$$

*Step IV.* By back-substitution, we get

$$u = 1, z = -7, y = 4 \text{ and } x = 5.$$

**Example 3.21.** *Using Gauss elimination method, solve the equations : $x + 2y + 3z - u = 10$, $2x + 3y - 3z - u = 1$, $2x - y + 2z + 3u = 7$, $3x + 2y - 4z + 3u = 2$.*

**Sol.** We have $\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 3 & -3 & -1 \\ 2 & -1 & 2 & 3 \\ 3 & 2 & -4 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ u \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \\ 7 \\ 2 \end{bmatrix}$

Operate $R_2 - 2R_1, R_3 - 2R_1, R_4 - 3R_1$

$$\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & -1 & -9 & 1 \\ 0 & -5 & -4 & 5 \\ 0 & -4 & -13 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ u \end{bmatrix} = \begin{bmatrix} -10 \\ -19 \\ -13 \\ -28 \end{bmatrix}$$

Operate $R_3 - 5R_2, R_4 - 4R_2$ $\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & -1 & -9 & 1 \\ 0 & 0 & 41 & 0 \\ 0 & 0 & 23 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ u \end{bmatrix} = \begin{bmatrix} 10 \\ -19 \\ 82 \\ 48 \end{bmatrix}$

Thus, we have $41z = 82$ *i.e.,* $z = 2$.

$$23z + 2u = 48 \text{ i.e., } 46 + 2u = 48, \qquad \therefore\ u = 1$$
$$-y - 9z + u = -19 \text{ i.e., } -y - 18 + 1 = -19, \qquad \therefore\ y = 2$$
$$x + 2y + 3z - u = 10 \text{ i.e., } x + 4 + 6 - 1 = 10, \qquad \therefore\ x = 1$$

Hence $x = 1, y = 2, z = 2, u = 1$.

**(4) Gauss-Jordan method.** This is a modification of the 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*** *i.e. each equation involving only one unknown. From these equations, the unknowns x, y, z can be obtained readily.*



Thus in this method, the labour of back-substitution for finding the unknowns is saved at the cost of additional calculations.

**Obs.** For a system of 10 equations, the number of multiplications required for Gauss-Jordan method is about 500 whereas for Gauss elimination method we need only 333 multiplications. This shows that though Gauss-Jordan method appears to be easier but requires 50% more operations than the Gauss elimination method. As such Gauss elimination method is preferred for large systems.

**Example 3.22.** *Apply Gauss-Jordan method to solve the equations $x + y + z = 9$ ; $2x - 3y + 4z = 13$ ; $3x + 4y + 5z = 40$.* (V.T.U., B.E., 2009)

**Sol.** We have

$$x + y + z = 9 \quad ...(i)$$
$$2x - 3y + 4z = 13 \quad ...(ii)$$
$$3x + 4y + 5z = 40 \quad ...(iii)$$

*Step I.* To eliminate $x$ from (*ii*) and (*iii*), operate (*ii*) $-$ 2(*i*) and (*iii*) $-$ 3(*i*) :

$$x + y + z = 9$$
$$-5y + 2z = -5 \quad ...(iv)$$
$$y + 2z = 13 \quad ...(v)$$
$$...(vi)$$

*Step II.* To eliminate $y$ from (*iv*) and (*vi*), operate (*iv*) $+ \frac{1}{5}$ (*v*) and (*vi*) $+ \frac{1}{5}$ (*v*) :

$$x + \frac{7}{5}z = 8 \quad ...(vii)$$
$$-5y + 2z = -5 \quad ...(viii)$$
$$\frac{12}{5}z = 12 \quad ...(ix)$$

*Step III.* To eliminate $z$ from (*vii*) and (*viii*), operate (*vii*) $- \frac{7}{12}$ (*ix*) and (*viii*) $- \frac{5}{6}$ (*ix*) :

$$x = 1$$
$$-5y = -15$$
$$\frac{12}{5}z = 12$$

Hence the solution is $x = 1, y = 3, z = 5$.

**Otherwise :** Rewriting the equations as $\begin{bmatrix} 1 & 1 & 1 \\ 2 & -3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 13 \\ 40 \end{bmatrix}$

Operate $R_2 - 2R_1, R_3 - 3R_1$, $\begin{bmatrix} 1 & 1 & 1 \\ 0 & -5 & 2 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ -5 \\ 13 \end{bmatrix}$

Operate $R_3 + \frac{1}{5} R_2$, $\begin{bmatrix} 1 & 1 & 1 \\ 0 & -5 & 2 \\ 0 & 0 & 12/5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ -5 \\ 12 \end{bmatrix}$

Operate $-R_2, 5R_3$ $\qquad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 5 & -2 \\ 0 & 0 & 12 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 5 \\ 60 \end{bmatrix}$

Operate $R_2 + \frac{1}{6}R_3, \frac{1}{12}R_3$ $\qquad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 15 \\ 5 \end{bmatrix}$

Operate $\frac{1}{5}R_2$ $\qquad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 9 \\ 3 \\ 5 \end{bmatrix}$

Operate $R_1 - R_2 - R_3$ $\qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$

Hence $x = 1, y = 3, z = 5$.

**Obs.** Here the process of elimination of variables amounts to *reducing the given coefficient metric to a* **diagonal matrix** *by elementary row transformations only.*

**Example 3.23.** *Solve the equations of Example 3.20, by Gauss-Jordan method.*

**Sol.** We have

$$10x - 7y + 3z + 5u = 6 \quad \text{...(i)}$$
$$-6x + 8y - z - 4u = 5 \quad \text{...(ii)}$$
$$3x + y + 4z + 11u = 2 \quad \text{...(iii)}$$
$$5x - 9y - 2z + 4u = 7 \quad \text{...(iv)}$$

*Step I.* To eliminate $x$, operate $\left[(ii) - \left(\frac{-6}{10}\right)(i)\right], \left[(iii) - \left(\frac{3}{10}\right)(i)\right], \left[(iv) - \left(\frac{5}{10}\right)(i)\right]$:

$$10x - 7y + 3z + 5u = 6 \quad \text{...(v)}$$
$$3.8y + 0.8z - u = 8.6 \quad \text{...(vi)}$$
$$3.1y + 3.1z + 9.5u = 0.2 \quad \text{...(vii)}$$
$$-5.5y - 3.5z + 1.5u = 4 \quad \text{...(viii)}$$

*Step II.* To eliminate $y$, operate $\left[(v) - \left(\frac{-7}{3.8}\right)(vi)\right], \left[(vii) - \left(\frac{3.1}{3.8}\right)(vi)\right], \left[(viii) - \left(\frac{5.5}{3.8}\right)(vi)\right]$:

$$10x + 4.4736842z + 3.1578947u = 21.842105 \quad \text{...(ix)}$$
$$3.8y + 0.8z - u = 8.6 \quad \text{...(x)}$$
$$2.4473684z + 10.315789u = -6.8157895 \quad \text{...(xi)}$$
$$-2.3421053z + 0.0526315u = 16.447368 \quad \text{...(xii)}$$



*Step III.* To eliminate $z$, operate $\left[(ix) - \left(\frac{4.473684}{2.4473684}\right)(xi)\right], \left[(x) - \left(\frac{0.8}{2.4473684}\right)(xi)\right], \left[(xii) - \left(\frac{-2.3421053}{2.4473684}\right)(xi)\right]$:

$$10x - 15.698923u = 34.301075$$
$$3.8y - 4.3720429u = 10.827957$$
$$2.4473684z + 10.315789u = -6.8157895$$
$$9.9247309u = 9.9245975$$

*Step IV.* From the last equation $u = 1$ nearly.

Substitution of $u = 1$ in the above three equations gives $x = 5, y = 4, z = -7$.

**(5) Factorization method**[*]. This method is based on the fact that every square matrix $A$ can be expressed as the product of a lower triangular matrix and an upper triangular matrix, provided all the principal minors of $A$ are non-singular, *i.e.* if $A = [a_{ij}]$, then

$$a_{11} \neq 0, \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \neq 0, \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \neq 0, \text{ etc.}$$

Also such a factorisation if it exists, is unique.

Now consider the equations

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

which can be written as $AX = B$ ...(1)

where $\quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$

Let $\quad A = LU,$ ...(2)

where $\quad L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$ and $U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$

Then (1) becomes $\quad LUX = B$ ...(3)

Writing $\quad UX = V,$ ...(4), (3) becomes $\quad LV = B$

which is equivalent to the equations $\quad v_1 = b_1, l_{21}v_1 + v_2 = b_2, l_{31}v_1 + l_{32}v_2 + v_3 = b_3$

Solving these for $v_1, v_2, v_3$, we know $V$. Then, (4) becomes

$$u_{11}x_1 + u_{12}x_2 + u_{13}x_3 = v_1, \; u_{22}x_2 + u_{23}x_3 = v_2, \; u_{33}x_3 = v_3,$$

from which $x_3, x_2$ and $x_1$ can be found by *back-substitution*.

*To compute the matrices L and U*, we write (2) as

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

[*]Another name given to this decomposition is **Doolittle's method.**

Multiplying the matrices on the left and equating corresponding elements from both sides, we obtain

(i) $u_{11} = a_{11}$, $u_{12} = a_{12}$, $u_{13} = a_{13}$

(ii) $l_{21}u_{11} = a_{21}$ or $l_{21} = a_{21}/a_{11}$ ; $l_{31}u_{11} = a_{31}$ or $l_{31} = a_{31}/a_{11}$

(iii) $l_{21}u_{12} + u_{22} = a_{22}$ or $u_{22} = a_{22} - \dfrac{a_{21}}{a_{11}} a_{12}$

$l_{21}u_{13} + u_{23} = a_{23}$ or $u_{23} = a_{23} - \dfrac{a_{21}}{a_{11}} a_{13}$

(iv) $l_{31}u_{12} + l_{32}u_{22} = a_{32}$ or $l_{32} = \dfrac{1}{u_{22}}\left[a_{32} - \dfrac{a_{31}}{a_{11}} a_{12}\right]$

(v) $l_{31}u_{13} + l_{32}u_{23} + u_{33} = a_{33}$ which gives $u_{33}$.

*Thus we compute the elements of L and U in the following set order :*

(i) First row of $U$,
(ii) First column of $L$,
(iii) Second row of $U$,
(iv) Second column of $L$,
(v) Third row of $U$.

This procedure can easily be generalised.

**Obs.** This method is superior to Gauss elimination method and is often used for the solution of linear systems and for finding the inverse of a matrix. The number of operations involved in terms of multiplications for a system of 10 equations by this method is about 110 as compared 333 operations of the Gauss method. Among the direct methods, factorization method is also preferred as the software for computers.

**Example 3.24.** *Apply factorization method to solve the equations :*

$3x + 2y + 7z = 4$ ; $2x + 3y + z = 5$ ; $3x + 4y + z = 7$. *(Madras, B.E., 2000 S)*

**Sol.** Let $\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} 3 & 2 & 7 \\ 2 & 3 & 1 \\ 3 & 4 & 1 \end{bmatrix}$ (*i.e.* A),

so that

(i) $R_1$ of $U$ : $u_{11} = 3$, $u_{12} = 2$, $u_{13} = 7$.

(ii) $C_1$ of $L$ : $l_{21}u_{11} = 2$, $\therefore\ l_{21} = 2/3$,
$l_{31}u_{11} = 3$, $\therefore\ l_{31} = 1$.

(iii) $R_2$ of $U$ : $l_{21}u_{12} + u_{22} = 3$, $\therefore\ u_{22} = 5/3$,
$l_{21}u_{13} + u_{23} = 1$, $\therefore\ u_{23} = -11/3$.

(iv) $C_2$ of $L$ : $l_{31}u_{12} + l_{32}u_{22} = 4$ $\therefore\ l_{32} = 6/5$.

(v) $R_3$ of $U$ : $l_{31}u_{13} + l_{32}u_{23} + u_{33} = 1$ $\therefore\ u_{33} = -8/5$.

Thus $A = \begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1 & 6/5 & 1 \end{bmatrix}\begin{bmatrix} 3 & 2 & 7 \\ 0 & 5/3 & -11/3 \\ 0 & 0 & -8/5 \end{bmatrix}$

Writing $UX = V$, the given system becomes

$$\begin{bmatrix} 1 & 0 & 0 \\ 2/3 & 1 & 0 \\ 1 & 6/5 & 1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 7 \end{bmatrix}$$



Solving this system, we have $v_1 = 4$,

$$\frac{2}{3} v_1 + v_2 = 5 \quad \text{or} \quad v_2 = \frac{7}{3}$$

$$v_1 + \frac{6}{5} v_2 + v_3 = 7 \quad \text{or} \quad v_3 = \frac{1}{5}$$

Hence the original system becomes

$$\begin{bmatrix} 3 & 2 & 7 \\ 0 & 5/3 & -11/3 \\ 0 & 0 & -8/5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 7/3 \\ 1/5 \end{bmatrix}$$

*i.e.* $$3x + 2y + 7z = 4, \quad \frac{5}{3}y - \frac{11}{3}z = \frac{7}{3}, \quad -\frac{8}{5}z = \frac{1}{5}$$

By back-substitution, we have

$$z = -1/8,\ y = 9/8 \quad \text{and} \quad x = 7/8.$$

**Example 3.25.** *Solve the equations of Example 3.20 by Factorization method.*

**Sol.** Let $\begin{bmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{bmatrix} = \begin{bmatrix} 10 & -7 & 3 & 5 \\ -6 & 8 & -1 & -4 \\ 3 & 1 & 4 & 11 \\ 5 & -9 & -2 & 4 \end{bmatrix}$ (*i.e.* A)

so that

(i) $R_1$ of $U$ : $u_{11} = 10, u_{12} = -7, u_{13} = 3, u_{14} = 5$

(ii) $C_1$ of $L$ : $l_{21} = -0.6, l_{31} = 0.3, l_{41} = 0.5$

(iii) $R_2$ of $U$ : $u_{22} = 3.8, u_{23} = 0.8, u_{24} = -1$

(iv) $C_2$ of $L$ : $l_{32} = 0.81579, l_{42} = -1.44737$

(v) $R_3$ of $U$ : $u_{33} = 2.44737, u_{34} = 10.31579$

(vi) $C_3$ of $L$ : $l_{43} = -0.95699$

(vii) $R_4$ of $U$ : $u_{44} = 9.92474$

Thus $$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.6 & 1 & 0 & 0 \\ 0.3 & 0.81579 & 1 & 0 \\ 0.5 & -1.44737 & -0.95699 & 1 \end{bmatrix}\begin{bmatrix} 10 & -7 & 3 & 5 \\ 0 & 3.8 & 0.8 & -1 \\ 0 & 0 & 2.44737 & 10.31579 \\ 0 & 0 & 0 & 9.92474 \end{bmatrix}$$

Writing $UX = V$, the given system becomes

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -0.6 & 1 & 0 & 0 \\ 0.3 & 0.81577 & 1 & 0 \\ 0.5 & -1.44737 & -0.95699 & 1 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ 2 \\ 7 \end{bmatrix}$$

Solving this system, we get

$$v_1 = 6,\ v_2 = 8.6,\ v_3 = -6.81579,\ v_4 = 9.92474.$$

Hence the original system becomes

$$\begin{bmatrix} 10 & -7 & 3 & 5 \\ 0 & 3.8 & 0.8 & -1 \\ 0 & 0 & 2.44737 & 10.31579 \\ 0 & 0 & 0 & 9.92474 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ u \end{bmatrix} = \begin{bmatrix} 6 \\ 8.6 \\ -6.81579 \\ 9.92474 \end{bmatrix}$$

*i.e.*
$$10x - 7y + 3z + 5u = 6, \quad 3.8y + 0.8z - u = 8.6,$$
$$2.44737z + 10.31579u = -6.81579, \; u = 1.$$

By back-substitution, we get

$$u = 1, z = -7, y = 4, x = 5.$$

## PROBLEMS 3.3

Solve the following equations by Cramer's rule :

1. $x + 3y + 6z = 2 \,;\, 3x - y + 4z = 9 \,;\, x - 4y + 2z = 7.$
2. $x + y + z = 6.6 \,;\, x - y + z = 2.2 \,;\, x + 2y + 3z = 15.2.$
3. $x^2z^3/y = e^8 \,;\, y^2z/x = e^4 \,;\, x^3y/z^4 = 1.$
4. $2vw - wu + uv = 3uvw \,;\, 3vw + 2wu + 4uv = 19uvw \,;\, 6vw + 7wu - uv = 17uvw.$
5. $3x + 2y - z + t = 1 \,;\, x - y - 2z + 4t = 3 \,;\, 2x - 3y + z - 2t = -2 \,;\, 5x - 2y + 3z + 2t = 0.$

Solve the following equations by matrix inversion method :

6. $x + y + z = 3 \,;\, x + 2y + 3z = 4 \,;\, x + 4y + 9z = 6.$ (*Bhopal, B.E., 2003*)
7. $x + y + z = 1 \,;\, x + 2y + 3z = 6 \,;\, x + 3y + 4z = 6.$ (*P.T.U., B. Tech., 2006*)
8. $2x - y + 3z = 8 \,;\, x - 2y - z = -4 \,;\, 3x + y - 4z = 0.$ (*Bombay, B. Tech., 2005*)
9. $2x_1 + x_2 + 2x_3 + x_4 = 6 \,;\, 4x_1 + 3x_2 + 3x_3 - 3x_4 = -1 \,;\, 6x_1 - 6x_2 + 6x_3 + 12x_4 = 36,\; 2x_1 + 2x_2 - x_3 + x_4 = 10.$ (*U.P.T.U., B. Tech., 2001*)
10. In a given electrical network, the equations for the currents $i_1, i_2, i_3$ are
$$3i_1 + i_2 + i_3 = 8 \,;\, 2i_1 - 3i_2 - 2i_3 = -5 \,;\, 7i_1 + 2i_2 - 5i_3 = 0.$$
Calculate $i_1$ and $i_3$ by (*a*) Cramer's rule, (*b*) matrix inversion.

Solve the following equations by Gauss elimination method :

11. $2x + y + z = 10 \,;\, 3x + 2y + 3z = 18 \,;\, x + 4y + 9z = 16.$ (*P.T.U., B. Tech., 2003*)
12. $2x + 2y + z = 12 \,;\, 3x + 2y + 2z = 8 \,;\, 5x + 10y - 8z = 10.$ (*W.B.T.U., B. Tech., 2004*)
13. $2x - y + 3z = 9 \,;\, x + y + z = 6 \,;\, x - y + z = 2.$ (*Bhopal, B.E., 2009*)
14. $2x_1 + 4x_2 + x_3 = 3 \,;\, 3x_1 + 2x_2 - 2x_3 = -2 \,;\, x_1 - x_2 + x_3 = 6.$ (*Marathwada, B. Tech., 2008*)
15. $5x_1 + x_2 + x_3 + x_4 = 4 \,;\, x_1 + 7x_2 + x_3 + x_4 = 12 \,;\, x_1 + x_2 + 6x_3 + x_4 = -5 \,;\, x_1 + x_2 + x_3 + 4x_4 = -6.$

Solve the following equations by Gauss-Jordan method :

16. $2x + y + z = 10 \,;\, 3x + 2y + 3z = 18 \,;\, x + 4y + 9z = 16.$ (*V.T.U., B. Tech., 2008*)
17. $2x - 3y + z = -1 \,;\, x + 4y + 5z = 25 \,;\, 3x - 4y + z = 2.$ (*Kerala, B. Tech., 2003*)
18. $x + y + z = 9 \,;\, 2x + y - z = 0 \,;\, 2x + 5y + 7z = 52.$ (*V.T.U., B.E., 2009*)
19. $x + 3y + 3z = 16,\; x + 4y + 3z = 18,\; x + 3y + 4z = 19.$ (*Anna, B. Tech., 2005*)
20. $2x_1 + x_2 + 5x_3 + x_4 = 5 \,;\, x_1 + x_2 - 3x_3 + 4x_4 = -1 \,;$
$3x_1 + 6x_2 - 2x_3 + x_4 = 8 \,;\, 2x_1 + 2x_2 + 2x_3 - 3x_4 = 2.$ (*Madras, MCA, 1997*)

Solve the following equations by factorization method :

21. $2x + 3y + z = 9 \,;\, x + 2y + 3z = 6 \,;\, 3x + y + 2z = 8.$



22. $10x + y + z = 12 \,;\, 2x + 10y + z = 13 \,;\, 2x + 2y + 10z = 14.$
23. $10x + y + 2z = 13 \,;\, 3x + 10y + z = 14 \,;\, 2x + 3y + 10z = 15.$ (*Andhra B.E., 2004*)
24. $2x_1 - x_2 + x_3 = -1 \,;\, 2x_2 - x_3 + x_4 = 1 \,;\, x_1 + 2x_3 - x_4 = -1 \,;\, x_1 + x_2 + 2x_4 = 3.$

## 3.5 ITERATIVE METHODS OF SOLUTION

The preceding methods of solving simultaneous linear equations are known as *direct methods*, as these methods yield the solution after a certain amount of fixed computation. On the other hand, an iterative method is that in which we start from an approximation to the true solution and obtain better and better approximations from a computation cycle repeated as often as may be necessary for achieving a desired accuracy. Thus in an iterative method, the amount of computation depends on the degree of accuracy required.

For large systems, iterative methods may be faster than the direct methods. Even the round-off errors in iterative methods are smaller. In fact, iteration is a self correcting process and any error made at any stage of computation gets automatically corrected in the subsequent steps.

Simple iterative methods can be devised for systems in which the coefficients of the leading diagonal are large as compared to others. We now describe three such methods :

(1) **Jacobi's iteration method.** Consider the equations

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \qquad \ldots(1)$$

If $a_1, b_2, c_3$ are large as compared to other coefficients, solve for $x, y, z$ respectively. Then the system can be written as

$$\left.\begin{aligned} x &= \frac{1}{a_1}(d_1 - b_1y - c_1z) \\ y &= \frac{1}{b_2}(d_2 - a_2x - c_2z) \\ z &= \frac{1}{c_3}(d_3 - a_3x - b_3y) \end{aligned}\right\} \qquad \ldots(2)$$

Let us start with the initial approximations $x_0, y_0, z_0$ for the values of $x, y, z$ respectively. Substituting these on the right sides of (2), the first approximations are given by

$$x_1 = \frac{1}{a_1}(d_1 - b_1y_0 - c_1z_0)$$

$$y_1 = \frac{1}{b_2}(d_2 - a_2x_0 - c_2z_0)$$

$$z_1 = \frac{1}{c_3}(d_3 - a_3x_0 - b_3y_0)$$

Substituting the values $x_1, y_1, z_1$ on the right sides of (2), the second approximations are given by

$$x_2 = \frac{1}{a_1}(d_1 - b_1y_1 - c_1z_1)$$

$$y_2 = \frac{1}{b_2}(d_2 - a_2x_1 - c_2z_1)$$

$$z_2 = \frac{1}{c_3}(d_3 - a_3x_1 - b_3y_1)$$

This process is repeated till the difference between two consecutive approximations is negligible.

**Obs.** In the absence of any better estimates for $x_0, y_0, z_0$, these may each be taken as zero.

**Example 3.26.** *Solve, by Jacobi's iteration method, the equations $20x + y - 2z = 17$; $3x + 20y - z = -18$; $2x - 3y + 20z = 25$.* *(Bhopal, B.E., 2009)*

**Sol.** We write the given equations in the form

$$\left.\begin{aligned} x &= \frac{1}{20}(17 - y + 2z) \\ y &= \frac{1}{20}(-18 - 3x + z) \\ z &= \frac{1}{20}(25 - 2x + 3y) \end{aligned}\right\} \quad ...(i)$$

We start from an approximation $x_0 = y_0 = z_0 = 0$.

Substituting these on the right sides of the equations (*i*), we get

$$x_1 = \frac{17}{20} = 0.85,\quad y_1 = -\frac{18}{20} = -0.9,\quad z_1 = \frac{25}{20} = 1.25$$

Putting these values on the right sides of the equations (*i*), we obtain

$$x_2 = \frac{1}{20}(17 - y_1 + 2z_1) = 1.02$$

$$y_2 = \frac{1}{20}(-18 - 3x_1 + z_1) = -0.965$$

$$z_2 = \frac{1}{20}(25 - 2x_1 + 3y_1) = 1.03$$

Substituting these values on the right sides of the equations (*i*), we have

$$x_3 = \frac{1}{20}(17 - y_2 + 2z_2) = 1.00125$$

$$y_3 = \frac{1}{20}(-18 - 3x_2 + z_2) = -1.0015$$

$$z_3 = \frac{1}{20}(25 - 2x_2 + 3y_2) = 1.00325$$



Substituting these values, we get

$$x_4 = \frac{1}{20}(17 - y_3 + 2z_3) = 1.0004$$

$$y_4 = \frac{1}{20}(-18 - 3x_3 + z_3) = -1.000025$$

$$z_4 = \frac{1}{20}(25 - 2x_3 + 3y_3) = 0.9965$$

Putting these values, we have

$$x_5 = \frac{1}{20}(-17 - y_4 + 2z_4) = 0.999966$$

$$y_5 = \frac{1}{20}(-18 - 3x_4 + z_4) = -1.000078$$

$$z_5 = \frac{1}{20}(25 - 2x_4 + 3y_4) = 0.999956$$

Again substituting these values, we get

$$x_6 = \frac{1}{20}(-17 - y_5 + 2z_5) = 1.0000$$

$$y_6 = \frac{1}{20}(-18 - 3x_5 + z_5) = -0.999997$$

$$z_6 = \frac{1}{20}(25 - 2x_5 + 3y_5) = 0.999992$$

The values in the 5th and 6th iterations being practically the same, we can stop. Hence the solution is

$$x = 1, y = -1, z = 1.$$

**Example 3.27.** *Solve by Jacobi's iteration method, the equations $10x + y - z = 11.19$, $x + 10y + z = 28.08$, $-x + y + 10z = 35.61$, correct to two decimal places.* *(Anna, B.Tech., 2007)*

**Sol.** Rewriting the given equations as

$$x = \frac{1}{10}(11.19 - y + z),\ y = \frac{1}{10}(28.08 - x - z),\ z = \frac{1}{10}(35.61 + x - y)$$

We start from an approximation, $x_0 = y_0 = z_0 = 0$.

*First iteration*

$$x_1 = \frac{11.19}{10} = 1.119,\ y_1 = \frac{28.08}{10} = 2.808,\ z_1 = \frac{35.61}{10} = 3.561$$

*Second iteration*

$$x_2 = \frac{1}{10}(11.19 - y_1 + z_1) = 1.19$$

$$y_2 = \frac{1}{10}(28.08 - x_1 - z_1) = 2.34$$

$$z_2 = \frac{1}{10}(35.61 + x_1 - y_1) = 3.39$$

*Third iteration*

$$x_3 = \frac{1}{10}(11.19 - y_2 + z_2) = 1.22$$

$$y_3 = \frac{1}{10}(28.08 - x_2 - z_2) = 2.35$$

$$z_3 = \frac{1}{10}(35.61 + x_2 - y_2) = 3.45$$

*Fourth iteration*

$$x_4 = \frac{1}{10}(11.19 - y_3 + z_3) = 1.23$$

$$y_4 = \frac{1}{10}(28.08 - x_3 - z_3) = 2.34$$

$$z_4 = \frac{1}{10}(35.61 + x_3 - y_3) = 3.45$$

*Fifth iteration*

$$x_5 = \frac{1}{10}(11.19 - y_4 + z_4) = 1.23$$

$$y_5 = \frac{1}{10}(28.08 - x_4 - z_4) = 2.34$$

$$z_5 = \frac{1}{10}(35.61 + x_4 - y_4) = 3.45$$

Hence $x = 1.23, y = 2.34, z = 3.45$

**Example 3.28.** *Solve the equations*

$$10x - 2x_2 - x_3 - x_4 = 3$$
$$-2x_1 + 10x_2 - x_3 - x_4 = 15$$
$$-x_1 - x_2 + 10x_3 - 2x_4 = 27$$
$$-x_1 - x_2 - 2x_3 + 10x_4 = -9,$$ *by Gauss-Jacobi iteration method.*

**Sol.** Rewriting the given equation as

$$x_1 = \frac{1}{10}(3 + 2x_2 + x_3 + x_4)$$

$$x_2 = \frac{1}{10}(15 + 2x_1 + x_3 + x_4)$$

$$x_3 = \frac{1}{10}(27 + x_1 + x_2 + 2x_4)$$

$$x_4 = \frac{1}{10}(-9 + x_1 + x_2 + 2x_3)$$

We start from an approximation $x_1 = x_2 = x_3 = x_4 = 0$.

*First iteration*

$$x_1 = 0.3, x_2 = 1.5, x_3 = 2.7, x_4 = -0.9.$$



*Second iteration*

$$x_1 = \frac{1}{10}[3 + 2(1.5) + 2.7 + (-0.9)] = 0.78$$
$$x_2 = \frac{1}{10}[15 + 2(0.3) + 2.7 + (-0.9)] = 1.74$$
$$x_3 = \frac{1}{10}[27 + 0.3 + 1.5 + 2(-0.9)] = 2.7$$
$$x_4 = \frac{1}{10}[-9 + 0.3 + 1.5 + 2(-0.9)] = -0.18$$

Proceeding in this way, we get

*Third iteration*

$$x_1 = 0.9, x_2 = 1.908, x_3 = 2.916, x_4 = -0.108$$

*Fourth iteration*

$$x_1 = 0.9624, x_2 = 1.9608, x_3 = 2.9592, x_4 = -0.036$$

*Fifth iteration*

$$x_1 = 0.9845, x_2 = 1.9848, x_3 = 2.9851, x_4 = -0.0158$$

*Sixth iteration*

$$x_1 = 0.9939, x_2 = 1.9938, x_3 = 2.9938, x_4 = -0.006$$

*Seventh iteration*

$$x_1 = 0.9939, x_2 = 1.9975, x_3 = 2.9976, x_4 = -0.0025$$

*Eighth iteration*

$$x_1 = 0.999, x_2 = 1.999, x_3 = 2.999, x_4 = -0.001$$

*Ninth iteration*

$$x_1 = 0.9996, x_2 = 1.9996, x_3 = 2.9996, x_4 = -0.004$$

*Tenth iteration*

$$x_1 = 0.9998, x_2 = 1.9998, x_3 = 2.9998, x_4 = -0.0001$$

Hence $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 0$.

**(2) Gauss-Seidal iteration method.** This is a modification of Jacobi's method. As before the system of equations :

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \quad ...(1)$$

is written as

$$\left.\begin{aligned} x &= \frac{1}{a_1}(d_1 - b_1y - c_1z) \\ y &= \frac{1}{b_2}(d_2 - a_2x - c_2z) \\ z &= \frac{1}{c_3}(d_3 - a_3x - b_3y) \end{aligned}\right\} \quad ...(2)$$

Here also we start with the initial approximations $x_0, y_0, z_0$ for $x, y, z$ respectively which may each be taken as zero. Substituting $y = y_0, z = z_0$ in the first of the equations (2), we get

$$x_1 = \frac{1}{a_1}(d_1 - b_1y_0 - c_1z_0)$$

Then putting $x = x_1, z = z_0$ in the second of the equations (2), we have

$$y_1 = \frac{1}{b_2}(d_2 - a_2x_1 - c_2z_0)$$

Next substituting $x = x_1, y = y_1$ in the third of the equations (2), we obtain

$$z_1 = \frac{1}{c_3}(d_3 - a_3x_1 - b_3y_1)$$

and so on *i.e.* as soon as a new approximation for an unknown is found, it is immediately used in the next step.

This process of iteration is repeated till the values of $x, y, z$ are obtained to desired degree of accuracy.

**Obs. 1.** Since the most recent approximations of the unknowns are used while proceeding to the next step, the *convergence in the Gauss-Seidal method is twice as fast as in Jacobi's method.*

**Obs. 2.** *Jacobi and Gauss-Seidal methods converge for any choice of the initial approximations if in each equation of the system,* **the absolute value of the largest co-efficient is almost equal to or in atleast one equation greater than the sum of the absolute values of all the remaining coefficients.**

**Example 3.29.** *Apply Gauss-Seidal iteration method to solve the equations of Example 3.26.* (V.T.U., B. Tech., 2006)

**Sol.** We write the given equations in the form

$$x = \frac{1}{20}(17 - y + 2z) \quad ...(i)$$

$$y = \frac{1}{20}(-18 - 3x + z) \quad ...(ii)$$

$$z = \frac{1}{20}(25 - 2x + 3y) \quad ...(iii)$$

*First iteration*

Putting $y = y_0, z = z_0$ in (*i*), we get $x_1 = \frac{1}{20}(17 - y_0 + 2z_0) = 0.8500$

Putting $x = x_1, z = z_0$ in (*ii*), we have $y_1 = \frac{1}{20}(-18 - 3x_1 + z_0) = -1.0275$

Putting $x = x_1, y = y_1$ in (*iii*), we obtain $z_1 = \frac{1}{20}(25 - 2x_1 + 3y_1) = 1.0109$

*Second iteration*

Putting $y = y_1, z = z_1$ in (*i*), we get $x_2 = \frac{1}{20}(17 - y_1 + 2z_1) = 1.0025$

Putting $x = x_2, z = z_1$ in (*ii*), we obtain $y_2 = \frac{1}{20}(-18 - 3x_2 + z_1) = -0.9998$

Putting $x = x_2, y = y_2$ in (*iii*), we get $z_2 = \frac{1}{20}(25 - 2x_2 + 3y_2) = 0.9998$



*Third iteration*, we get

$$x_3 = \frac{1}{20}(17 - y_2 + 2z_2) = 1.0000$$

$$y_3 = \frac{1}{20}(-18 - 3x_3 + z_2) = -1.0000$$

$$z_3 = \frac{1}{20}(25 - 2x_3 + 3y_3) = 1.0000$$

The values in the 2nd and 3rd iterations being practically the same, we can stop. Hence the solution is $x = 1, y = -1, z = 1$.

**Example 3.30.** *Solve the equations $27x + 6y - z = 85$, $x + y + 54z = 110$ ; $6x + 15y + 2z = 72$ by Gauss-Jacobi method and Gauss-Seidel method.* (Anna, B.Tech., 2006)

**Sol.** Rewriting the given equations as

$$x = \frac{1}{27}(85 - 6y + z) \quad ...(i)$$

$$y = \frac{1}{15}(72 - 6x - 2z) \quad ...(ii)$$

$$z = \frac{1}{54}(110 - x - y) \quad ...(iii)$$

(*a*) *Gauss-Jacobi's method*

We start from an approximation $x_0 = y_0 = z_0 = 0$

*First iteration*

$$x_1 = \frac{85}{27} = 3.148, y_1 = \frac{72}{15} = 4.8, z_1 = \frac{110}{54} = 2.037$$

*Second iteration*

$$x_2 = \frac{1}{27}(85 - 6y_1 + z_1) = 2.157$$

$$y_2 = \frac{1}{15}(72 - 6x_1 - 2z_1) = 3.269$$

$$z_2 = \frac{1}{54}(110 - x_1 - y_1) = 1.890$$

*Third iteration*

$$x_3 = \frac{1}{27}(85 - 6y_2 + z_2) = 2.492$$

$$y_3 = \frac{1}{15}(72 - 6x_2 - 2z_2) = 3.685$$

$$z_3 = \frac{1}{54}(110 - x_2 - y_2) = 1.937$$

*Fourth iteration*

$$x_4 = \frac{1}{27}(85 - 6y_3 + z_3) = 2.401$$

$$y_4 = \frac{1}{15}(72 - 6x_3 - 2z_3) = 3.545$$

$$z_4 = \frac{1}{54}(110 - x_3 - y_3) = 1.923$$

*Fifth iteration*

$$x_5 = \frac{1}{27}(85 - 6y_4 + z_4) = 2.432$$

$$y_5 = \frac{1}{15}(72 - 6x_4 - 2z_4) = 3.583$$

$$z_4 = \frac{1}{54}(110 - x_3 - y_3) = 1.927$$

Repeating this process, the successive iterations are.

$$x_6 = 2.423, y_6 = 3.570, z_6 = 1.926$$
$$x_7 = 2.426, y_7 = 3.574, z_7 = 1.926$$
$$x_8 = 2.425, y_8 = 3.573, z_8 = 1.926$$
$$x_9 = 2.426, y_9 = 3.573, z_9 = 1.926$$

Hence $x = 2.426, y = 3.573, z = 1.926$

*(b) Gauss-Seidal method*

*First iteration*

Putting $y = y_0 = 0, z = z_0 = 0$ in (*i*), $\quad x_1 = \frac{1}{27}(85 - 6y_0 + z_0) = 3.14$

Putting $x = x_1, z = z_0$ in (*ii*), $\quad y_1 = \frac{1}{15}(72 - 6x_1 - 2z_0) = 3.541$

Putting $x = x_1, y = y_1$ in (*iii*), $\quad z_1 = \frac{1}{54}(110 - x_1 - y_1) = 1.913$

*Second iteration*

$$x_2 = \frac{1}{27}(85 - 6y_1 + z_1) = 2.432$$

$$y_2 = \frac{1}{15}(72 - 6x_2 - 2z_1) = 3.572$$

$$z_2 = \frac{1}{54}(110 - x_2 - y_2) = 1.926$$

*Third iteration*

$$x_3 = \frac{1}{27}(85 - 6y_2 + z_2) = 2.426$$

$$y_3 = \frac{1}{15}(72 - 6x_3 - 2z_2) = 3.573$$

$$z_3 = \frac{1}{54}(110 - x_3 - y_3) = 1.926$$



*Fourth iteration*

$$x_4 = \frac{1}{27}(85 - 6y_3 + z_3) = 2.426$$

$$y_4 = \frac{1}{15}(72 - 6x_4 - 2z_3) = 3.573$$

$$z_4 = \frac{1}{54}(110 - x_4 - y_4) = 1.926.$$

Hence $x = 2.426$, $y = 3.573$, $z = 1.926$.

**Obs.** We have seen that *the convergence is quite fast in Gauss-Seidal method as compared to Gauss-Jacobi method.*

**Example 3.31.** *Solve the equations of Example 3.28 by Gauss-Seidal iteration method.* (*Bhopal, B.E. 2009*)

**Sol.** Rewriting the given equations as

$$x_1 = 0.3 + 0.2x_2 + 0.1x_3 + 0.1x_4 \quad ...(i)$$
$$x_2 = 1.5 + 0.2x_1 + 0.1x_3 + 0.1x_4 \quad ...(ii)$$
$$x_3 = 2.7 + 0.1x_1 + 0.1x_2 + 0.2x_4 \quad ...(iii)$$
$$x_4 = -0.9 + 0.1x_1 + 0.1x_2 + 0.2x_3 \quad ...(iv)$$

*First iteration*

Putting $x_2 = 0$, $x_3 = 0, x_4 = 0$ in (*i*), we get $x_1 = 0.3$
Putting $x_1 = 0.3$, $x_3 = 0, x_4 = 0$ in (*ii*), we obtain $x_2 = 1.56$
Putting $x_1 = 0.3$, $x_2 = 1.56, x_4 = 0$ in (*iii*), we obtain $x_3 = 2.886$
Putting $x_1 = 0.3$, $x_2 = 1.56, x_3 = 2.886$ in (*iv*), we get $x_4 = -0.1368$

*Second iteration*

Putting $x_2 = 1.56$, $x_3 = 2.886, x_4 = -0.1368$ in (*i*), we obtain $x_1 = 0.8869$
Putting $x_1 = 0.8869$, $x_3 = 2.886, x_4 = -0.1368$ in (*ii*), we obtain $x_2 = 1.9523$
Putting $x_1 = 0.8869$, $x_2 = 1.9523, x_4 = -0.1368$ in (*iii*), we have $x_3 = 2.9566$
Putting $x_1 = 0.8869$, $x_2 = 1.9523, x_3 = 2.9566$ in (*iv*), we get $x_4 = -0.0248$.

*Third iteration*

Putting $x_2 = 1.9523$, $x_3 = 2.9566, x_4 = -0.0248$ in (*i*), we obtain $x_1 = 0.9836$
Putting $x_1 = 0.9836$, $x_3 = 2.9566, x_4 = -0.0248$ in (*ii*), we obtain $x_2 = 1.9899$
Putting $x_1 = 0.9836$, $x_2 = 1.9899, x_4 = -0.0248$ in (*iii*), we get $x_3 = 2.9924$
Putting $x_1 = 0.9836$, $x_2 = 1.9899, x_3 = 2.9924$ in (*iv*), we get $x_4 = -0.0042$.

*Fourth iteration.* Proceeding as above

$$x_1 = 0.9968, x_2 = 1.9982, x_3 = 2.9987, x_4 = -0.0008.$$

*Fifth iteration* is $\quad x_1 = 0.9994, x_2 = 1.9997, x_3 = 2.9997, x_4 = -0.0001.$

*Sixth iteration* is $\quad x_1 = 0.9999, x_2 = 1.9999, x_3 = 2.9999, x_4 = -0.0001$

Hence the solution is $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 0$.

**(3) Relaxation method***. Consider the equations

$$a_1x + b_1y + c_1z = d_1$$
$$a_2x + b_2y + c_2z = d_2$$
$$a_3x + b_3y + c_3z = d_3$$

*This method was originally developed by R.V. Southwell in 1935, for application to structural engineering problems.

We define the residuals $R_x$, $R_y$, $R_z$ by the relations

$$\left.\begin{aligned} R_x &= d_1 - a_1x - b_1y - c_1z \\ R_y &= d_2 - a_2x - b_2y - c_2z \\ R_z &= d_3 - a_3x - b_3y - c_3z \end{aligned}\right\} \quad ...(1)$$

To start with we assume $x = y = z = 0$ and calculate the initial residuals. Then the residuals are reduced step by step, by giving increments to the variables. For this purpose, we construct the following *operation table* :

| | $\delta R_x$ | $\delta R_y$ | $\delta R_z$ |
|---|---|---|---|
| $\delta x = 1$ | $-a_1$ | $-a_2$ | $-a_3$ |
| $\delta y = 1$ | $-b_1$ | $-b_2$ | $-b_3$ |
| $\delta z = 1$ | $-c_1$ | $-c_2$ | $-c_3$ |

We note from the equations (1) that if $x$ is increased by 1 (keeping $y$ and $z$ constant), $R_x$, $R_y$ and $R_z$ decrease by $a_1$, $a_2$, $a_3$ respectively. This is shown in the above table alongwith the effects on the residuals when $y$ and $z$ are given unit increments. (The table is the transpose of the coefficient matrix).

At each step, the numerically largest residual is reduced to almost zero. To reduce a particular residual, the value of the corresponding variable is changed ; *e.g.* to reduce $R_x$ by $p$, $x$ should be increased by $p/a_1$.

When all the residuals have been reduced to almost zero, the increments in $x$, $y$, $z$ are added separately to give the desired solution.

**Obs. 1.** As a check, the computed values of $x$, $y$, $z$ are substituted in (1) and the residuals are calculated. If these residuals are not all negligible, then there is some mistake and the entire process should be rechecked.

**Obs. 2.** *Relaxation method can be applied successfully only if the diagonal elements of the coefficient matrix dominate the other coefficients in the corresponding row i.e.* if in the equations (1)

$$|a_1| \geq |b_1| + |c_1|$$
$$|b_2| \geq |a_2| + |c_2|$$
$$|c_3| \geq |a_3| + |b_3|$$

where > sign should be valid for at least one row.

**Example 3.32.** *Solve, by Relaxation method, the equations :*

$9x - 2y + z = 50$ ; $x + 5y - 3z = 18$ ; $-2x + 2y + 7z = 19$. *(Madras, B.E., 2000)*

**Sol.** The residuals are given by

$$R_x = 50 - 9x + 2y - z \,;$$
$$R_y = 18 - x - 5y + 3z \,;$$
$$R_z = 19 + 2x - 2y - 7z$$



The operations table is

| | $\delta R_x$ | $\delta R_y$ | $\delta R_z$ |
|---|---|---|---|
| $\delta x = 1$ | $-9$ | $-1$ | 2 |
| $\delta y = 1$ | 2 | $-5$ | $-2$ |
| $\delta z = 1$ | $-1$ | 3 | $-7$ |

The relaxation table is

| | $R_x$ | $R_y$ | $R_z$ | |
|---|---|---|---|---|
| $x = y = z = 0$ | 50 | 18 | 19 | ...(*i*) |
| $\delta x = 5$ | 5 | 13 | 29 | ...(*ii*) |
| $\delta z = 4$ | 1 | 25 | 1 | ...(*iii*) |
| $\delta y = 5$ | 11 | 0 | $-9$ | ...(*iv*) |
| $\delta x = 1$ | 2 | $-1$ | $-7$ | ...(*v*) |
| $\delta z = -1$ | 3 | $-4$ | 0 | ...(*vi*) |
| $\delta y = -0.8$ | 1.4 | 0 | 1.6 | ...(*vii*) |
| $\delta z = 0.23$ | 1.17 | 0.69 | $-0.09$ | ...(*viii*) |
| $\delta x = 0.13$ | 0 | 0.56 | 0.17 | ...(*ix*) |
| $\delta y = 0.112$ | 0.224 | 0 | $-0.054$ | ...(*x*) |

$$\Sigma\delta x = 6.13, \ \Sigma\delta y = 4.31, \ \Sigma\delta z = 3.23.$$

Thus $\qquad x = 6.13, \ y = 4.31, \ z = 3.23$

[**Explanation.** In (*i*), the largest residual is 50. To reduce it, we give an increment $\delta x = 5$ and the resulting residuals are shown in (*ii*). Of these $R_x = 29$ is the largest and we give an increment $\delta z = 4$ to get the results in (*iii*). In (*vi*) $R_y = -4$ is the (numerically) largest and we give an increment $\delta y = -4/5 = -0.8$ to obtain the results in (*vii*). Similarly the other steps have been carried out.]

**Example 3.33.** *Solve the equations :*

$10x - 2y - 3z = 205$ ; $-2x + 10y - 2z = 154$ ; $-2x - y + 10z = 120$ *(Rohtak, B. Tech., 2005)*

*by Relaxation method.*

**Sol.** The residuals are given by

$$R_x = 205 - 10x + 2y + 3z \,;$$
$$R_y = 154 + 2x - 10y + 2z \,;$$
$$R_z = 120 + 2x + y - 10z.$$

The operations table is

| | $\delta R_x$ | $\delta R_y$ | $\delta R_z$ |
|---|---|---|---|
| $\delta x = 1$ | $-10$ | 2 | 2 |
| $\delta y = 1$ | 2 | $-10$ | 1 |
| $\delta z = 1$ | 3 | 2 | $-10$ |

The relaxation table is

| | $R_x$ | $R_y$ | $R_z$ |
|---|---|---|---|
| $x = y = z = 0$ | 205 | 154 | 120 |
| $\delta x = 20$ | 5 | 194 | 160 |
| $\delta y = 19$ | 43 | 4 | 179 |
| $\delta z = 18$ | 97 | 40 | $-1$ |
| $\delta x = 10$ | $-3$ | 60 | 19 |
| $\delta y = 6$ | 9 | 0 | 25 |
| $\delta z = 2$ | 15 | 4 | 5 |
| $\delta x = 2$ | $-5$ | 8 | 9 |
| $\delta z = 1$ | $-2$ | 10 | $-1$ |
| $\delta y = 1$ | 0 | 0 | 0 |

$$\Sigma\delta x = 32,\ \Sigma\delta y = 26,\ \Sigma\delta z = 21.$$

Hence $\qquad x = 32, y = 26, z = 21.$

## PROBLEMS 3.4

1. Solve by Jacobi's method, the equations : $5x - y + z = 10$ ; $2x + 4y = 12$ ; $x + y + 5z = -1$ starting with the solution (2, 3, 0).

Solve the following equations by Gauss-Jacobi method :

2. $15x_1 + x_2 - x_3 = 14$ ; $x_1 + 20x_2 + x_3 = 23$ ; $2x_1 - 3x_2 + 18x_3 = 37$.
3. $13x + 5y - 3z + u = 18$ ; $2x + 12y + z - 4u = 13$ ; $x - 4y + 10z + u = 29$ ; $2x + y - 3z + 9u = 31$.

Solve the following equations by Gauss-Seidal method :

4. $2x + y + 6z = 9$ ; $8x + 3y + 2z = 13$ ; $x + 5y + z = 7$.
5. $10x + y + z = 12$ ; $2x + 10y + z = 13$ ; $2x + 2y + 10z = 14$. (V.T.U., MCA, 200
6. $10x + 2y + z = 9$ ; $2x + 20y - 2z = -44$ ; $-2x + 3y + 10z = 22$. (Manipal, B.E., 200
7. $7x_1 + 52x_2 + 13x_3 = 104$ ; $83x_1 + 11x_2 - 4x_3 = 95$ ; $3x_1 + 8x_2 + 29x_3 = 71$. (Hazaribagh, B.E., 200
8. $3x_1 - 0.1x_2 - 0.2x_3 = 7.85$ ; $0.1x_1 + 7x_2 - 0.3x_3 = -19.3$ ; $0.3x_1 - 0.2x_2 + 10x_3 = 71.4$. (Mumbai, B.E., 200

Solve, by Relaxation method, the following equations :

9. $3x + 9y - 2z = 11$ ; $4x + 2y + 13z = 24$ ; $4x - 4y + 3z = -8$. (Bhopal, B.E., 200
10. $10x - 2y - 2z = 6$ ; $-x + 10y - 2z = 7$ ; $-x - y + 10z = 8$. (P.T.U., B. Tech., 200
11. $-9x + 3y + 4z + 100 = 0$ ; $x - 7y + 3z + 80 = 0$ ; $2x + 3y - 5z + 60 = 0$.

## 3.6. (1) ILL-CONDITIONED EQUATIONS

A linear system is said to be *ill-conditioned* if small changes in the coefficients of the equations result in large changes in the values of the unknowns. On the contrary, a system is *well-conditioned* if small changes in the coefficients of the system also produce small changes in the solution. We often come across ill-conditioned systems in practical applications. Ill-conditioning of a system is usually expected when the determinant of the coefficient matrix is small. The coefficient matrix of an ill-conditioned system is called an *ill-conditioned* matrix.



While solving simultaneous equations, we also come across two forms of *instabilities* : *Inherent* and *Induced*. Inherent instability of a system is a property of the given problem and occurs due to the problem being ill-conditioned. It can be avoided by reformulation of the problem suitably. Induced instability occurs because of the incorrect choice of method.

**(2) Iterative method to improve accuracy of an ill-conditioned system.** Consider the system of equations

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \quad ...(1)$$

Let $x', y', z'$ be an approximate solution. Substituting these values on the left-hand sides, we get new values of $d_1, d_2, d_3$ as $d_1', d_2', d_3'$ so that the new system is

$$\left.\begin{aligned} a_1x' + b_1y' + c_1z' &= d_1' \\ a_2x' + b_2y' + c_2z' &= d_2' \\ a_3x' + b_3y' + c_3z' &= d_3' \end{aligned}\right\} \quad ...(2)$$

Subtracting each equation in (2) from the corresponding equations in (1), we obtain

$$\left.\begin{aligned} a_1x_e + b_1y_e + c_1z_e &= k_1 \\ a_2x_e + b_2y_e + c_2z_e &= k_2 \\ a_3x_e + b_3y_e + c_3z_e &= k_3 \end{aligned}\right\} \quad ...(3)$$

where $x_e = x - x'$, $y_e = y - y'$, $z_e = z - z'$ and $k_i = d_i - d_i'$

We now solve the system (3) for $x_e, y_e, z_e$ giving $x = x' + x_e$, $y = y' + y_e$ and $z = z' + z_e$ which will be better approximations for $x, y, z$. We can repeat the procedure for improving the accuracy.

**Example 3.34.** *Establish whether the system $1.01x + 2y = 2.01$ ; $x + 2y = 2$ is well-conditioned or not ?*

**Sol.** Its solution is $x = 1$ and $y = 0.5$.

Now consider the system $x + 2.01y = 2.04$ and $x + 2y = 2$

which has the solution $x = -6$ and $y = 4$.

Hence the system is ill-conditioned.

**Example 3.35.** *An approximate solution of the system $2x + 2y - z = 6$ ; $x + y + 2z = 8$ ; $-x + 3y + 2z = 4$ is given by $x = 2.8$, $y = 1$, $z = 1.8$. Using the above iterative method, improve this solution.*

**Sol.** Substituting the approximate values $x' = 2.8$, $y' = 1$, $z' = 1.8$ in the given equations, we get

$$\left.\begin{aligned} 2(2.8) + 2(1) - 1.8 &= 5.8 \\ 2.8 + 1 + 2(1.8) &= 7.4 \\ -2.8 + 3(1) + 2(1.8) &= 3.8 \end{aligned}\right\} \quad ...(i)$$

Subtracting each equation in (*i*) from the corresponding given equations, we obtain

$$\left.\begin{aligned} 2x_e + 2y_e - z_e &= 0.2 \\ x_e + y_e + 2z_e &= 0.6 \\ -x_e + 3y_e + 2z_e &= 0.2 \end{aligned}\right\} \quad ...(ii)$$

where $x_e = x - 2.8$, $y_e = y - 1$, $z_e = z - 1.8$.

Solving the equations (*ii*), we get $x_e = 0.2$, $y_e = 0$, $z_e = 0.2$.

This gives the better solution $x = 3$, $y = 1$, $z = 2$, which incidently is the exact solution.

## PROBLEMS 3.5

1. Establish whether the system of equations

$$10x + 8y + 9z + 6w = 33,$$
$$6x + 7y + 5z + 5w = 23,$$
$$8x + 10y + 7z + 7w = 32,$$
$$9x + 7y + 10z + 5w = 31$$

is well-conditioned or not ?

2. An approximate solution of the equations $x + 4y + 7z = 5$ ; $2x + 5y + 8z = 7$ ; $3x + 6y + 9z = 9.1$ is given by $x = 1.8$, $y = -1.2$, $z = 1$. Improve this solution by using the iterative method.

## 3.7. COMPARISON OF VARIOUS METHODS

Direct and iterative methods have their advantages and disadvantages and a choice of method depends on a particular system of equations. The direct methods yield a solution in a finite number of steps for any non-singular set of equations, while in an iterative method the amount of computation depends on the accuracy desired. In general, it is preferable to use a direct method for the solution of a linear system. However for large systems, an iterative method yields the solution faster and should therefore, be preferred.

Gauss elimination method requires more of recording and is quite time consuming for operations. As such it is more expensive from the programming point of view. Among the direct methods, Crout's triangularisation method is used more often for the solution of a linear system and as a software for computers.

The rounding off errors also get propagated in the elimination method whereas in the iteration techniques only the rounding off errors committed in the final iteration have any effect. In general, the iteration methods have smaller round-off errors for iteration is a self correcting technique. Thus the use of an iterative method for ill-conditioned system is preferable.

On the other hand, an iterative method may not always converge. When it converges the iterative method is definitely better than the direct methods.

We come across two types of instabilities while solving a linear system of equations i.e. Inherent instability and Induced instability.

Inherent instability occurs due to the set of equations being ill-conditioned and as such is a property of the problem itself. It can, however, be avoided by a suitable reformulation of the problem.

On the other hand, induced instability occurs due to an incorrect choice of the method of solution.

## 3.8. SOLUTION OF NON-LINEAR SIMULTANEOUS EQUATIONS

**Newton-Raphson method.** Consider the equations

$$f(x, y) = 0, \; g(x, y) = 0 \qquad ...(1)$$

If an initial approximation $(x_0, y_0)$ to a solution has been found by graphical method or otherwise, then a better approximation $(x_1, y_1)$ can be obtained as follows :



Let $x_1 = x_0 + h$, $y_1 = y_0 + k$, so that

$$f(x_0 + h, y_0 + k) = 0, \; g(x_0 + h, y_0 + k) = 0 \qquad ...(2)$$

Expanding each of the functions in (2) by Taylor's series to first degree terms, we get approximately

$$\left.\begin{aligned} f_0 + h\frac{\partial f}{\partial x_0} + k\frac{\partial f}{\partial y_0} &= 0 \\ g_0 + h\frac{\partial g}{\partial x_0} + k\frac{\partial g}{\partial y_0} &= 0 \end{aligned}\right\} \qquad ...(3)$$

where $f_0 = f(x_0, y_0)$, $\dfrac{\partial f}{\partial x_0} = \left(\dfrac{\partial f}{\partial x}\right)_{x_0, y_0}$ etc.

Solving the equations (3) for $h$ and $k$, we get a new approximation to the root as

$$x_1 = x_0 + h, \; y_1 = y_0 + k.$$

This process is repeated till we get the values to the desired accuracy.

**Obs. 1.** This method will not converge unless the starting values of the roots chosen are close to the actual roots.

**2.** The method can be extended to 3 equations in 3 variables. But it is very cumbersome to obtain a meaningful solution unless the entire information about the equations and their physical context is available.

**Otherwise.** Whenever it is possible, one of the variables may be eliminated from the given equations giving a single polynomial equation in the other variable. Then find this variable to desired degree of accuracy by Newton-Raphson method. Sometimes the above polynomial equation is seen to have a root by trial. If so, reduce this equation to the next lower degree equation and find its other root by Newton-Raphson method. Having found this variable to required degree of accuracy, the other variable can at once, be found from one of the given equations.

**Example 3.36.** *Solve the system of non-linear equations*

$$x^2 + y = 11, \; y^2 + x = 7. \qquad \text{(Pune, B.E., 2000)} \quad ...(1)$$

**Sol.** An initial approximation to the solution is obtained from a rough graph of (1), as $x_0 = 3.5$ and $y_0 = -1.8$.

We have $f = x^2 + y - 11$ and $g = y^2 + x - 7$ so that

$$\frac{\partial f}{\partial x} = 2x, \qquad \frac{\partial f}{\partial y} = 1$$

$$\frac{\partial g}{\partial x} = 1, \qquad \frac{\partial g}{\partial y} = 2y.$$

Then Newton-Raphson's equations (3) above will be

$$7h + k = 0.55, \; h - 3.6k = 0.26.$$

Solving these, we get $h = 0.0855$, $k = -0.0485$

$\therefore$ The better approximation to the root is

$$x_1 = x_0 + h = 3.5855, \; y_1 = y_0 + k = -1.8485.$$

Repeating the above process, replacing $(x_0, y_0)$ by $(x_1, y_1)$, we obtain $x_2 = 3.5844$, $y_2 = -1.8482$.

**Otherwise.** Eliminating $y$ from the given equations, we get

$$x^4 - 22x^2 + x + 114 = 0$$

By trial, $x = 3$ is its root.

$\therefore$ The reduced equation is $x^3 + 3x^2 - 13x - 38 = 0$

To find the other root, we apply Newton-Raphson method to

$$f(x) = x^3 + 3x^2 - 13x - 38.$$

Taking $x_0 = 3.5$, we get $\quad x_2 = 3.5844.$

Thus $\quad y = 11 - x^2$ gives $\quad y = -1.848$ for $x = 3.5844$

Also $\quad y = -2$ for $x = 3.$

**Example 3.37.** *Solve the equations $2x^2 + 3xy + y^2 = 3$, $4x^2 + 2xy + y^2 = 30$ correct to three decimal places, using Newton-Raphson method, given that $x_0 = -3$ and $y_0 = 2$.*

**Sol.** We have $f = 2x^2 + 3xy + y^2 - 3$ and $g = 4x^2 + 2xy + y^2 - 30$, so that

So that
$$\frac{\partial f}{\partial x} = 4x + 3y, \quad \frac{\partial f}{\partial y} = 3x + 2y$$

$$\frac{\partial g}{\partial x} = 8x + 2y, \quad \frac{\partial g}{\partial y} = 2x + 2y$$

Now
$$f_0 = 2x_0^2 + 3x_0y_0 + y_0^2 - 3 = 1$$
$$g_0 = 4x_0^2 + 2x_0y_0 + y_0^2 - 30 = -2$$
$$\frac{\partial f}{\partial x_0} = -6, \quad \frac{\partial f}{\partial y_0} = -5\,; \quad \frac{\partial g}{\partial x_0} = -20, \quad \frac{\partial g}{\partial y_0} = -2$$

Then Newton-Raphson equations (3) above will be

$$20h + 2k = -2;\ 6h + 5k = 1$$

Solving these equations, we get $h = -\dfrac{3}{22} = -0.1364$, $k = \dfrac{4}{11} = 0.3636$

$\therefore$ The better approximation is

$$x_1 = x_0 + h = -3 - 0.1364 = -3.1364$$
$$y_1 = y_0 + k = 2 + 0.3636 = 2.3636$$

Repeating the above process and replacing $(x_0, y_0)$ by $(x_1, y_1)$, we obtain $x_2 = -3.131$, $y_2 = 2.362$

Again proceeding as above and replacing $(x_1, y_1)$ by $(x_2, y_2)$, we obtain $x_3 = -3.1309$, $y_3 = 2.3617$

Since the values $x_2, y_2$ and $x_3, y_3$ are approximately equal, the solution correct to three decimal places, is $x = -3.131$, $y = 2.362$.

## PROBLEMS 3.6

1. Solve the equations $x^2 + y = 5$, $y^2 + x = 3$.
2. Solve the non-linear equations $x = 2(y + 1)$, $y^2 = 3xy - 7$ correct to three decimals.
3. Find a root of the equations $xy = x + 9$, $y^2 = x^2 + 7$.
4. Use Newton-Raphson method to solve the equations $x = x^2 + y^2$, $y = x^2 - y^2$ correct to two decimals, starting with the approximation (0.8, 0.4).
5. Solve the non-linear equations $x^2 - y^2 = 4$, $x^2 + y^2 = 16$ numerically with $x_0 = y_0 = 2.828$ using Newton-Raphson method. Carry out two iterations. *(V.T.U., MCA, 2007)*



## 3.9 OBJECTIVE TYPE OF QUESTIONS

## PROBLEMS 3.7

*Select the correct answer or fill up the blanks in the following questions :*

1. As soon as a new value of a variable is found by iteration, it is used immediately in the following equations, this method is called
   (*a*) Gauss-Jordan method (*b*) Gauss-Seidal method
   (*c*) Jacobi's method (*d*) Relaxation method.
2. The difference between direct and iterative method of solving simultaneous linear equations is ............ .
3. In solving simultaneous equations by Gauss-Jordan method, the coefficient matrix is reduced to ............ matrix.
4. The condition for the convergence of Gauss-Seidal matrix is that in each equation of the system ............ .
5. A matrix in which $a_{ij} = 0$ for $i \neq j$ is called ............ .
6. Solutions of simultaneous non-linear equations can be obtained using
   (*a*) Method of iteration (*b*) Newton-Raphson method
   (*c*) None of the above.
7. To which form the coefficient matrix is transformed when $AX = B$ is solved by Gauss-elimination method.
8. Guass-Seidal iteration converges only if the coefficient matrix is diagonally dominant. (True or False)
9. What is 'partial pivoting' and 'complete pivoting' in the solution of linear simultaneous equations.
10. The convergence in Gauss-Seidal method is ...... than that in Jacobi's method :
    (*a*) more fast (*b*) more slow
    (*c*) slow (*d*) equal. *(Bhopal, B.E. 2007)*
11. By Gauss elimination method, solve $x + y = 2$ and $2x + 3y = 5$. *(Anna, B.Tech., 2007)*

# 4

# MATRIX INVERSION & EIGEN VALUE PROBLEM



## 4.1. INTRODUCTION

There are two main numerical problems which arise in connection with the matrices. One of these is the problem of finding the inverse of a matrix. The other problem is that of finding the eigen-values and the corresponding eigen-vectors of a matrix. When a student first encounters an Eigen-value problem, it appears to him somewhat artificial and theoretical only. In fact the computation of Eigen-values is required in many engineering and scientific problems. For instance, the frequencies of the vibrations of beams are the Eigen values of a matrix. Eigen-values are also required while finding the frequencies associated with

(*i*) the vibrations of a system of masses and springs,

(*ii*) the symmetric vibrations of an annular membrane,

(*iii*) the oscillations of a triple pendulum,

(*iv*) the torsional oscillations of a uniform cantilever,

(*v*) the torsional oscillations of a multi-cylinder engine etc.

Once the physical formulation in any of the above situations is completed, all these problems have the same mathematical approach : that of finding an eigen-value for a numerical matrix.

## 4.2. MATRIX INVERSION

In § 3.2(4), we have already defined the inverse of a non-singular square matrix $A$, to be another matrix $B$ of the same order such that $AB = BA = I$, $I$ being a unit matrix of the same order.

The inverse of a matrix $A$ is written as $A^{-1}$ so that $AA^{-1} = A^{-1}A = I$







Thus the inverse of a matrix exists if and only if it is a non-singular square matrix. Also inverse of a matrix, when it exists in unique.

There are several methods of finding the inverse of a matrix. Of these, the method of obtaining the inverse with the help of an adjoint has already been illustrated by Example 3.9. But it requires a lot of calculations. As such, we shall now, describe some other methods which require less of computational labour and can be easily extended to matrices of higher order.

## 4.3. GAUSS ELIMINATION METHOD

The method involves the same procedure as explained in §3.4(3). Here we take a unit matrix of the same order as the given matrix $A$ and write it as $AI$.

Now making simultaneous row operations on $AI$, we try to convert $A$ into an upper triangular matrix and then to a unit matrix. Ultimately when $A$ is transformed into a unit matrix, the adjacent matrix (emerged out from the transformation of $I$) gives the inverse of $A$. To increase the accuracy, the largest element in $A$ is taken as the pivot element for performing the row operations.

## 4.4. GAUSS-JORDAN METHOD

This is similar to the Gauss elimination method except that instead of first converting $A$ into upper triangular form, it is directly converted into the unit matrix.

*In practice, the two matrices A and I are written side by side and the same transformations are performed on both. As soon as A is reduced to I, the other matrix represents* $A^{-1}$.

**Example 4.1.** *Using Gauss-Jordan method, find the inverse of the matrix*

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

(*Kurukshetra, B. Tech., 2006*)

**Sol.** Writing the given matrix side by side with the unit matrix of order 3, we have

$$\left[\begin{array}{ccc:ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 1 & 3 & -3 & 0 & 1 & 0 \\ -2 & -4 & -4 & 0 & 0 & 1 \end{array}\right] \quad \text{(Operate } R_2 - R_1 \text{ and } R_3 + 2R_1\text{)}$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 2 & -6 & -1 & 1 & 0 \\ 0 & -2 & 2 & 2 & 0 & 1 \end{array}\right] \quad \text{(Operate } \tfrac{1}{2}R_2 \text{ and } \tfrac{1}{2}R_3\text{)}$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -1/2 & 1/2 & 0 \\ 0 & -1 & 1 & 1 & 0 & 1/2 \end{array}\right] \quad \text{(Operate } R_1 - R_2 \text{ and } R_3 + R_2\text{)}$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & 6 & 3/2 & -1/2 & 0 \\ 0 & 1 & -3 & -1/2 & 1/2 & 0 \\ 0 & 0 & -2 & 1/2 & 1/2 & 1/2 \end{array}\right]$$

[Operate $R_1 + 3R_3$, $R_2 - \frac{3}{2}R_3$ and $(\frac{1}{2})R_3$]

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 3 & 1 & 3/2 \\ 0 & 1 & 0 & -5/4 & -1/4 & -3/4 \\ 0 & 0 & 1 & -1/4 & -1/4 & -1/4 \end{array}\right]$$

Hence the inverse of the given matrix is $\begin{bmatrix} 3 & 1 & 3/2 \\ -5/4 & -1/4 & -3/4 \\ -1/4 & -1/4 & -1/4 \end{bmatrix}$

■ **Example 4.2.** *Using Gauss-Jordan method, find the inverse of the matrix* $\begin{bmatrix} 2 & 2 & 3 \\ 2 & 1 & 1 \\ 1 & 3 & 5 \end{bmatrix}$.

*(Anna, B. Tech., 2004)*

**Sol.** Writing the given matrix side by side with the unit matrix of order 3, we have

$$\left[\begin{array}{ccc:ccc} 2 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 3 & 5 & 0 & 0 & 1 \end{array}\right] \qquad \left(\text{Operate } \tfrac{1}{2} R_1\right)$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 1 & 3/2 & 1/2 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 3 & 5 & 0 & 0 & 1 \end{array}\right] \qquad (\text{Operate } R_2 - 2R_1, R_3 - R_1)$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 1 & 3/2 & 1/2 & 0 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 0 & 2 & 7/2 & -1/2 & 0 & 1 \end{array}\right] \qquad (\text{Operate } R_1 + R_2, R_3 + 2R_2)$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & -1/2 & -1/2 & 1 & 0 \\ 0 & -1 & -2 & -1 & 1 & 0 \\ 0 & 0 & -1/2 & -5/2 & 2 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & -1/2 & -1/2 & 1 & 0 \\ 0 & 1 & 2 & 1 & -1 & 0 \\ 0 & 0 & -1/2 & -5/2 & 2 & 1 \end{array}\right] \qquad (\text{Operate } (-2) R_3)$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & -1/2 & -1/2 & 1 & 0 \\ 0 & 1 & 2 & 1 & -1 & 0 \\ 0 & 0 & 1 & 5 & -4 & -2 \end{array}\right] \qquad \left(\text{Operate } R_1 + \tfrac{1}{2} R_3, R_2 - 2R_3\right)$$

$$\sim \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & -9 & 7 & 4 \\ 0 & 0 & 1 & 5 & -4 & -2 \end{array}\right]$$

Hence the inverse of the given matrix is $\begin{bmatrix} 2 & -1 & -1 \\ -9 & 7 & 4 \\ 5 & -4 & -2 \end{bmatrix}$.

## 4.5. FACTORIZATION METHOD

In this method, we factorize the given matrix as $A = LU$ ...(1)

where $L$ is a lower triangular matrix with unit diagonal elements and $U$ is an upper triangular matrix *i.e.*



$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$ [§ 3.4(5)]

Now (1) gives $A^{-1} = (LU)^{-1} = U^{-1}L^{-1}$ ...(2)

*To find* $L^{-1}$, let $L^{-1} = X$, where $X$ is a lower triangular matrix.

Then $LX = I$

*i.e.*
$$\begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} x_{11} & 0 & 0 \\ x_{21} & x_{22} & 0 \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying the matrices on the L.H.S. and equating the corresponding elements, we have

$$x_{11} = 1, x_{22} = 1, x_{33} = 1 \qquad \text{...(3)}$$

$$l_{21}x_{11} + x_{21} = 0, l_{31}x_{11} + l_{32}x_{21} + x_{31} = 0$$

and
$$l_{32}x_{22} + x_{32} = 0 \qquad \text{...(4)}$$

(3) gives $x_{11} = x_{22} = x_{33} = 1$

(4) $x_{21} = -l_{21}x_{11}, x_{31} = -(l_{31} + l_{32}x_{21})$ and $x_{32} = -l_{32}$

Thus $L^{-1} = X$ is completely determined.

*To find* $U^{-1}$, let $U^{-1} = Y$, where $Y$ is an upper triangular matrix.

Then $YU = I$

*i.e.*
$$\begin{bmatrix} y_{11} & y_{12} & y_{13} \\ 0 & y_{22} & y_{23} \\ 0 & 0 & y_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplying the matrices on the L.H.S. and then equating the corresponding elements, we have ...(5)

$$y_{11}u_{11} = 1, y_{22}u_{22} = 1, y_{33}u_{33} = 1$$

$$\left.\begin{array}{r} y_{11}u_{12} + y_{12}u_{22} = 0, y_{11}u_{13} + y_{12}u_{23} + y_{13}u_{33} = 0 \\ y_{22}u_{23} + y_{23}u_{33} = 0 \end{array}\right\} \qquad \text{...(6)}$$

and

From (5), $y_{11} = 1/u_{11}, y_{22} = 1/u_{22}, y_{33} = 1/u_{33}$

From (6), $y_{12} = -y_{11}u_{12}/u_{22}, y_{13} = -(y_{11}u_{13} + y_{12}u_{23})/u_{33}; y_{23} = -y_{22}u_{23}/u_{33}$.

$\therefore$ We get $U^{-1} = Y$, completely.

Hence, by (2), we obtain $A^{-1}$.

■ **Example 4.3.** *Using the factorization method, find the inverse of the matrix*

$$A = \begin{bmatrix} 50 & 107 & 36 \\ 25 & 54 & 20 \\ 31 & 66 & 21 \end{bmatrix}$$

**Sol.** Taking $L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$ and $U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$,

$$A = LU, \quad \begin{bmatrix} 50 & 107 & 36 \\ 25 & 54 & 20 \\ 31 & 66 & 21 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

$$\therefore \quad \begin{aligned} &50 = u_{11},\ 107 = u_{12},\ 36 = u_{13};\\ &25 = l_{21}u_{11},\ 54 = l_{21}u_{12} + u_{22},\ 20 = l_{21}u_{13} + u_{23};\\ &31 = l_{31}u_{11},\ 66 = l_{21}u_{12} + l_{32}u_{22},\ 21 = l_{31}u_{13} + l_{32}u_{23} + u_{33}\end{aligned}$$

or $$u_{11} = 50,\ u_{12} = 107,\ u_{13} = 36,\ l_{21} = 1/2,\ u_{22} = 1/2,\ u_{23} = 2,$$
$$l_{31} = 31/50,\ l_{32} = -17/25,\ u_{33} = 1/25.$$

Thus $$L = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 31/50 & -17/25 & 1 \end{bmatrix} \text{ and } U = \begin{bmatrix} 50 & 107 & 36 \\ 0 & 1/2 & 2 \\ 0 & 0 & 1/25 \end{bmatrix}$$

*To find* $L^{-1}$, let $L^{-1} = X$. Then $LX = I$

i.e. $$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 31/50 & -17/25 & 1 \end{bmatrix}\begin{bmatrix} x_{11} & 0 & 0 \\ x_{21} & x_{22} & 0 \\ x_{31} & x_{32} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\therefore\ x_{11} = 1,\ \tfrac{1}{2}x_{11} + x_{21} = 0,\ x_{22} = 1,\ \frac{31}{50}x_{11} - \frac{17}{25}x_{21} + x_{31} = 0,\ -\frac{17}{25}x_{22} + x_{32} = 0,\ x_{33} = 1$$

or $$x_{11} = x_{22} = x_{33} = 1,\ x_{21} = -\frac{1}{2},\ x_{31} = -\frac{24}{25},\ x_{32} = \frac{17}{25}.$$

Thus $$L^{-1} = X = \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -24/25 & 17/25 & 1 \end{bmatrix}$$

*To find* $U^{-1}$, let $U^{-1} = Y$. Then $YU = I$

i.e. $$\begin{bmatrix} y_{11} & y_{12} & y_{13} \\ 0 & y_{22} & y_{23} \\ 0 & 0 & y_{33} \end{bmatrix}\begin{bmatrix} 50 & 107 & 36 \\ 0 & 1/2 & 2 \\ 0 & 0 & 1/25 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\therefore\quad 50y_{11} = 1,\ 50y_{12} + 107y_{22} = 0,\ 50y_{13} + 107y_{23} + 36y_{33} = 0$$
$$\frac{1}{2}y_{22} = 1,\ \frac{1}{2}y_{23} + 2y_{33} = 0,\ \frac{1}{25}y_{33} = 1.$$

or $$y_{11} = 1/50,\ y_{22} = 2,\ y_{33} = 25,\ y_{12} = -107/25,\ y_{23} = -100,\ y_{13} = 196.$$

So that $$U^{-1} = \begin{bmatrix} 1/50 & -107/25 & 196 \\ 0 & 2 & -100 \\ 0 & 0 & 25 \end{bmatrix}$$

Hence, $$A^{-1} = U^{-1}L^{-1} = \begin{bmatrix} 1/50 & -107/25 & 196 \\ 0 & 2 & -100 \\ 0 & 0 & 25 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ -24/25 & 17/25 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} -186 & 129 & 196 \\ 95 & -66 & -100 \\ -24 & 17 & 25 \end{bmatrix}$$

## 4.6 PARTITION METHOD

According to this method, if the inverse of a matrix $A_n$ of order $n$ is known, then the inverse of a matrix $A_{n+1}$ of order $(n+1)$ can be determined by adding $(n+1)$th row and $(n+1)$th column to $A_n$.



Suppose $$A = \begin{bmatrix} A_1 & : & A_2 \\ \hline A_3' & : & \alpha \end{bmatrix} \text{ and } A^{-1} = \begin{bmatrix} X_1 & : & X_2 \\ \hline X_3' & : & x \end{bmatrix}$$

where $A_2$, $X_2$ are column vectors and $A_3'$, $X_3'$ are row vectors (*i.e.* transposes of column vectors $A_3$, $X_3$) and $\alpha$, $x$ are ordinary numbers.

Also we assume that $A_1^{-1}$ is known. Actually $A_3$ and $X_3$ are column vectors since their transposes are row vectors.

Now $AA^{-1} = I_{n+1}$ gives

$$A_1X_1 + A_1X_3' = I_n \quad ...(1)$$
$$A_1X_2 + A_2x = 0 \quad ...(2)$$
$$A_3'X_1 + \alpha X_3' = 0 \quad ...(3)$$
$$A_3'X_2' + \alpha x = 1 \quad ...(4)$$

From (2), $X_2 = -A_1^{-1}A_2x$ ...(5)

and using this, (4) gives $(\alpha - A_3'A_1^{-1}A_2)x = 1$. ...(6)

Hence $x$ and then $X_2$ can be found.

Also from (1), $X_1 = A_1^{-1}(I_n - A_2X_3')$ ...(7)

and using this, (3) gives $(\alpha - A_3'A_1^{-1}A_2)X_3' = -A_3'A_1^{-1}$ ...(8)

whence $X_3'$ and then $X_1$ are determined.

Thus, having found $X_1$, $X_2$, $X_3'$ and $x$, $A^{-1}$ is completely known.

**Obs.** The partition method is also known as the '*Escalator method*'.

**Example 4.4.** *Using the partition method, find the inverse of* $A = \begin{bmatrix} 13 & 14 & 6 & 4 \\ 8 & -1 & 13 & 9 \\ 6 & 7 & 3 & 2 \\ 9 & 5 & 16 & 11 \end{bmatrix}$.

**Sol.** We have $$A = \left[\begin{array}{ccc:c} 13 & 14 & 6 & 4 \\ 8 & -1 & 13 & 9 \\ 6 & 7 & 3 & 2 \\ 9 & 5 & 16 & 11 \end{array}\right] = \begin{bmatrix} A_1 & : & A_2 \\ A_3' & : & \alpha \end{bmatrix}$$

so that $$A_1 = \begin{bmatrix} 13 & 14 & 6 \\ 8 & -1 & 13 \\ 6 & 7 & 3 \end{bmatrix},\quad A_2 = \begin{bmatrix} 4 \\ 9 \\ 2 \end{bmatrix}$$

$$A_3' = [9 \quad 5 \quad 16] \text{ and } \alpha = 11.$$

We find $$A_1^{-1} = \frac{1}{94}\begin{bmatrix} 94 & 0 & -188 \\ -54 & -3 & 121 \\ -62 & 7 & 125 \end{bmatrix}$$

Let $$A^{-1} = \begin{bmatrix} X_1 & : & X_2 \\ \hline X_3' & : & \alpha \end{bmatrix}. \text{ Then } AA^{-1} = I.$$

Hence $A_3' A_1^{-1} A_2 = [9 \quad 5 \quad 16] \dfrac{1}{94}\begin{bmatrix} 94 & 0 & -188 \\ -54 & -3 & 121 \\ -62 & 7 & 125 \end{bmatrix}\begin{bmatrix} 4 \\ 9 \\ 2 \end{bmatrix} = \dfrac{1}{94}[9 \quad 5 \quad 16]\begin{bmatrix} 0 \\ -1 \\ 65 \end{bmatrix} = \dfrac{1035}{94}$

$\therefore \quad (\alpha - A_3' A_1^{-1} A_2)x = 1$ [(6) of § 4.6]

becomes, $\left(11 - \dfrac{1035}{95}\right)x = 1$ *i.e.* $x = -94$

Also $X_2 = -A_1^{-1} A_2 x = -\dfrac{1}{94}\begin{bmatrix} 94 & 0 & -188 \\ -54 & -3 & 121 \\ -62 & 7 & 125 \end{bmatrix}\begin{bmatrix} 4 \\ 9 \\ 2 \end{bmatrix}(-94) = \begin{bmatrix} 0 \\ -1 \\ 65 \end{bmatrix}$ [(5) of § 4.7]

Then $(\alpha - A_3' A_1^{-1} A_2) X_3' = -A_3' A_1^{-1}$ [(8) of § 4.6]

becomes $\left(11 - \dfrac{1035}{94}\right) X_3' = -\dfrac{1}{94}[-416, 97, 913]$ whence $X_3' = [-416, 97, 913]$

Finally $X_1 = A_1^{-1}(I - A_2 X_3')$ [(7) of § 4.6]

where $A_2 X_3' = \begin{bmatrix} 4 \\ 9 \\ 2 \end{bmatrix}[-416, 97, 913] = \begin{bmatrix} -1664 & 388 & 3652 \\ -3744 & 873 & 8217 \\ -832 & 194 & 1826 \end{bmatrix}$

$$\therefore \quad X_1 = \frac{1}{94}\begin{bmatrix} 94 & 0 & -188 \\ -54 & -3 & 121 \\ -62 & 7 & 125 \end{bmatrix}\begin{bmatrix} 1665 & -388 & -3652 \\ 3744 & -872 & -8217 \\ 832 & -194 & -1825 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 \\ -5 & 1 & 11 \\ 287 & -67 & -630 \end{bmatrix}$$

$$\text{Hence } A^{-1} = \begin{bmatrix} X_1 & X_2 \\ X_3' & x \end{bmatrix} = \begin{bmatrix} 1 & 0 & -2 & 0 \\ -5 & 1 & 11 & -1 \\ 287 & -67 & -630 & 65 \\ -416 & 97 & 913 & -94 \end{bmatrix}$$

**Example 4.5.** *If A and C are non-singular matrices, then show that*

$$\begin{bmatrix} A & 0 \\ B & C \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ -C^{-1}BA^{-1} & C^{-1} \end{bmatrix}$$

*Hence find inverse of* $\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 4 & 0 \\ \hline 0 & 1 & 0 & 3 \end{array}\right]$. *(Mumbai, B.Tech., 2005)*

**Sol.** Let the given matrix be $M = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$ and its inverse be $M^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$ both in the partioned form where $A, B, C, P, Q, R, S$ are all matrices.

$$\therefore \quad MM^{-1} = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix}\begin{bmatrix} P & Q \\ R & S \end{bmatrix} = I$$

or $$\begin{bmatrix} AP + 0R & AQ + 0S \\ BP + CR & BQ + CS \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$



$\therefore$ Equating corresponding elements, we have

$$AP + 0R = I,\ AQ + 0S = 0,\ BP + CR = 0,\ BQ + CS = I.$$

Second relation gives $AQ = 0$ *i.e.* $Q = 0$ as $A$ is non-singular.

First relation gives $AP = I$ *i.e.* $P = A^{-1}$.

First third equation, $BP + CR = 0$ *i.e.* $CR = -BP = -BA^{-1}$

$\therefore \quad C^{-1}CR = -C^{-1}BA^{-1}$ or $IR = -C^{-1}BA^{-1}$ or $R = -C^{-1}BA^{-1}$

From fourth equation, $BQ + CS = I$, or $CS = I$ or $S = C^{-1}$

Hence $$M^{-1} = \begin{bmatrix} A^{-1} & 0 \\ -C^{-1}BA^{-1} & C^{-1} \end{bmatrix}$$

## 4.7. ITERATIVE METHOD

Suppose we wish to compute $A^{-1}$ and we know that $B$ is an approximate inverse of $A$. Then the error matrix is given by $E = AB - I$

or $$AB = I + E$$

$\therefore$ $$(AB)^{-1} = (I + E)^{-1} \text{ i.e. } B^{-1}A^{-1} = (I + E)^{-1}$$

or $$A^{-1} = B(I + E)^{-1} = B(I - E + E^2 - \ldots\ldots),$$

provided the series converges.

Thus we can find further approximations of $A^{-1}$, by using $A^{-1} = B(1 - E + E^2 - \ldots)$

**Example 4.6.** *Using the iterative method, find the inverse of*

$$A = \begin{bmatrix} 1 & 10 & 1 \\ 2 & 0 & 1 \\ 3 & 3 & 2 \end{bmatrix} \text{ taking } B = \begin{bmatrix} 0.4 & 2.4 & -1.4 \\ 0.14 & 0.14 & -0.14 \\ -0.85 & -3.8 & 2.8 \end{bmatrix}$$

**Sol.** Here $$E = AB - I = \begin{bmatrix} 1 & 10 & 1 \\ 2 & 0 & 1 \\ 3 & 3 & 2 \end{bmatrix}\begin{bmatrix} 0.4 & 2.4 & -1.4 \\ 0.14 & 0.14 & -0.14 \\ -0.85 & -3.8 & 2.8 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -0.05 & 0 & 0 \\ -0.05 & 0 & 0 \\ -0.08 & 0.02 & -0.02 \end{bmatrix}$$

$\therefore$ $$E^2 = \begin{bmatrix} 0.0025 & 0 & 0 \\ 0.0025 & 0 & 0 \\ 0.0064 & -0.0004 & -0.0004 \end{bmatrix}$$

To the second approximation, we have

$$A^{-1} = B(1 - E + E^2) = B - BE + BE^2$$

$$= \begin{bmatrix} 0.4 & 2.4 & -1.4 \\ 0.0025 & 0.14 & -0.14 \\ -0.85 & -3.8 & -2.8 \end{bmatrix} - \begin{bmatrix} -0.02 & -0.12 & 0.07 \\ -0.02 & -0.12 & 0.07 \\ -0.0122 & -0.1132 & 0.0532 \end{bmatrix}$$

$$+ \begin{bmatrix} 0.001 & 0.006 & -0.0035 \\ 0.001 & 0.006 & -0.0035 \\ 0.0014 & 0.0095 & -0.0053 \end{bmatrix} = \begin{bmatrix} 0.421 & 2.526 & -1.474 \\ 0.161 & 0.266 & -0.214 \\ -0.836 & -3.677 & 2.742 \end{bmatrix}.$$

## PROBLEMS 4.1

Use Gauss-Jordan method to find the inverse of the following matrices :

1. $\begin{bmatrix} 2 & 0 & 1 \\ 3 & 2 & 5 \\ 1 & -1 & 0 \end{bmatrix}$ *(Anna, B. Tech., 2005)*

2. $\begin{bmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}$ *(Andhra, B. Tech., 2001)*

3. $\begin{bmatrix} 8 & -4 & 0 \\ -4 & 8 & -4 \\ 0 & -4 & 8 \end{bmatrix}$

Use factorization method, to find the inverse of the following matrices :

4. $\begin{bmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$ *(V.T.U., MCA, 2007)*

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

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

7. $\begin{bmatrix} 10 & 2 & 1 \\ 2 & 20 & -2 \\ -2 & 3 & 10 \end{bmatrix}$

Apply the partition method to obtain the inverse of the following matrices :

8. $\begin{bmatrix} 1 & 1 & 1 \\ 4 & 3 & -1 \\ 3 & 5 & 3 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 3 & 3 & 2 \\ 1 & 4 & 3 & 4 \\ 1 & 3 & 4 & 5 \\ 2 & 5 & 3 & 2 \end{bmatrix}$

10. Using iterative method, find the inverse of the matrix $A = \begin{bmatrix} 5 & 2 \\ 3 & -1 \end{bmatrix}$ taking $B = \begin{bmatrix} 0.1 & 0.2 \\ 0.3 & -0.4 \end{bmatrix}$.

11. Apply iterative method to find more accurate inverse of $A = \begin{bmatrix} 1 & 10 & 1 \\ 2 & 0 & 1 \\ 3 & 3 & 2 \end{bmatrix}$, assuming the initial inverse matrix to be $\begin{bmatrix} 0.43 & 2.43 & -1.43 \\ 0.14 & 0.14 & -0.14 \\ -0.85 & -3.85 & 0.85 \end{bmatrix}$.

## 4.8. (1) EIGEN VALUES AND EIGEN VECTORS

If $A$ is any square matrix of order $n$ with elements $a_{ij}$, we can find a column matrix $X$ and a constant $\lambda$ such that $AX = \lambda X$ or $AX - \lambda IX = 0$ or $[A - \lambda I]X = 0$

This matrix equation represents $n$ homogeneous linear equations

$$\begin{aligned} (a_{11} - \lambda)x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + \ldots\ldots + a_{2n}x_n &= 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots\ldots + (a_{nn} - \lambda)x_n &= 0 \end{aligned} \qquad \ldots(1)$$

which will have a non-trivial solution only if the coefficient determinant vanishes, *i.e.*

$$\begin{vmatrix} a_{11} - \lambda & a_{12} \ldots\ldots\ldots\ldots a_{1n} \\ a_{21} & a_{22} - \lambda \ldots\ldots a_{2n} \\ a_{n1} & a_{n2} \ldots\ldots\ldots\ldots a_{nn} - \lambda \end{vmatrix} = 0 \qquad \ldots(2)$$



On expansion, it gives an $n$th degree equation in $\lambda$, called the *characteristic equation* of the matrix $A$. Its roots $\lambda_i$ ($i = 1, 2, \ldots\ldots n$) are called the *eigen values* or *latent roots* and corresponding to each eigen value, the equation (2) will have a non-zero solution

$$X = [x_1, x_2, \ldots\ldots\ldots, x_n]'$$

which is known as the *eigen vector*. Such an equation can ordinarily be solved easily. However for larger systems better methods are to be applied.

**(2) Cayley-Hamilton theorem.** *Every square matrix satisfies its own characteristic equation* *i.e.* if the characteristic equation for the $n$th order square matrix $A$ is

$$|A - \lambda I| = (-1)^n \lambda^n + k_1 \lambda^{n-1} + \ldots\ldots + k_n = 0$$

then

$$(-1)^n A^n + k_1 A^{n-1} + k_n = 0.$$

**Example 4.7.** *Find the eigen values and eigen vectors of the matrix* $\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

**Sol.** The characterstic equation is $[A - \lambda I] = 0$

*i.e.* $$\begin{bmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{bmatrix} = 0 \quad \text{or} \quad \lambda^2 - 7\lambda + 6 = 0$$

or $$(\lambda - 6)(\lambda - 1) = 0 \quad \therefore \quad \lambda = 6, 1.$$

Thus the eigen values are 6 and 1.

If $x, y$ be the components of an eigen vector corresponding to the eigen value $\lambda$, then

$$[A - \lambda I]\, X = \begin{bmatrix} 5-\lambda & 4 \\ 1 & 2-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

Corresponding to $\lambda = 6$, we have $\begin{bmatrix} -1 & 4 \\ 1 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$

which gives only one independent equation $-x + 4y = 0$

$\therefore \ \dfrac{x}{4} = \dfrac{y}{1}$ giving the eigen vector (4, 1).

Corresponding to $\lambda = 1$, we have $\begin{bmatrix} 4 & 4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$ which gives only one independent equation $x + y = 0$.

$\therefore \ \dfrac{x}{1} = \dfrac{y}{-1}$ giving the eigen vector $(1, -1)$.

**Example 4.8.** *Find the eigen values and eigen vectors of the matrix*

$$A = \begin{bmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{bmatrix}$$ *(J.N.T.U., B. Tech., 2006)*

**Sol.** The characteristic equation is

$$|A - \lambda I| = \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} = -\lambda^3 + 18\lambda^2 - 45\lambda = 0$$

or $$\lambda(\lambda - 3)(\lambda - 15) = 0 \quad \therefore \quad \lambda = 0, 3, 15.$$

Thus the eigen values of $A$ are 0, 3, 15.

If $x, y, z$ be the components of an eigen vector corresponding to the eigen value $\lambda$, we have

$$(A - \lambda I)X = \begin{bmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \quad ...(i)$$

Putting $\lambda = 0$, we have $8x - 6y + 2z = 0$, $-6x + 7y - 4z = 0$, $2x - 4y + 3z = 0$.

These equations determine a single linearly independent solution which may be taken as (1, 2, 2) so that every non-zero multiple of this vector is an eigen vector corresponding to $\lambda = 0$ ...(ii)

Similarly, the eigen vectors corresponding to $\lambda = 3$ and $\lambda = 15$ are the arbitrary non-zero multiples of the vectors (2, 1, – 2) and (2, – 2, 1) which are obtained from (i).

Hence the three eigen vectors may be taken as (1, 2, 2), (2, 1, – 2), (2, – 2, 1).

**Obs.** The eigen vector $[x, y, z]'$ such that $x^2 + y^2 + z^2 = 1$ is said to be **normalized.** In particular, if we choose $x = 1/3$, $y = 2/3$, $z = 2/3$ in (ii), the corresponding normalized eigen vector will be (1/3, 2/3, 2/3).

**Example 4.9.** *Using Cayley-Hamilton theorem, find the inverse of the matrix*

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

(Rajasthan, B.E., 200_)

**Sol.** The characteristic equation of the matrix is

$$\begin{bmatrix} 2-\lambda & 1 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 1 & 2-\lambda \end{bmatrix} = 0 \quad \text{or} \quad \lambda^3 - 5\lambda^2 + 7\lambda - 3 = 0$$

By Cayley-Hamilton theorem, we have $A^3 - 5A^2 + 7A - 3I = 0$ ...(i)

Multiplying (i) by $A^{-1}$, we get

$$A^2 - 5A + 7I - 3A^{-1} = 0 \quad \text{or} \quad A^{-1} = \frac{1}{3}(A^2 - 5A + 7) \quad ...(ii)$$

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

$$\therefore \quad A^2 - 5A + 7I = \begin{bmatrix} 5 & 4 & 4 \\ 0 & 1 & 0 \\ 4 & 4 & 5 \end{bmatrix} - 5\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix} + 7\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 \\ 0 & 3 & 0 \\ -1 & -1 & 2 \end{bmatrix}$$

Hence from (ii), $A^{-1} = \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ 0 & 3 & 0 \\ -1 & -1 & 2 \end{bmatrix}$.



## 4.9. PROPERTIES OF EIGEN VALUES

We state below, some of the important properties of eigen values for ready reference :

*I. The sum of the eigen values of a matrix A, is the sum of the elements of its principal diagonal.*

*II. If $\lambda$ is an eigen value of a matrix A, then $1/\lambda$ is the eigen value of $A^{-1}$.*

*III. If $\lambda$ is an eigen value of an orthogonal matrix, then $1/\lambda$ is also its eigen value.*

*IV. If $\lambda_1, \lambda_2, ......, \lambda_n$ are the eigen values of a matrix A, then $A^m$ has the eigen values $\lambda_1^m, \lambda_2^m, ......, \lambda_n^m$ (m being a positive integer).*

*V. If a square matrix A has n linearly independent eigen vectors, then a matrix P can be found such that $P^{-1}AP$ is a diagonal matrix whose diagonal elements are the eigen values of A.*

The transformation of $A$ by a non-singular matrix $P$ to $P^{-1}AP$ is called a *similarity transformation.*

*VI. Any similarity transformation applied to a matrix leaves its eigen values unchanged.*

## 4.10. BOUNDS FOR EIGEN VALUES

*If $\lambda$ be an eigen value of a matrix A, then for some $k$ $(1 \le k \le n)$,*

$$|\lambda - a_{kk}| \le |a_{k1}| + |a_{k2}| + ...... + |a_{kn}| = \rho_k \text{ (say)},$$

*i.e. All the eigen values of A lie in the union of the n circles with centers $a_{kk}$ and radii $\rho_k$.*

*Proof.* Let $\lambda$ be an eigen value of an arbitrary square matrix $A$ and $X$ be the corresponding eigen vector. Then $AX = \lambda X$

or
$$a_{11}x_1 + a_{12}x_2 + ...... + a_{1n}x_n = \lambda x_1$$
$$......................................................$$
$$a_{k1}x_1 + a_{k2}x_2 + ...... + a_{kn}x_n = \lambda x_k$$
$$......................................................$$
$$a_{n1}x_1 + a_{n2}x_2 + ...... + a_{nn}x_n = \lambda x_n$$

If $x_k$ be the largest component of $X$, then $|x_m/x_k| \le 1$ $(m = 1, 2, ......, n)$ ...(1)

Dividing the $k$th equation by $x_k$, we obtain

$$a_{k1}(x_1/x_k) + ...... + a_{k,k-1}(x_{k-1}/x_k) + a_{kk} + ...... + a_{kn}(x_n/x_k) = \lambda$$

or
$$\lambda - a_{kk} = a_{k1}(x_1/x_k) + ...... + a_{k,k-1}(x_{k-1}/x_k) + ...... + a_{kn}(x_n/x_k)$$

Taking absolute values on both sides and using the theorem $|a + b| \le |a| + |b|$, we obtain $|\lambda - a_{kk}| \le |a_{k1}| + ...... + |a_{k,k-1}| + ...... + |a_{kn}| = \rho_k$ (say) [by (1)]

This shows that all the eigen values of $A$ lie within or on the union of the circles with centres $a_{kk}$ and radii $\rho_k$.

As $A$ and $A'$ have the same eigen values, the above theorem is also true for columns. *These circles are called the* **Gerschgorin circles.**

The bounds thus obtained being all independent ; all the eigen values of $A$ must lie in the intersection of these bounds. *These bounds are called the* **Gerschgorin bounds.**

The above theorem gives us the possible location of the eigen values and also helps us to estimate their bounds. *If any of the Gerschgorin circles is isolated, then it contains exactly one eigen value.*

## 4.11. POWER METHOD

In many engineering problems, it is required to compute the numerically largest eigen value and the corresponding eigen vector. In such cases, the following *iterative method* is quite convenient which is also well-suited for machine computations.

If $X_1, X_2 ...... X_n$ be the eigen vectors corresponding to the eigen-values $\lambda_1, \lambda_2, ......, \lambda_n$, then an arbitrary column vector can be written as

$$X = k_1X_1 + k_2X_2 + ...... + k_nX_n$$

Then
$$AX = k_1 AX_1 + k_2 AX_2 + ...... + k_n AX_n$$
$$= k_1 \lambda_1X_1 + k_2 \lambda_2X_2 + ...... + k_n \lambda_nX_n$$

Similarly
$$A^2X = k_1\lambda_1^2X_1 + k_2\lambda_2^2X_2 + ...... + k_n\lambda_n^2X_n$$

and
$$A^rX = k_1 \lambda_1^rX_1 + k_2 \lambda_2^rX_2 + ...... + k_n \lambda_n^rX_n$$

If $|\lambda_1| > |\lambda_2| > ...... > |\lambda_n|$, then $\lambda_1$ is the largest root and the contribution of the term $k_1\lambda_1^rX_1$ to the sum on the right increases with $r$ and, therefore, every time we multiply a column vector by $A$, it becomes nearer to the eigen vector $X_1$. Then we make the largest component of the resulting column vector unity to avoid the factor $k_1$.

*Thus we start with a column vector $X$ which is as near the solution as possible and evaluate $AX$ which is written as $\lambda^{(1)} X^{(1)}$ after normalisation. This gives the first approximation $\lambda^{(1)}$ to the eigen-value and $X^{(1)}$ to the eigen vector. Similarly we evaluate $AX^{(1)} = \lambda^{(2)} X^{(2)}$ which gives the second approximation. We repeat this process till $[X^{(r)} - X^{(r-1)}]$ becomes negligible. Then $\lambda^{(r)}$ will be the largest eigen-value and $X^{(r)}$, the corresponding eigen vector.*

This iterative procedure for finding the dominant eigen value of a matrix is known as *Rayleigh's power method*.

**Obs.** Rewriting $AX = \lambda X$ as $A^{-1}AX = \lambda A^{-1}X$ or $X = \lambda A^{-1}X$.

We have $A^{-1}X = \dfrac{1}{\lambda} X$

If we use this equation, then *the above method yields the smallest eigen value.*

**Example 4.10.** *Determine the largest eigen value and the corresponding eigen vector of the matrix* $\begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$.

**Sol.** Let the initial approximation to the eigen vector corresponding to the largest eigen value of $A$ be $X = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

Then
$$AX = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} = 5\begin{bmatrix} 1 \\ 0.2 \end{bmatrix} = \lambda^{(1)} X^{(1)}$$

So the first aporoximation to the eigen value is $\lambda^{(1)} = 5$ and the corresponding eigen vector is $X^{(1)} = \begin{bmatrix} 1 \\ 0.2 \end{bmatrix}$.

Now
$$AX^{(1)} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.2 \end{bmatrix} = \begin{bmatrix} 5.8 \\ 1.4 \end{bmatrix} = 5.8\begin{bmatrix} 1 \\ 0.241 \end{bmatrix} = \lambda^{(2)} X^{(2)}$$

Thus the second aporoximation to the eigen-value is $\lambda^{(2)} = 5.8$ and the corresponding eigen vector is $X^{(2)} = \begin{bmatrix} 1 \\ 0.241 \end{bmatrix}$, repeating the above process, we get



$$AX^{(2)} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.241 \end{bmatrix} = 5.966\begin{bmatrix} 1 \\ 0.248 \end{bmatrix} = \lambda^{(3)} X^{(3)}$$

$$AX^{(3)} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.249 \end{bmatrix} = 5.994\begin{bmatrix} 1 \\ 0.250 \end{bmatrix} = \lambda^{(4)} X^{(4)}$$

$$AX^{(4)} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.250 \end{bmatrix} = 5.999\begin{bmatrix} 1 \\ 0.25 \end{bmatrix} = \lambda^{(5)} X^{(5)}$$

$$AX^{(5)} = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0.25 \end{bmatrix} = 6\begin{bmatrix} 1 \\ 0.25 \end{bmatrix} = \lambda^{(6)} X^{(6)}$$

Clearly $\lambda^{(5)} = \lambda^{(6)}$ and $X^{(5)} = X^{(6)}$ upto 3 decimal places. Hence the largest eigen-value is 6 and the corresponding eigen vector is $\begin{bmatrix} 1 \\ 0.25 \end{bmatrix}$.

**Example 4.11.** *Find the largest eigen value and the corresponding eigen vector of the matrix* $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ *using power method. Take $[1, 0, 0]^T$ as initial eigen vector.*

(V.T.U., B.E., 2009)

**Sol.** Let the initial approximation to the required eigen vector be $X = [1, 0, 0]'$.

Then
$$AX = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = 2\begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} = \lambda^{(1)} X^{(1)}$$

So the first approximation to the eigen value is 2 and the corresponding eigen vector $X^{(1)} = [1, -0.5, 0]'$.

Hence
$$AX^{(1)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 2.5 \\ -2 \\ 0.5 \end{bmatrix} = 2.5\begin{bmatrix} 1 \\ -0.8 \\ 0.2 \end{bmatrix} = \lambda^{(2)} X^{(2)}$$

Repeating the above process, we get

$$AX^{(2)} = 2.8\begin{bmatrix} 1 \\ -1 \\ 0.43 \end{bmatrix} = \lambda^{(3)} X^{(3)}; \quad AX^{(3)} = 3.43\begin{bmatrix} 0.87 \\ -1 \\ 0.54 \end{bmatrix} = \lambda^{(4)} X^{(4)};$$

$$AX^{(4)} = 3.41\begin{bmatrix} 0.80 \\ -1 \\ 0.61 \end{bmatrix} = \lambda^{(5)} X^{(5)}; \quad AX^{(5)} = 3.41\begin{bmatrix} 0.76 \\ -1 \\ 0.65 \end{bmatrix} = \lambda^{(6)} X^{(6)}$$

$$AX^{(6)} = 3.41\begin{bmatrix} 0.74 \\ -1 \\ 0.67 \end{bmatrix} = \lambda^{(7)} X^{(7)}$$

Clearly $\lambda^{(6)} = \lambda^{(7)}$ and $X^{(6)} = X^{(7)}$ approximately.

Hence the largest eigen value is 3.41 and the corresponding eigen-vector is $[0.74, -1, 0.67]'$.

**Example 4.12.** *Obtain by power method, the numerically dominant eigen value and eigen vector of the matrix.*

$$A = \begin{bmatrix} 15 & -4 & -3 \\ -10 & 12 & -6 \\ -20 & 4 & -2 \end{bmatrix}.$$ (*Anna, B.Tech., 2007*)

**Sol.** Let the initial approximation to the eigen vector be $X = [1, 1, 1]'$. Then

$$AX = \begin{bmatrix} 15 & -4 & -3 \\ -10 & 12 & -6 \\ -20 & 4 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ -4 \\ -18 \end{bmatrix} = -18\begin{bmatrix} -0.444 \\ 0.222 \\ 1 \end{bmatrix} = \lambda^{(1)}X^{(1)}$$

So the first approximation to eigen value is $-18$ and the corresponding eigen vector is $[-0.444, 0.222, 1]'$.

Now $$AX^{(1)} = \begin{bmatrix} 15 & -4 & -3 \\ -10 & 12 & -6 \\ -20 & 4 & -2 \end{bmatrix}\begin{bmatrix} -0.444 \\ 0.222 \\ 1 \end{bmatrix} = -10.548\begin{bmatrix} 1 \\ -0.105 \\ -0.736 \end{bmatrix} = \lambda^{(2)}X^{(2)}$$

$\therefore$ The second approximation to the eigen value is $-10.548$ and the eigen vector is $[1, -0.105, -0.736]'$.

Repeating the above process

$$AX^{(2)} = -18.948\begin{bmatrix} -0.930 \\ 0.361 \\ 1 \end{bmatrix} = \lambda^{3}X^{(3)};\ AX^{(3)} = -18.394\begin{bmatrix} 1 \\ -0.415 \\ -0.981 \end{bmatrix} = \lambda^{(4)}X^{(4)}$$

$$AX^{(4)} = -19.698\begin{bmatrix} -0.995 \\ 0.462 \\ 1 \end{bmatrix} = \lambda^{(5)}X^{(5)};\ AX^{(5)} = -19.773\begin{bmatrix} 1 \\ -480 \\ -0.999 \end{bmatrix} = \lambda^{(6)}X^{(6)}$$

$$AX^{(6)} = -19.922\begin{bmatrix} -0.997 \\ 0.490 \\ 1 \end{bmatrix} = \lambda^{(7)}X^{(7)};\ AX^{(7)} = -19.956\begin{bmatrix} 1 \\ -0.495 \\ -0.999 \end{bmatrix} = \lambda^{(8)}X^{(8)}$$

Since $\lambda^{(7)} = \lambda^{(8)}$ and $X^{(7)} = X^{(8)}$ approximately, therefore the dominant eigen value and the corresponding eigen vector are given by

$$\lambda^{(8)}X^{(8)} = 19.956\begin{bmatrix} -1 \\ 0.495 \\ 0.999 \end{bmatrix} \text{ i.e., } 20\begin{bmatrix} -1 \\ 0.5 \\ 1 \end{bmatrix}$$

Hence the dominant eigen value is 20 and eigen vector is $[-1, 0.5, 1]'$.

## PROBLEMS 4.2

1. Find the eigen values and eigen vectors of the matrices

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

2. Find the latent roots and the latent vectors of the matrices

   (*a*) $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ (*J.N.T.U., B. Tech., 2005*) (*b*) $\begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$ (*U.P.T.U., B. Tech., 2006*)



   (*c*) $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$ (*V.T.U., B.E., 2009*)

3. Using Cayley-Hamilton theorem, find the inverse of

   (*i*) $\begin{bmatrix} 1 & 1 & 2 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (*U.P.T.U., B. Tech., 2006*)

   (*ii*) $\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$ (*Rajasthan, B.E., 2005*)

4. Find, by power method, the larger eigen-value of the following matrices :

   (*a*) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ (*Anna, B. Tech., 2005*) (*b*) $\begin{bmatrix} 4 & 1 \\ 1 & 3 \end{bmatrix}$

5. Find the largest eigen-value and the corresponding eigen-vector of the matrices :

   (*a*) $\begin{bmatrix} 1 & 3 & -1 \\ 3 & 2 & 4 \\ -1 & 4 & 10 \end{bmatrix}$ (*Madras, B. Tech., 2006*)

   (*b*) $\begin{bmatrix} 1 & -3 & 2 \\ 4 & 4 & -1 \\ 6 & 3 & 5 \end{bmatrix}$ (*Anna, B. Tech., 2005*)

   (*c*) $\begin{bmatrix} 25 & 1 & 2 \\ 1 & 3 & 0 \\ 2 & 0 & -4 \end{bmatrix}$ taking $[1, 0, 0]^T$ as initial eigen vector. (*V.T.U., B. Tech., 2008*)

**Obs.** The iteration method is a special method as it gives the largest or the smallest eigen value only. Now we shall describe three modern methods for finding all the eigen values of a real symmetric matrix $A$.

The eigen values of $A$ are given by the diagonal elements when $A$ is reduced to either the diagonal matrix $D$ or the lower triangular matrix $L$ or the upper triangular matrix $U$.

Thus the methods of finding eigen values of $A$ are based on reducing $A$ to $D$ or $L$ or $U$.

## 4.12. JACOBI'S METHOD

Let $A$ be a given real symmetric matrix. Its eigen values are real and there exists a real orthogonal matrix $B$ such that $B^{-1}AB$ is a diagonal matrix $D$. Jacobi's method consists of diagonalising $A$ by applying a series of orthogonal transformations $B_1, B_2, \ldots, B_r$ such that their product $B$ satisfies the equation $B^{-1}AB = D$.

For this purpose, we choose the numerically largest non-diagonal element $a_{ij}$ and form a $2 \times 2$ submatrix $A_1 = \begin{bmatrix} a_{ij} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix}$.

where $a_{ij} = a_{ji}$, which can easily be diagonalised.

Consider an orthogonal matrix $B_1 = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ so that $B_1^{-1} = B_1'$.

Then $B^{-1}A_1B_1 = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

$$= \begin{bmatrix} a_{ii}\cos^2\theta + a_{jj}\sin^2\theta + a_{ij}\sin 2\theta, & a_{ij}\cos 2\theta + \frac{1}{2}(a_{jj} - a_{ii})\sin 2\theta \\ a_{ij}\cos 2\theta + \frac{1}{2}(a_{jj} - a_{ii})\sin 2\theta, & a_{ii}\sin^2\theta + a_{jj}\cos^2\theta - a_{ij}\sin 2\theta \end{bmatrix} \quad ...(1)$$

Now this matrix will reduce to the diagonal form, if $a_{ij}\cos 2\theta + \frac{1}{2}(a_{jj} - a_{ii})\sin 2\theta = 0$

*i.e.* if
$$\tan 2\theta = \frac{2a_{ij}}{a_{ii} - a_{jj}} \quad ...(2)$$

This equation gives four values of $\theta$, but to get the least possible rotation, we choose $-\pi/4 \le \theta \le \pi/4$.

Thus (1) reduces to a diagonal matrix.

As a next step, the largest non-diagonal element (in magnitude) in the new rotated matrix is found and the above procedure is repeated using the orthogonal matrix $B_2$.

In this way, a series of such transformations are performed so as to annihiliate the non-diagonal elements. After making $r$ transformations, we obtain

$$B_r^{-1} B_{r-1}^{-1} \ldots B_1^{-1} A B_1 \ldots\ldots B_{r-1} B_r = B^{-1} AB$$

As $r \to \infty$, $B^{-1}AB$ approaches a diagonal matrix whose diagonal elements are the eigen values of $A$.

Also the corresponding columns of $B = B_1B_2 \ldots\ldots B_r$, are the eigen vectors of $A$.

**Example 4.13.** *Using Jacobi's method, find all the eigen values and the eigen vectors of the matrix*

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

**Sol.** Here the largest non-diagonal element is $a_{13} = a_{31} = 2$. Also $a_{11} = 1$ and $a_{33} = 1$.

$\therefore$
$$\tan 2\theta = \frac{2a_{13}}{a_{11} - a_{33}} = \frac{2 \times 2}{1 - 1} \to \infty$$

*i.e.*
$$2\theta = \pi/2 \quad \text{or} \quad \theta = \pi/4$$

Then
$$B_1 = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \text{ and } B_1^{-1} = B'$$

$\therefore$ The first transformation gives

$$D_1 = B_1^{-1}AB_1 = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & \sqrt{2} & 2 \\ \sqrt{2} & 3 & \sqrt{2} \\ 2 & \sqrt{2} & 1 \end{bmatrix} \times \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}$$

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



Now the largest non-diagonal element is $a_{12} = a_{21} = 2$. Also $a_{11} = 3$ and $a_{22} = 3$.

$\therefore$
$$\tan 2\theta = \frac{2a_{12}}{a_{11} - a_{22}} = \frac{2 \times 2}{0} \to \infty.$$

*i.e.*
$$2\theta = \pi/2 \quad \text{or} \quad \theta = \pi/4.$$

Then
$$B_2 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$\therefore$ The second transformation gives

$$B_2^{-1} D_1 B_2 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix} \times \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Hence the eigen values of the given matrix are 5, 1, – 1 and the corresponding eigen-vectors are the columns of

$$B = B_1B_2 = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/2 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

**Note.** A disadvantage of Jacobi's method is that the element annihiliated by a transformation, may not remain zero during the subsequent transformations. *Given's suggested a reduction which does not disturb zeros already formed. But instead of leading to a diagonal matrix as in Jacobi's method, the Given's method leads to a tri-diagonal matrix.* The eigen values and eigen vectors of the original matrix have to be found from those of the tri-diagonal matrix.

**Example 4.14.** *Obtain using Jacobi's method, all the eigen values and eigen vectors of the matrix*

$$A = \begin{bmatrix} 1 & 1 & 0.5 \\ 1 & 1 & 0.25 \\ 0.5 & 0.25 & 2 \end{bmatrix}$$

**Sol.** Here the largest non-diagonal element is $a_{12} = 1$.

Also $a_{11} = 1$, $a_{22} = 1$.

$\therefore$
$$\tan 2\theta = \frac{2a_{12}}{a_{11} - a_{12}} = \frac{2 \times 1}{0} \to \infty.$$

*i.e.,*
$$2\theta = \frac{\pi}{2} \quad \text{or} \quad \theta = \frac{\pi}{4}.$$

Then
$$B_1 = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } B_1^{-1} = B_1'$$

$\therefore$ The first transformation is

$$D_1 = B_1^{-1}AB_1 = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1/2 \\ 1 & 1 & 1/4 \\ 1/2 & 1/4 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & 3\sqrt{2}/8 \\ 0 & 0 & -\sqrt{2}/8 \\ 3\sqrt{2}/8 & -\sqrt{2}/8 & 2 \end{bmatrix}$$

Now the largest non-diagonal element of $+ D_1$ is $a_{13} = 3\sqrt{2}/8$. Also $\alpha_{11} = 2, \alpha_{33} = 2$.

$$\therefore \qquad \tan 2\theta = \frac{2\alpha_{13}}{\alpha_{11} - \alpha_{33}} \to \infty, \text{ i.e., } 2\theta = \frac{\pi}{2} \text{ or } \theta = \frac{\pi}{4}.$$

Then
$$B_2 = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}$$

$\therefore$ The second transformation gives

$$D_2 = B_2^{-1} D_1 B_2 = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 & 3\sqrt{2}/8 \\ 0 & 0 & -\sqrt{2}/8 \\ 3\sqrt{2}/8 & -\sqrt{2}/8 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}$$

$$= \begin{bmatrix} 2.530 & -0.125 & 0 \\ -0.125 & 0 & -0.125 \\ 0 & -0.125 & 1.47 \end{bmatrix}$$

Repeating the above steps, we obtain

$$B_3 = \begin{bmatrix} 0.998 & 0.049 & 0 \\ -0.049 & 0.998 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and
$$D_3 = B_3^{-1} D_2 B_3 = \begin{bmatrix} 2.536 & -0.000 & 0.006 \\ -0.000 & -0.006 & -0.125 \\ 0.006 & -0.125 & 1.469 \end{bmatrix}$$

Hence the eigen values of $A$ are 2.536, – 0.006, 1.469 approximately and the corresponding eigen vectors are the columns of $B = B_1 B_2 B_3 = \begin{bmatrix} 0.531 & -0.721 & -0.444 \\ 0.461 & 0.686 & -0.562 \\ 0.710 & 0.094 & 0.698 \end{bmatrix}$

## 4.13. GIVEN'S METHOD

If $A$ be a real symmetric matrix, then Given's method consists of the following steps:

*Step I : To reduce A to a tri-diagonal symmetric matrix :*

To begin with, consider the matrix $A_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$ ...(1)



and the orthogonal rotation matrix $B_1$ in the plane (2, 3) as $B_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 1 & \sin\theta & \cos\theta \end{bmatrix}$

$$\text{Then } B_1^{-1} A_1 B_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & \cos\theta & -\sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 1 \\ 0 & \cos\theta & -\sin\theta \\ 1 & \sin\theta & \cos\theta \end{bmatrix}$$

In the resulting matrix, (1, 3) element $= -a_{12}\sin\theta + a_{13}\cos\theta$. It will be zero, if $-a_{12}\sin\theta + a_{13}\cos\theta = 0$ *i.e.* if $\tan\theta = a_{13}/a_{12}$. ...(2)

Thus with this value of $\theta$, the above transformation gives zeros in (1, 3) and (3, 1) positions.

Now we perform rotation in the plane (2, 4) and put the resulting element (1, 4) = 0. This would not affect the zeros obtained earlier. Proceeding in this way, the transformations are applied to the matrix so as to annihilate the elements (1, 3), (1, 4), (1, 5), ......, (1, $n$); (2, 4), (2, 5), ......, (2, $n$) in this order. Finally we arrive at the tri-diagonal matrix

$$P = \begin{bmatrix} p_1 & q_1 & 0 & 0 \ldots\ldots 0 \\ q_1 & p_2 & q_2 & 0 \ldots\ldots 0 \\ 0 & q_2 & p_3 & q_3 \ldots\ldots 0 \\ \ldots\ldots & & p_{n-1} & q_{n-1} \\ 0 & 0 & 0 & 0 \ldots q_{n-1} \ldots p_n \end{bmatrix}$$

*Step II. To find the eigen values of a tri-diagonal matrix.*

Let the resulting tri-diagonal matrix after first transformation be

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{12} & \alpha_{22} & \alpha_{23} \\ 0 & \alpha_{23} & \alpha_{33} \end{bmatrix} \qquad ...(3)$$

Then the eigen values of (1) and (3) are the same. To obtain the eigen values of (3), we have

$$\begin{bmatrix} \alpha_{11} - \lambda & \alpha_{12} & 0 \\ \alpha_{12} & \alpha_{22} - \lambda & \alpha_{23} \\ 0 & \alpha_{23} & \alpha_{33} - \lambda \end{bmatrix} = 0 = f_3(\lambda) \quad \text{[say]}$$

$$\therefore \qquad f_0(\lambda) = 1,\ f_1(\lambda) = \alpha_{11} - \lambda = \alpha_{11} - \lambda f_0(\lambda)$$

$$f_2(\lambda) = \begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} \\ \alpha_{12} & \alpha_{22} - \lambda \end{vmatrix} = (\alpha_{22} - \lambda) f_1(\lambda) - \alpha_{12}^2 f_0(\lambda)$$

Expanding $f_3(\lambda)$ in terms of the third row, we get

$$f_3(\lambda) = (\alpha_{33} - \lambda) \begin{vmatrix} \alpha_{11} - \lambda & \alpha_{12} \\ \alpha_{12} & \alpha_{22} - \lambda \end{vmatrix} - \alpha_{23} \begin{vmatrix} \alpha_{11} - \lambda & 0 \\ \alpha_{12} & \alpha_{23} \end{vmatrix} \qquad ...(4)$$

*i.e.*
$$f_3(\lambda) = (\alpha_{33} - \lambda) f_2(\lambda) - (\alpha_{23})^2 f_1(\lambda)$$

In general, the recurrence formula is

$$f_k(\lambda) = (\alpha_{kk} - \lambda) f_{k-1}(\lambda) - (\alpha_{k-1,k})^2 f_{k-2}(\lambda),\ 2 \le k \le n$$

The equation $f_k(\lambda) = 0$ is the characteristic equation which can be solved by any standard method. Thus the roots of (5) will be the eigen values of the given symmetric matrix.

*Step III. To find the eigen vectors of the tri-diagonal matrix.*

If $Y$ be an eigen vector of the tri-diagonal matrix $P$ and if $B_1, B_2, \ldots B_j$ are the orthogonal matrices employed in reducing the matrix $A$ to the form $P$, then the corresponding eigen vector of $A$ is given by $X = B_1B_2 \ldots B_j\,Y$.

**Obs. 1.** *The number of rotations required for the Given's method are equivalent to the number of non-tri-diagonal elements of the matrix.* In the case of a 3 × 3 matrix, only one rotation is required; whereas for a 4 × 4 matrix, 3 rotations are needed and so on.

The amount of computation goes on decreasing from one rotation to the next, as the order of the matrix for computation also starts reducing.

**Obs. 2.** *The sequence of functions $f_0(\lambda), f_1(\lambda), f_2(\lambda), \ldots\ldots f_k(\lambda)$ is called the* **Strum sequence.** A table of this sequence for various values of $\lambda$ is prepaired and the number of changes in sign of the Strum sequence is calculated. *Then the difference between these number of changes of sign for consecutive values of $\lambda$ gives an approximate location of the eigen values.* Once the location of the eigen values is known, their exact values can be found by any iterative method *e.g.* Newton-Raphson method.

**Example 4.15.** *Using Given's method, reduce the following matrix to the tri-diagonal form :*

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

**Sol.** There being only one non-tri-diagonal element $a_{13}(= 3)$ which has to be reduced to zero, only one rotation is required.

To annihilate $a_{13}$, we define the orthogonal matrix in the plane (2, 3) as :

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

where $\theta$ is found from the formula

$$\tan\theta = \frac{a_{13}}{a_{12}} = \frac{3}{1} = 3 \quad \text{and hence} \quad \sin\theta = 3/\sqrt{10} \text{ and } \cos\theta = 1/\sqrt{10}.$$

$$\therefore \qquad A_1 = B^{-1}AB = \begin{bmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{12} & \alpha_{22} & \alpha_{23} \\ 0 & \alpha_{23} & \alpha_{33} \end{bmatrix}$$

where $\alpha_{11} = 2,\ \alpha_{12} = a_{12}\cos\theta + a_{13}\sin\theta = \sqrt{10} = 3.16,$

$\alpha_{22} = a_{22}\cos^2\theta + 2a_{23}\sin\theta\cos\theta + a_{33}\sin^2\theta = 4.3$

$\alpha_{23} = (a_{33} - a_{22})\sin\theta\cos\theta + a_{23}(\cos^2\theta - \sin^2\theta) = -1.9$

$\alpha_{33} = a_{22}\sin^2\theta + a_{33}\cos^2\theta = 3.9.$

Hence $A$ is reduced to the tri-diagonal matrix $\begin{bmatrix} 2 & 3.16 & 0 \\ 3.16 & 4.3 & -1.9 \\ 0 & -1.9 & 3.9 \end{bmatrix}$

**Note.** An alternative procedure for reduction of a symmetric matrix to the tri-diagonal form has been suggested by Householder. This method, though more complicated, requires half as much computation, as the Given's method. In any case, it is a substantial improvement on the Given's procedure since it reduces an entire row and column by a single transformation. In this method, the matrix is reduced to tri-diagonal form using elementary orthogonal transformations.



## 4.14. HOUSE-HOLDER'S METHOD

Consider an $n$th order real symmetric matrix $A = [a_{ij}]$. This method consists in pre and post-multiplying $A$ by a real symmetric orthogonal matrix $P$ such $PAP$ reduces to the tri-diagonal form.

Let the matrix $P$ be of the form $\quad P = I - 2ww'$

where $w$ is a column matrix such that ...(1)

$$ww' = w_1^2 + w_2^2 + \ldots\ldots + w_n^2 = 1$$

Then $\quad P' = (I - 2ww')' = I - 2ww' = P$ ...(2)

and $\quad P'P = (I - 2ww')'(I - 2ww')$

$$= I - 4ww' + 4ww'\,.\,ww' = I \qquad \text{[by (2)]}$$

Thus $P$ is a symmetric orthogonal matrix.

Now take $w$ with first $(k - 1)$ zero components, so that

$$w_k' = [0, 0, \ldots\ldots, 0, x_k, \ldots\ldots, x_n]$$

Since $w_k'w_k = 1$, we have $x_k^2 + x_{k+1}^2 + \ldots\ldots x_n^2 = 1$ ...(3)

Then $P_k^{-1}AP_k = P_k'AP_k = P_kAP_k$

We now form successively $A_k = P_kA_{k-1}P_k$ ; $k = 2, 3, \ldots\ldots, n - 1$.

As a first transformation, we determine $x$'s so that zeros are created in the positions (1, 3), (1, 4), ......, (1, $n$) and (3, 1), (4, 1), ......, ($n$, 1).

As a second transformation, we find $x$'s so that zeros are created in the positions (2, 4), (2, 5), ......, (2, $n$) and (4, 2), (5, 2), ......, ($n$, 2).

After $(n - 2)$ such transformations, we arrive at a tri-diagonal matrix.

**Example 4.16.** *Using House-holder's method, reduce the following matrix to the tri-diagonal form :*

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

**Sol.** Let $\quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$

Here we take $\quad w_2' = [0, x_2, x_3]$

so that

$$P = 1 - 2w_2w_2' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 - 2x_2^2 & -2x_2x_3 \\ 0 & -2x_2x_3 & 1 - 2x_3^2 \end{bmatrix}$$

Now the element (1, 3) of $PAP$ can become zero only if the corresponding element in $AP$ is zero. The first row elements of $AP$ are $a_{11}, a_{12} - 2p_1x_2, a_{13} - 2p_1x_3$ where $p_1 = a_{12}x_2 + a_{13}x_3$.

$\therefore$ We require that $a_{13} - 2p_1x_3 = 0$ ...(i)

Since the sum of the squares of the elements in any row is invariant under an orthogonal transformation, we have $a_{11}^2 + a_{12}^2 + a_{13}^2 = a_{11}^2 + (a_{12} - 2p_1x_2)^2 + 0$

or

$$a_{12} - 2p_1x_2 = \pm\sqrt{(a_{12}^2 + a_{13}^2)} \qquad \text{...(ii)}$$

For the given matrix, (i) and (ii) become ...(iii)

$$3 - 2p_1x_3 = 0$$

$$4 - 2p_1x_2 = \pm\sqrt{(4^2 + 3^2)} = \pm 5 \qquad ...(iv)$$

where $p_1 = 4x_2 + 3x_3$

Multiplying $(iii)$ by $x_3$ and $(iv)$ by $x_2$ and adding, we get

$$3x_3 + 4x_2 - 2p_1(x_3^2 + x_2^2) = \pm 5x_2 \qquad [\because x_2^2 + x_3^2]$$

i.e.
$$p_1 - 2p_1 = \pm 5x_2 \quad \text{or} \quad p_1 = \mp 5x_2 \qquad ...(v)$$

Substituting in $(iv)$, we obtain $4 - 2(\mp 5x_2)x_2 = \pm 5$

which gives $x_2 = 1/\sqrt{10}$ or $x_2 = 3/\sqrt{10}$

From $(iii)$,
$$x_3 = \frac{3}{2p_1} = \mp \frac{3}{10x_2} \qquad [\text{by } (v)]$$

Since $x_3$ contains $x_2$ in the denominator, we obtain best accuracy if $x_2$ is large.

$\therefore$ Choosing $x_2 = \dfrac{3}{\sqrt{10}}, x_3 = \mp \dfrac{3}{10} \times \dfrac{\sqrt{10}}{3} = \pm \dfrac{1}{\sqrt{10}}$

Taking + ve sign, we get $x_2 = \dfrac{3}{\sqrt{10}}$ and $x_3 = \dfrac{1}{\sqrt{10}}$.

$\therefore$
$$P = I - 2w_2w_2'$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 2\begin{bmatrix} 0 \\ 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix} [0, 3/\sqrt{10}, 1/\sqrt{10}]$$

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

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

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -4/5 & -3/5 \\ 0 & -3/5 & -4/5 \end{bmatrix}\begin{bmatrix} 1 & -5 & 0 \\ 4 & -2 & 1 \\ 3 & -11/5 & -2/5 \end{bmatrix} = \begin{bmatrix} 1 & -5 & 0 \\ -5 & 73/25 & -14/25 \\ 0 & -14/25 & -11/25 \end{bmatrix}$$

which is the required tri-diagonal matrix.

## PROBLEMS 4.3

1. Using Jacobi's method, find all the eigen values and the eigen vectors of the matrices:

   $(a)\ \begin{bmatrix} 5 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 5 \end{bmatrix}$ $\qquad (b)\ \begin{bmatrix} 2 & 3 & 1 \\ 3 & 2 & 2 \\ 1 & 2 & 1 \end{bmatrix}$

2. Reduce the matrix $\begin{bmatrix} 1 & 1 & 1/2 \\ 1 & 1 & 1/4 \\ 1/2 & 1/4 & 0 \end{bmatrix}$ to the tri-diagonal form, using the Householder's method



3. Apply the Householder's method, to find the eigen values of the matrix $\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$.

4. Transform the matrix $\begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$ to the tri-diagonal form using Given's method. Hence find the largest eigen value and the corresponding eigen vector of the tri-diagonal matrix.

5. Find the eigen values of the matrix $\begin{bmatrix} 2 & -i & 0 \\ i & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$.

## 4.15. OBJECTIVE TYPE OF QUESTIONS

## PROBLEMS 4.4

*Select the correct answer or fill up the blanks in the following questions :*

1. The eigen values of a triangular matrix are ......... .
2. Inverse of $\begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$ is ......... .
3. The most suitable initial eigen vector out of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, to find the larger eigen value of the matrix $\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$ in one iteration, is ......... .
4. Two eigen-values of the matrix $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$ are equal to 1 each, then the eigen values of $A^{-1}$ are ......... .
5. Eigen values of $\begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$ are ......... .
6. If $\lambda$ is an eigen value of a matrix $A$, then $1/\lambda$ is the eigen value of ......... .
7. The product of two eigen values of the matrix $\begin{bmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{bmatrix}$ is 16, then the third eigen value is ......... .
8. Power method works satisfacotorily only if the matrix $A$ has a ...... eigen value.
9. Eigen values of the matrix $\begin{bmatrix} 1 & 1-i \\ 1+i & 1 \end{bmatrix}$ are ......... .
10. If $\lambda$ is an eigen value of an orthogonal matrix, then $1/\lambda$ is also its ......... .

11. Dominant eigen values of $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ by Power method are ......... .
12. The eigen values of an idempotent matrix are ......... .
13. If the eigen values of a matrix $A$ are $-4, 3, 1$, then the dominant eigen value of $A$ is .....
14. If $A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix}$, then $A^{-1} =$ ......... .
15. The eigen value that can the obtained by using Power method is ...... .
16. If $\lambda$ is the largest eigen value of the matrix $A$, then the relation giving the smallest eigen value is ......

# 5

# EMPIRICAL LAWS & CURVE-FITTING



## 5.1. (1) INTRODUCTION

In many branches of Applied Mathematics, it is required to express a given data, obtained from observations, in the form of a law connecting the two variables involved. Such a law inferred by some scheme, is known as the *empirical law*. For example, it may be desired to obtain the law connecting the length and the temperature of a metal bar. At various temperatures, the length of the bar is measured. Then, by one of the methods explained below, a law is obtained that represents the relationship existing between temperature and length for the observed values. This relation can then be used to predict the length at an arbitrary temperature.

**(2) Scatter diagram.** To find a relationship between the set of paired observations $x$ and $y$ (say), we plot their corresponding values on the graph, taking one of the variables along the $x$-axis and other along the $y$-axis *i.e.* $(x_1, y_1), (x_2, y_2) ......, (x_n, y_n)$. The resulting diagram showing a collection of dots is called a *scatter diagram*. A smooth curve that approximates the above set of points is known as the *approximating curve*.

**(3) Curve fitting.** Several equations of different types can be obtained to express the given data approximately. But the problem is to find the equation of the curve of 'best fit' which may be most suitable for predicting the unknown values. The process of finding such an equation of 'best fit' is known as *curve-fitting*.

(4) If there are $n$ pairs of observed values then it is possible to fit the given data to an equation that contains $n$ arbitrary constants for we can solve $n$ simultaneous equations for $n$ unknowns. If it were desired to obtain an equation representing these data but having less then $n$ arbitrary constants, then we can have recourse to any of the four methods : *Graphical method, Method of least-squares, Method of group averages and Method of moments*. The graphical method and the method of averages fail to give the values of the unknown constants

uniquely and accurately, while the other methods do. The method of least squares is probably the best to fit a unique curve to a given data. It is widely used in applications and can be easily implemented on a computer.

## 5.2. GRAPHICAL METHOD

*When the curve representing the given data is a* **linear law** $y = mx + c$ *; we proceed as follows :*

(*i*) *Plot the given points on the graph paper* taking a suitable *scale.*

(*ii*) *Draw the straight line of best fit* such that the point are evenly distributed about the line.

(*iii*) *Taking two suitable points* $(x_1, y_1)$ *and* $(x_2, y_2)$ *on the line, calculate* $m$, the slope of the line and $c$, its intercept on $y$-axis.

When the points representing the observed values do not approximate to a straight line, a smooth curve is drawn through them. From the shape of the graph, we try to infer the law of the curve and then reduce it to the form $y = mx + c$.

## 5.3. LAWS REDUCIBLE TO THE LINEAR LAW

We give below some of the laws in common use, indicating the way these can be reduced to the linear form by suitable substitutions :

**(1)** *When the law is* $y = mx^n + c$

Taking $x^n = X$ and $y = Y$, the above law becomes $Y = mX + c$

**(2)** *When the law is* $y = ax^n$.

Taking logarithms of both sides, it becomes $\log_{10} y = \log_{10} a + n \log_{10} x$

Putting $\log_{10} x = X$ and $\log_{10} y = Y$, it reduces to the form

$$Y = nX + c, \text{ where } c = \log_{10} a.$$

**(3)** *When the law is* $y = ax^n + b \log x$.

Writing it as $\dfrac{y}{\log x} = a\dfrac{x^n}{\log x} + b$ and taking $x^n/\log x = X$ and $y/\log x = Y$, the given law becomes, $Y = aX + b$.

**(4)** *When the law is* $y = ae^{bx}$.

Taking logarithms, it becomes $\log_{10} y = (b \log_{10} e) x + \log_{10} a$.

Putting $x = X$ and $\log_{10} y = Y$, it takes the form $Y = mX + c$

where $m = b \log_{10} e$ and $c = \log_{10} a$.

**(5)** *When the law is* $xy = ax + by$.

Dividing by $x$, we have $y = b\dfrac{y}{x} + a$.

Putting $y/x = X$ and $y = Y$, it reduces to the form $Y = bX + a$.

**Example 5.1.** *R is the resistance to motion of a train at speed V ; find a law of the type* $R = a + bV^2$ *connecting R and V, using the following data :*

| | | | | | |
|---|---|---|---|---|---|
| *V (km/hr)* : | 10 | 20 | 30 | 40 | 50 |
| *R (kg/ton)* : | 8 | 10 | 15 | 21 | 30 |



**Sol.** Given law is $R = a + bV^2$ ...(*i*)

Taking $V^2 = x$ and $R = y$,

(*i*) becomes, $y = a + bx$ ...(*ii*)

which is a *linear law.*

Table for the values of $x$ and $y$ is as follows :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 100 | 400 | 900 | 1600 | 2500 |
| $y$ : | 8 | 10 | 15 | 21 | 30 |

Plot these points. Draw the straight line of best fit through these points (Fig. 5.1).

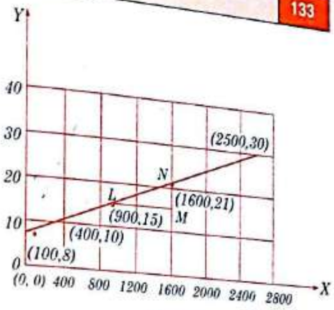


Fig. 5.1

Slope of this line ($= b$)

$$= \frac{MN}{LM} = \frac{21 - 15}{1600 - 900} = \frac{6}{700}$$

$$= 0.0085 \text{ nearly.}$$

Since $L(900, 15)$ lies on (*ii*),

$\therefore$ $15 = a + 0.0085 \times 900,$

whence $a = 15 - 7.65 = 7.35$ nearly.

**Example 5.2.** *The following values of x and y are supposed to follow the law* $y = ax^2 + b \log_{10} x$. *Find graphically the most probable values of the constants a and b.*

| $x$ | 2.85 | 3.88 | 4.66 | 5.69 | 6.65 | 7.77 | 8.67 |
|---|---|---|---|---|---|---|---|
| $y$ | 16.7 | 26.4 | 35.1 | 47.5 | 60.6 | 77.5 | 93.4 |

**Sol.** Given law is $y = ax^2 + b \log_{10} x$

*i.e.*
$$\frac{y}{\log_{10} x} = a\frac{x^2}{\log_{10} x} + b \qquad ...(i)$$

Putting $x^2/\log_{10} x = X$ and $y/\log_{10} x = Y$,

(*i*) becomes $Y = aX + b$ ...(*ii*)

This is a *linear law.*

Table for the values of $X$ and $Y$ is as follows :

| $X = x^2/\log_{10} x$ | 17.93 | 25.56 | 32.49 | 42.87 | 53.75 | 67.80 | 80.83 |
|---|---|---|---|---|---|---|---|
| $Y = y/\log_{10} x$ | 35.59 | 44.83 | 52.50 | 69.90 | 73.65 | 87.04 | 99.56 |
| Points | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ |

Plot these points and draw the straight line of best fit through these points (Fig. 5.2).

Slope of this line ($= a$) $= \dfrac{MP_5}{P_3M} = \dfrac{73.65 - 52.50}{53.75 - 32.49} = \dfrac{21.15}{21.26} = 0.99$

Since $P_3$ lies on (*ii*), therefore,

whence $52.50 = 0.99 \times 32.49 + b$

$b = 20.2$

Hence (*i*) becomes $y = (0.99)\, x^2 + (20.2) \log_{10} x$.

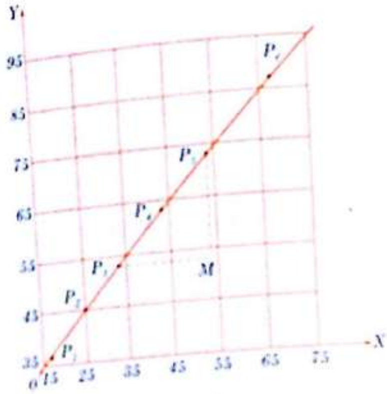


Fig. 5.2

**Example 5.3.** *The values of x and y obtained in an experiment are as follows:*

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ | 2.30 | 3.10 | 4.00 | 4.92 | 5.91 | 7.20 |
| $y$ | 33.0 | 39.1 | 50.3 | 67.2 | 85.6 | 125.0 |

*The probable law is $y = ae^{bx}$. Test graphically the accuracy of this law and if the law holds good, find the best values of the constants.*

**Sol.** Given law is $y = ae^{bx}$

Taking logarithms to base 10, we have

$$\log_{10} y = \log_{10} a + (b \log_{10} e) x$$

Putting $x = X$ and $\log_{10} y = Y$, it becomes

$$y = (b \log_{10} e) X + \log_{10} a$$

Table for the values of $X$ and $Y$ is as under :

| $X = x$ | 2.30 | 3.10 | 4.00 | 4.92 | 5.91 | 7.20 |
|---|---|---|---|---|---|---|
| $Y = \log_{10} y$ | 1.52 | 1.59 | 1.70 | 1.83 | 1.93 | 2.1 |
| Points | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ |

**Scale** : 1 small division along $x$-axis = 0.1

10 small divisions along $y$-axis = 0.1.

Plot these points and draw the line of best fit. As these points are lying almost along a straight line, the given law is nearly accurate (Fig. 5.3).

Now slope of this line $(= b \log_{10} e) = \dfrac{MN}{NM} = 0.12$

whence $$b = \frac{0.12}{\log_{10} e} = 0.12 \times 2.303 = 0.276$$



Since the point $L$ (4, 1.71) lies on (*ii*), therefore,

$$1.71 = 0.12 \times 4 + \log_{10} a \quad \text{whence} \quad a = 17 \text{ nearly.}$$

Hence the curve of best fit is $y = 17\, e^{0.276\, x}$.

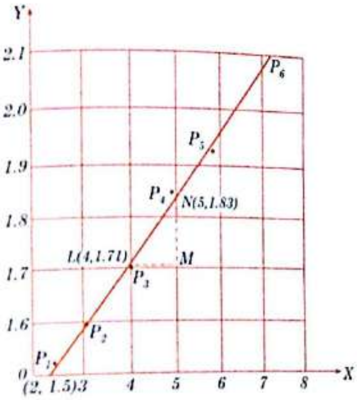


Fig. 5.3

## PROBLEMS 5.1

1. If $p$ is the pull required to lift the weight by means of a pulley block, find a linear law of the form $p = a + bw$, connecting $p$ *and* $w$, using the following data :

| | | | | |
|---|---|---|---|---|
| $w$(lb) : | 50 | 70 | 100 | 120 |
| $p$(lb) : | 12 | 15 | 21 | 25 |

Compute $p$, when $w$ = 150 lb.

2. Convert the following equations to their linear forms :
(*i*) $y = ax + bx^2$ (*ii*) $y = b/[x(x - a)]$.

3. The resistance $R$ of a carbon filament lamp was measured at various values of the voltage $V$ and the following observations were made :

| | | | | | |
|---|---|---|---|---|---|
| Voltage $V$... | 62 | 70 | 78 | 84 | 92 |
| Resistance $R$... | 73 | 70.7 | 69.2 | 67.8 | 66.3 |

Assuming a law of the form $R = \dfrac{a}{V} + b$, find by graphical method the best values of $a$ and $b$.

4. Verify if the values of $x$ and $y$, related as shown in the following table, obey the law $y = a + b\sqrt{x}$. If so, find graphically the values of $a$ and $b$.

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 500 | 1,000 | 2,000 | 4,000 | 6,000 |
| $y$ : | 0.20 | 0.33 | 0.38 | 0.45 | 0.51 |

5. The following table gives the pressure $p$ and the volume $v$ at various instants during the expansion of steam in a cylinder. Show that the equation of the expansion is of the form $pv^n = c$ and find the values of $n$ and $c$ approximately.

| | | | | | | |
|---|---|---|---|---|---|---|
| $p$ : | 200 | 100 | 50 | 30 | 20 | 10 |
| $v$ : | 1.0 | 1.7 | 2.9 | 4.8 | 5.9 | 10 |

6. The following values of $T$ and $l$ follow the law $T = al^n$. Test if this is so and find the best values of $a$ and $n$.

| | | | | |
|---|---|---|---|---|
| $T$ = | 1.0 | 1.5 | 2.0 | 2.5 |
| $l$ = | 25 | 56.2 | 100 | 1.56 |

7. Fit the curve $y = ae^{bx}$ to the following data :

| | | | |
|---|---|---|---|
| $x$ : | 0 | 2 | 4 |
| $y$ : | 5.1 | 10 | 31.1. |

8. The following are the results of an experiment on friction of bearings. The speed being kept constant, corresponding values of the coefficient of friction and the temperature are shown in the table :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t$ : | 120 | 110 | 100 | 90 | 80 | 70 | 60 |
| $\mu$ : | 0.0051 | 0.0059 | 0.0071 | 0.0085 | 0.00102 | 0.00124 | 0.00148 |

If $\mu$ and $t$ are given by the law $\mu = ae^{bt}$, find the values of $a$ and $b$ by plotting the graph for $\mu$ and $t$.

## 5.4. PRINCIPLE OF LEAST SQUARES

The graphical method has the obvious drawback of being unable to give a unique curve of fit. *The principle of least squares, however, provides an elegant procedure of fitting a unique curve to a given data.*

Let the curve $y = a + bx + cx^2 + \ldots\ldots + kx^m$ ...(1)

be fitted to the set of data points $(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$.

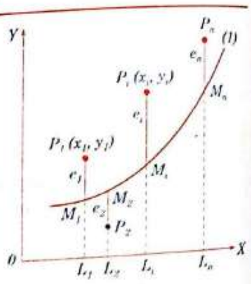

Fig. 5.4

Now we have to determine the constants $a, b, c, \ldots k$ such that it represents the curve of best fit. In case $n = m$, on substituting the values $(x_i, y_i)$ in (1), we get $n$ equations from which a unique set of $n$ constants can be found. But when $n > m$, we obtain $n$ equations which are more than the $m$ constants and hence cannot be solved for these constants. So we try to determine the values of $a, b, c, \ldots k$ which satisfy all the equations as nearly as possible and thus may give the best fit. In such cases, we apply the *principle of least squares*.

At $x = x_i$, the *observed (experimental)* value of the ordinate is $y_i$ and the corresponding value on the fitting curve (1) is $a + bx_i + cx_i^2 + \ldots + kx_i^m$ ($= \eta_i$, say) which is the *expected (or calculated) value* (Fig. 5.4). The difference of the observed and the expected values *i.e.* $y_i - \eta_i$ ($= e_i$) is called the *error (or residual)* at $x = x_i$. Clearly some of the errors $e_1, e_2, \ldots, e_n$ will be positive and others negative. Thus to give equal weightage to each error, we square each of these and form their sum *i.e.* $E = e_1^2 + e_2^2 + \ldots + e_n^2$.

*The curve of best fit is that for which e's are as small as possible i.e., the sum of the squares of the errors is a minimum.* This is known as the *Principle of least squares* and was suggested by a French mathematician *Adrien Marie Legendre* in 1806.



**Obs.** The principle of least squares does not help us to determine the form of the appropriate curve which can fit a given data. It only determines the best possible values of the constants in the equation when the form of the curve is known before hand. The selection of the curve is a matter of experience and practical considerations.

## 5.5. (1) METHOD OF LEAST SQUARES

For clarity, suppose it is required to fit the curve $y = a + bx + cx^2$ to a given set of observations $(x_1, y_1), (x_2, y_2), \ldots, (x_5, y_5)$. For any $x_i$, the observed value is $y_i$ and the expected value is $\eta_i = a + bx_i + cx_i^2$ so that the error $e_i = y_i - \eta_i$.

$\therefore$ The sum of the squares of these errors is

$$E = e_1^2 + e_2^2 + \ldots + e_5^2$$
$$= [y_1 - (a + bx_1 + cx_1^2)]^2 + [y_2 - (a + bx_2 + cx_2^2)]^2 + \ldots\ldots\ldots + [y_5 - (a + bx_5 + cx_5^2)]^2$$

For $E$ to be minimum, we have

$$\frac{\partial E}{\partial a} = 0 = -2[y_1 - (a + bx_1 + cx_1^2)]^2 - 2[y_2 - (a + bx_2 + cx_2^2)] - \ldots\ldots - 2[y_5 - (a + bx_5 + cx_5^2)] \quad \ldots(1)$$

$$\frac{\partial E}{\partial b} = 0 = -2x_1[y_1 - (a + bx_1 + cx_1^2)] - 2x_2[y_2 - (a + bx_2 + cx_2^2)] - \ldots\ldots - 2x_2[y_5 - (a + bx_5 + cx_5^2)] \quad \ldots(2)$$

$$\frac{\partial E}{\partial c} = 0 = -2x_1^2[y_1 - (a + bx_1 + cx_1^2)] - 2x_2^2[y_2 - (a + bx_2 + cx_2^2)] - \ldots\ldots - 2x_5^2[y_5 - (a + bx_5 + cx_5^2)] \quad \ldots(3)$$

Equation (1) simplifies to

$$y_1 + y_2 + \ldots + y_5 = 5a + b(x_1 + x_2 + \ldots + x_5) + c(x_1^2 + x_2^2 + \ldots + x_5^2)$$

*i.e.* $$\Sigma y_i = 5a + b\Sigma x_i + cx_i^2 \quad \ldots(4)$$

Equation (2) becomes

$$x_1y_1 + x_2y_2 + \ldots + x_5y_5 = a(x_1 + x_2 + \ldots + x_5) + b(x_1^2 + x_2^2 + \ldots + x_5^2) + c(x_1^3 + x_2^3 + \ldots + x_5^3)$$

*i.e.* $$\Sigma x_iy_i = a\Sigma x_i + b\Sigma x_i^2 + x\Sigma x_i^3. \quad \ldots(5)$$

Similarly (3) simplifies to $\Sigma x_i^2y_i = a\Sigma x_i^2 + b\Sigma x_i^3 + c\Sigma x_i^4$ ...(6)

The equations (4), (5) and (6) are known as *Normal equations* and can be solved as simultaneous equations in $a, b, c$. The values of these constants when substituted in (1) give the desired curve of best fit.

**Obs.** On calculating $\frac{\partial^2 E}{\partial a^2}, \frac{\partial^2 E}{\partial b^2}, \frac{\partial^2 E}{\partial c^2}$ and substituting the values of $a, b, c$ just obtained, we will observe that each is positive *i.e.* $E$ is a minimum.

### (2) Working procedure

(*a*) *To fit the straight line* $y = a + bx$

(*i*) Substitute the observed set of $n$ values in this equation.

(*ii*) Form normal equations for each constant *i.e.* $\Sigma y = na + b\Sigma x$, $\Sigma xy = a\Sigma x + b\Sigma x^2$

[The normal equation for the unknown $a$ is obtained by multiplying the equations by the coefficient of $a$ and adding. The normal equation for $b$ is obtained by multiplying the equations by the coefficient of $b$ (*i.e.* $x$) and adding.]

(*iii*) Solve these normal equations as simultaneous equations for $a$ and $b$.

(*iv*) Substitute the values of $a$ and $b$ in $y = a + bx$, which is the required line of best fit.

(*b*) *To fit the parabola* : $y = a + bx + cx^2$

(*i*) Form the normal equations $\Sigma y = na + b\Sigma x + c\Sigma x^2$

$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$ and $\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$

[The normal equation for $c$ has been obtained by multiplying the equations by the coefficient of $c$ (*i.e.* $x^2$) and adding.]

(*ii*) Solve these as simultaneous equations for $a, b, c$.

(*iii*) Substitute the values of $a, b, c$ in $y = a + bx + cx^2$, which is the required parabola of best fit.

(*c*) In general, the curve $y = a + bx + cx^2 + \ldots\ldots + kx^{m-1}$ can be fitted to a given data by writing $m$ normal equations.

**Example 5.4.** *If P is the pull required to lift a load W by means of a pulley block, find a linear law of the form $P = mW + c$ connecting P and W, using the following data :*

| | | | |
|---|---|---|---|
| $P = 12$ | 15 | 21 | 25 |
| $W = 50$ | 70 | 100 | 120 |

*where P and W are taken in kg-wt. Compute P when W = 150 kg.* (*U.P.T.U., B. Tech., 2004*)

**Sol.** The corresponding normal equations are

$$\left.\begin{aligned} \Sigma P &= 4c + m\Sigma W \\ \Sigma WP &= c\Sigma W + m\Sigma W^2 \end{aligned}\right\} \qquad \ldots(i)$$

The values of $\Sigma W$ etc. are calculated by means of the following table :

| $W$ | $P$ | $W^2$ | $WP$ |
|---|---|---|---|
| 50 | 12 | 2500 | 600 |
| 70 | 15 | 4900 | 1050 |
| 100 | 21 | 10000 | 2100 |
| 120 | 25 | 14400 | 3000 |
| Total = 340 | 73 | 31800 | 6750 |

$\therefore$ The equations (*i*) become $73 = 4c + 340m$ and $6750 = 340c + 31800m$

*i.e.* $2c + 170m = 365$ ...(*ii*)

and $34c + 3180m = 675$ ...(*iii*)

Multiplying (*ii*) by 17 and subtracting from (*iii*), we get $m = 0.1879$

$\therefore$ from (*ii*), $c = 2.2785$

Hence the line of best fit is $P = 2.2759 + 0.1879\ W$

When $W = 150$ kg, $P = 2.2785 + 0.1879 \times 150 = 30.4635$ kg.

**Example 5.5.** *Fit a straight line to the following data :*

| $x$ | 6 | 7 | 7 | 8 | 8 | 8 | 9 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $y$ | 5 | 5 | 4 | 5 | 4 | 3 | 4 | 3 | 3 |

(*J.N.T.U., B.Tech., 2008*)



**Sol.** Let the straight line be $y = ax + b$.

Then the normal equations are $\Sigma y = a\Sigma x + 9b$

$\Sigma xy = a\Sigma x^2 + b\Sigma x$ ...(*i*)

The values of $\Sigma x$, $\Sigma y$ etc. are calculated below :

| $x$ | $y$ | $xy$ | $x^2$ |
|---|---|---|---|
| 6 | 5 | 30 | 36 |
| 7 | 5 | 35 | 49 |
| 7 | 4 | 28 | 16 |
| 8 | 5 | 40 | 64 |
| 8 | 4 | 32 | 64 |
| 8 | 3 | 24 | 64 |
| 9 | 4 | 36 | 81 |
| 9 | 3 | 27 | 81 |
| 10 | 3 | 30 | 100 |
| $\Sigma x = 72$ | $\Sigma y = 36$ | $\Sigma xy = 282$ | $\Sigma x^2 = 588$ |

$\therefore$ The equations (*i*) become $36 = 72a + 9b$ and $282 = 588a + 72b$

*i.e.,* $8a + b = 4$ ...(*ii*)

$98a + 12b = 47$ ...(*iii*)

Multiplying (*ii*) by 12 and subtracting from (*iii*), we get $a = -0.5$.

From (*ii*), $b = 8$.

Hence the required line of best fit is $y = -0.5x + 8$.

**Example 5.6.** *Fit a second degree parabola to the following data :*

| $x$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $y$ | 1 | 1.8 | 1.3 | 2.5 | 6.3 |

(*P.T.U., 2006*)

**Sol.** Let $u = x - 2$ and $v = y$ so that the parabola of fit $y = a + bx + cx^2$ becomes

$$v = A + Bu + Cu^2 \qquad \ldots(i)$$

The normal equations are

$\Sigma v = 5A + B\Sigma u + C\Sigma u^2$ or $12.9 = 5A + 10C$

$\Sigma uv = A\Sigma u + B\Sigma u^2 + C\Sigma u^3$ or $11.3 = 10B$

$\Sigma u^2 v = A\Sigma u^2 + B\Sigma u^3 + C\Sigma u^4$ or $33.5 = 10A + 34C$

Solving these as simultaneous equations, we get

$A = 1.48, B = 1.13, C = 0.55$

$\therefore$ (*i*) becomes, $v = 1.48 + 1.13u + 0.55u^2$

Hence $y = 1.48 + 1.13\,(x - 2) + 0.55\,(x - 2)^2$

$y = 1.42 - 1.07\,x + 0.55\,x^2$

**Obs.** For the sake of convenience and ease in calculations, it is sometimes advisable to change the origin and scale with the substitutions $X = (x - A)/h$ and $Y = (y - B)/h$, where $A$ and $B$ are the assumed means (or middle values) of $x$ and $y$ series respectively and $h$ is the width of the interval.

**Example 5.7.** *Fit a second degree parabola to the following data :*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ = | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 | 3.5 | 4.0 |
| $y$ = | 1.1 | 1.3 | 1.6 | 2.0 | 2.7 | 3.4 | 4.1 |

(V.T.U., B.E., 2009)

**Sol.** We shift the origin to (2.5, 0) and take 0.5 as the new unit. This amounts to changing the variable $x$ to $X$, by the relation $X = 2x - 5$.

Let the parabola of fit be $y = a + bX + cX^2$.

The values of $\Sigma X$ etc. are calculated as below :

| $x$ | $X$ | $y$ | $Xy$ | $X^2$ | $X^2y$ | $X^3$ | $X^4$ |
|---|---|---|---|---|---|---|---|
| 1.0 | –3 | 1.1 | – 3.3 | 9 | 9.9 | – 27 | 81 |
| 1.5 | – 2 | 1.3 | – 2.6 | 4 | 5.2 | – 8 | 16 |
| 2.0 | – 1 | 1.6 | – 1.6 | 1 | 1.6 | – 1 | 1 |
| 2.5 | 0 | 2.0 | 0.0 | 0 | 0.0 | 0 | 0 |
| 3.0 | 1 | 2.7 | 2.7 | 1 | 2.7 | 1 | 1 |
| 3.5 | 2 | 3.4 | 6.8 | 4 | 13.6 | 8 | 16 |
| 4.0 | 3 | 4.1 | 12.3 | 9 | 36.9 | 27 | 81 |
| Total | 0 | 16.2 | 14.3 | 28 | 69.9 | 0 | 196 |

The normal equations are

$$7a + 28c = 16.2, \; 28b = 14.3, \; 28a + 196c = 69.9$$

Solving these as simultaneous equations, we get

$$a = 2.07, b = 0.511, c = 0.061.$$

$\therefore$ $$y = 2.07 + 0.511X + 0.061\,X^2$$

Replacing $X$ by $2x - 5$ in the above equation, we get

$$y = 2.07 + 0.511\,(2x - 5) + 0.061\,(2x - 5)^2$$

which simplifies to $\quad y = 1.04 - 0.193x + 0.244x^2$

This is the required parabola of best fit.

**Example 5.8.** *Fit a second degree parabola to the following data :*

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$ : | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
| $y$ : | 352 | 356 | 357 | 358 | 360 | 361 | 361 | 360 | 359 |

(U.P.T.U., 2001)

**Sol.** Taking $u = x - 1993$ and $v = y - 357$, the equation $y = a + bx + cx^2$ becomes

$$v = A + Bu + Cu^2 \qquad ...(i)$$

| $x$ | $u = x - 1993$ | $y$ | $v = y - 357$ | $uv$ | $u^2$ | $u^2v$ | $u^2$ | $u^4$ |
|---|---|---|---|---|---|---|---|---|
| 1989 | – 4 | 352 | – 5 | 20 | 16 | – 80 | – 64 | 256 |
| 1990 | – 3 | 356 | – 1 | 3 | 9 | – 9 | – 27 | 81 |
| 1991 | – 2 | 357 | 0 | 0 | 4 | 0 | – 8 | 16 |
| 1992 | – 1 | 358 | 1 | – 1 | 1 | 1 | – 1 | 1 |
| 1993 | 0 | 360 | 3 | 0 | 0 | 0 | 0 | 0 |
| 1994 | 1 | 361 | 4 | 4 | 1 | 4 | 1 | 1 |
| 1995 | 2 | 361 | 4 | 8 | 4 | 16 | 8 | 16 |
| 1996 | 3 | 360 | 3 | 9 | 9 | 27 | 27 | 81 |
| 1997 | 4 | 359 | 2 | 8 | 16 | 32 | 64 | 256 |
| Total | $\Sigma u = 0$ | | $\Sigma v = 11$ | $\Sigma uv = 51$ | $\Sigma u^2 = 60$ | $\Sigma u^2v = -9$ | $\Sigma u^3 = 0$ | $\Sigma u^4 = 708$ |



The normal equations are

$$\Sigma v = 9A + B\Sigma u + C\Sigma u^2 \quad \text{or} \quad 11 = 9A + 60C$$

$$\Sigma uv = A\Sigma u + B\Sigma u^2 + C\Sigma u^3 \quad \text{or} \quad 51 = 60B \quad \text{or} \quad B = \frac{17}{20}$$

$$\Sigma u^2v = A\Sigma u^2 + B\Sigma u^3 + C\Sigma u^4 \quad \text{or} \quad -9 = 60A + 708C$$

On solving these equations, we get $A = \frac{694}{231}, B = \frac{17}{20}, C = -\frac{247}{924}$

$\therefore$ (i) becomes $$v = \frac{694}{231} + \frac{17}{20}u - \frac{247}{924}u^2$$

or $$y - 357 = \frac{694}{231} + \frac{17}{20}(x - 1993) - \frac{247}{924}(x - 1993)^2$$

or $$y = \frac{694}{231} - \frac{33881}{20} - \frac{247}{924}(1993)^2 + \frac{17}{20}x + \frac{247 \times 3986}{924}x - \frac{247}{924}x^2$$

or $$y = 3 - 1694.05 - 1061792.32 + 357 + 0.85\,x + 1065.52x - 0.267x^2$$

Hence $$y = -\,1062526.37 + 1066.37x - 0.267x^2$$

## PROBLEMS 5.2

1. By the method of least squares, find the straight line that best fits the following data :

| | | | | | |
|---|---|---|---|---|---|
| $x$: | 1 | 2 | 3 | 4 | 5 |
| $y$: | 14 | 27 | 40 | 55 | 68 |

2. In some determinations of the value $v$ of carbon dioxide dissolved in a given volume of water at different temperatures $\theta$, the following pairs of values were obtained :

| | | | | |
|---|---|---|---|---|
| $\theta$ = | 0 | 5 | 10 | 15 |
| $v$ = | 1.80 | 1.45 | 1.18 | 1.00 |

Obtain by the method of least squares, a relation of the form $v = a + b\theta$ which best fits to these observations. (P.T.U., B. Tech., 2005)

3. Fit a straight line to the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| *Year x* | : | 1961 | 1971 | 1981 | 1991 | 2001 |
| *Production y* (in thousand tons) | : | 8 | 10 | 12 | 10 | 16 |

and find the expected production in 2006.

4. A simply supported beam carries a concentrated load $P$(lb) at its mid-point. Corresponding to various values of $P$, the maximum deflection $Y$ (in) is measured. The data are given below :

| | | | | | | |
|---|---|---|---|---|---|---|
| $P$: | 100 | 120 | 140 | 160 | 180 | 200 |
| $Y$: | 0.45 | 0.55 | 0.60 | 0.70 | 0.80 | 0.85 |

Find a law of the form $Y = a + bP$.

5. The result of measurement of electric resistance $R$ of a copper bar at various temperatures $t$° C are listed below :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t$ : | 19 | 25 | 30 | 36 | 40 | 45 | 50 |
| $R$: | 76 | 77 | 79 | 80 | 82 | 83 | 85 |

Find a relation $R = a + bt$ when $a$ and $b$ are constants to be determined by you.

6. A chemical company, wishing to study the effect of extraction time ($t$) on the efficiency of an extraction operation ($e$) obtained the data shown in the following table :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $t$: | 27 | 45 | 41 | 19 | 3 | 39 | 19 | 49 | 15 | 31 |
| $e$: | 57 | 64 | 80 | 46 | 62 | 72 | 52 | 77 | 57 | 68 |

Fit a straight line to the given data by the method of least squares.
(*J.N.T.U., B.Tech.*, 2004)

7. Find the parabola of the form $y = a + bx + cx^2$ which fits most closely with the observations :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$: | $-3$ | $-2$ | $-1$ | 0 | 1 | 2 | 3 |
| $y$: | 4.63 | 2.11 | 0.67 | 0.09 | 0.63 | 2.15 | 4.58 |

(*V.T.U., B.E.*, 2009)

8. By the method of least squares, fit a parabola of the form $y = a + bx + cx^2$, to the following data :

| | | | | | |
|---|---|---|---|---|---|
| $x$: | 2 | 4 | 6 | 8 | 10 |
| $y$: | 6.07 | 12.85 | 31.47 | 57.38 | 91.29 |

(*J.N.T.U., B.Tech.*, 2008)

9. Fit a parabola $y = a + bx + cx^2$ to the following data :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $y$: | 2 | 6 | 7 | 8 | 10 | 11 | 11 | 10 | 9 |

(*Mumbai, B.Tech.*, 2004)

10. The following table gives the results of the measurements of train resistance, $V$ is the velocity in miles per hour, $R$ is the resistance in pounds per ton :

| | | | | | | |
|---|---|---|---|---|---|---|
| $V$: | 20 | 40 | 60 | 80 | 100 | 120 |
| $R$: | 5.5 | 9.1 | 14.9 | 22.8 | 33.3 | 46.0 |

If $R$ is related to $V$ by the relation $R = a + bV + cV^2$, find $a$, $b$ and $c$.
(*U.P.T.U., B. Tech.*, 2002)

## 5.6 FITTING A CURVE OF THE TYPE

(1) $y = a + bx^2$ (2) $y = ax + bx^2$
(3) $y = ax + b/x$ (4) $ax^2 + b/x$.

**(1) $y = a + bx^2$**

Putting $x^2 = X$, we have $y = a + bX$ ...(i)

which is a linear equation. Its normal equations are

$$\Sigma y = na + b\Sigma X \; ; \; \Sigma yX = a\Sigma X + b\Sigma X^2$$

Solving these, we get $a$ and $b$. Substituting these values of $a$, $b$ and replacing $X$ by $x^2$ in (*i*), we obtain the desired equation of best fit.

**(2) $y = ax + bx^2$**

Rewriting this equation as $y/x = a + bx$ and putting $y/x = Y$, we have

$$Y = a + bx \quad ...(i)$$

Its normal equations are

$$\Sigma Y = na + b\Sigma x; \; \Sigma Yx = a\Sigma x + b\Sigma x^2$$

Solving these we get $a$ and $b$. Replacing $Y$ by $y/x$ in (*i*), we obtain the desired equation of best fit.



**(3) $y = ax + b/x$**

Rewriting this equation as $xy = ax^2 + b$ and putting

$x^2 = X$ and $xy = Y$, we have $Y = b + aX$ ...(i)

Its normal equations are

$$\Sigma Y = nb + a\Sigma X; \; \Sigma XY = b\Sigma X + a\Sigma X^2$$

Solving these equations, we get $a$ and $b$. Replacing $X$ by $x^2$ and $Y$ by $xy$ in (*i*), we obtain the desired equation of best fit.

**(4) $y = ax^2 + b/x$**

Rewriting this equation as $xy = ax^3 + b$ and putting $x^3 = X$ and $xy = Y$, we have

$$Y = b + aX \quad ...(i)$$

Its normal equations are

$$\Sigma Y = bn + a\Sigma X; \; \Sigma XY = b\Sigma X + X^2$$

Solving these equations, we get $a$ and $b$. Replacing $X$ by $x^3$ and $Y$ by $xy$, we obtain the desired equation of best fit.

**Example 5.9.** *Find the least squares fit of the form $y = a_0 + a_1x^2$ to the following data :*

| | | | | |
|---|---|---|---|---|
| $x$ : | $-1$ | 0 | 1 | 2 |
| $y$ : | 2 | 5 | 3 | 0 |

(*U.P.T.U., B.Tech.*, 2008)

**Sol.** Putting $x^2 = X$, we have $y = a_0 + a_1X$ ...(i)

$\therefore$ The normal equations are

$$\Sigma y = 4a + a_1\Sigma X; \; \Sigma Xy = a_0\Sigma X + a_1\Sigma X^2.$$

The values of $\Sigma X$, $\Sigma X^2$ etc. are calculated below :

| $x$ | $y$ | $X$ | $X^2$ | $XY$ |
|---|---|---|---|---|
| $-1$ | 2 | 1 | 1 | 2 |
| 0 | 5 | 0 | 0 | 0 |
| 1 | 3 | 1 | 1 | 3 |
| 2 | 0 | 4 | 16 | 0 |
| | $\Sigma y = 10$ | $\Sigma X = 6$ | $\Sigma X^2 = 18$ | $\Sigma XY = 5$ |

$\therefore$ The normal equations become $10 = 400 + 6a_1$; $5 = 600 + 18a_1$

Solving these equations we get, $a_0 = 4.167$, $a_1 = -1.111$.

Hence the curve of best fit is

$$y = 4.167 - 1.111X \quad i.e., \quad y = 4.167 - 1.111x^2.$$

**Example 5.10.** *Using the method of least squares, fit the curve $y = ax^2 + b/x$ to the following data :*

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | $-1.51$ | 0.99 | 8.88 | 7.66 |

**Sol.** Rewriting the given equation as $xy = ax^3 + b$ and putting $x^3 = X$ and $xy = Y$, we get

$$Y = aX + b \quad ...(i)$$

$\therefore$ The normal equations are

$$\Sigma Y = a\Sigma X + 4b \; ; \; \Sigma XY = a\Sigma X^2 + b\Sigma X$$

The values of $\Sigma X$, $\Sigma Y$ etc. are calculated below :

| $x$ | $y$ | $X = x^3$ | $Y = xy$ | $XY$ | $X^2$ |
|---|---|---|---|---|---|
| 1 | – 1.51 | 1 | – 1.51 | – 1.51 | 1 |
| 2 | 0.99 | 8 | 1.98 | 15.84 | 64 |
| 3 | 3.88 | 27 | 11.64 | 314.28 | 729 |
| 4 | 7.66 | 64 | 30.64 | 1960.96 | 4096 |
| | | $\Sigma X = 100$ | $\Sigma Y = 42.75$ | $\Sigma XY = 2289.57$ | $\Sigma X^2 = 4890$ |

$\therefore$ The normal equations become

$$42.75 = 100a + 4b$$
$$2289.57 = 4890a + 100b$$

Solving these equations, we get $a = 0.51$, $b = -2.06$

Hence the curve of best-fit is $Y = 0.51X + 2.06$

*i.e.*,
$$xy = 0.51x^3 - 2.06 \quad \text{or} \quad y = 0.51x^2 - \frac{2.06}{x}.$$

## 5.7. FITTING OF OTHER CURVES

**(1) $y = ax^b$**

Taking logarithms, $\log_{10} y = \log_{10} a + b \log_{10} x$

*i.e.* $\quad Y = A + bX \quad$ ...(*i*) where $X = \log_{10} x$, $Y = \log_{10} y$ and $A = \log_{10} a$.

$\therefore$ The normal equations for (*i*) are

$$\Sigma Y = nA + b\Sigma X, \; \Sigma XY = A\Sigma X + b\Sigma X^2$$

from which $A$ and $b$ can be determined. Then $a$ can be calculated from $A = \log_{10} a$.

**(2) $y = ae^{bx}$** *(Exponential curve)*

Taking logarithms, $\log_{10} y = \log_{10} a + bx \log_{10} e$

*i.e.* $\quad Y = A + Bx \quad$ where $Y = \log_{10} y$, $A = \log_{10} a$ and $B = b \log_{10} e$

Here the normal equations are $\Sigma Y = nA + B\Sigma x$, $\Sigma xY = A\Sigma x + B\Sigma x^2$

from which $A$, $B$ can be found and consequently $a$, $b$ can be calculated.

**(3) $xy^a = b$** *(or $pv^\gamma = k$)* *(Gas equation)*

Taking logarithms, $\log_{10} x + a \log_{10} y = \log_{10} b$

or
$$\log_{10} y = \frac{1}{a} \log_{10} b - \frac{1}{a} \log_{10} x.$$

This is of the form $Y = A + BX$ where $X = \log_{10} x$, $Y = \log_{10} y$, $A = \frac{1}{a} \log_{10} b$, $B = -\frac{1}{a}$.

Here also the problem reduces to finding a straight line of best fit through the given data.



**Example 5.11.** *An experiment gave the following values :*

| | | | | |
|---|---|---|---|---|
| *v (ft/min)* : | 350 | 400 | 500 | 600 |
| *t (min)* : | 61 | 26 | 7 | 2.6 |

*It is known that v and t are connected by the relation $v = at^b$. Find the best possible values of a and b.*

**Sol.** We have $\quad \log_{10} v = \log_{10} a + b \log_{10} t$

or $\quad Y = A + bX \quad$ where $X = \log_{10} t$, $Y = \log_{10} v$, $A = \log_{10} a$.

$\therefore$ The normal equations are

$$\Sigma Y = 4A + b\Sigma X \qquad ...(i)$$
$$\Sigma XY = A\Sigma X + b\Sigma X^2 \qquad ...(ii)$$

Now $\Sigma X$ etc. are calculated as in the following table :

| $v$ | $t$ | $X = \log_{10} t$ | $Y = \log_{10} v$ | $XY$ | $X^2$ |
|---|---|---|---|---|---|
| 350 | 61 | 1.7853 | 2.5441 | 4.542 | 3.187 |
| 400 | 26 | 1.4150 | 2.6021 | 3.682 | 2.002 |
| 500 | 7 | 0.8451 | 2.6990 | 2.281 | 0.714 |
| 600 | 2.6 | 0.4150 | 2.7782 | 1.153 | 0.172 |
| Total | | 4.4604 | 10.6234 | 11.658 | 6.075 |

$\therefore$ Equations (*i*) and (*ii*) become

$$4A + 4.46b = 10.623 \; ; \quad 4.46A + 6.075b = 11.658$$

Solving these, $A = 2.845$, $b = -0.1697$ $\therefore$ $a = \text{antilog } A = \text{antilog } 2.845 = 699.8$.

**Example 5.12.** *Predict the mean radiation dose at an altitude of 3000 feet by fitting an exponential curve to the given data :*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Altitude (x)* : | 50 | 450 | 780 | 1200 | 4400 | 4800 | 5300 |
| *Dose of radiation (y)* : | 28 | 30 | 32 | 36 | 51 | 58 | 69 |

*(J.N.T.U., B.E., 2003)*

**Sol.** Let $y = ab^x$ be the exponential curve.

Then $\quad \log_{10} y = \log_{10} a + x \log_{10} b$

or $\quad Y = A + Bx \quad$ where $Y = \log_{10} y$, $A = \log_{10} a$, $B = \log_{10} b$

$\therefore$ The normal equations are

$$\Sigma Y = 7A + B\,\Sigma x \qquad ...(i)$$
$$\Sigma x\, Y = A\Sigma x + B\,\Sigma x^2 \qquad ...(ii)$$

Now $\Sigma x$ etc are calculated as follows :

| $x$ | $y$ | $Y = \log_{10} y$ | $xY$ | $x^2$ |
|---|---|---|---|---|
| 50 | 28 | 1.447158 | 72.3579 | 2500 |
| 450 | 30 | 1.477121 | 664.7044 | 202500 |
| 780 | 32 | 1.505150 | 1174.0170 | 608400 |
| 1200 | 36 | 1.556303 | 1867.5636 | 1440000 |
| 4400 | 51 | 1.707570 | 7513.3080 | 19360000 |
| 4800 | 58 | 1.763428 | 8464.4544 | 23040000 |
| 5300 | 69 | 1.838849 | 9745.8997 | 28090000 |
| $\Sigma = 16980$ | | 11.295579 | 29502.305 | 72743400 |

$\therefore$ Equations (i) and (ii) become

$$11.295579 = 7A + 16980B$$
$$29502.305 = 16980A + 72743400B$$

Solving these equations, we get $A = 1.4521015,\ B = 0.0000666289$

$\therefore$ $\log_{10} y = Y = 1.4521015 + 0.0000666289\,x$

Hence $y(\text{at } x = 3000) = 44.874$ *i.e.* $44.9$ approx.

**Example 5.13.** *Fit a curve of the form $y = ae^{bx}$ to the following data :*

| | | | | |
|---|---|---|---|---|
| $x$ | 0 | 1 | 2 | 3 |
| $y$ | 1.05 | 2.10 | 3.[illegible] | 8.30 |

(J.N.T.U., B.Tech., 2009)

**Sol.** Taking logarithms of both sides, the given equation becomes

$$\log_{10} y = \log_{10} a + bx \log_{10} e$$

*i.e.*, $Y = A + bx$ where $Y = \log_{10} y,\ A = \log_{10} a,\ B = b \log_{10} e$

$\therefore$ The normal equations are

$$\Sigma Y = 4A + B\Sigma x;\ \Sigma xY = A\Sigma x + B\Sigma x^2.$$

Now $\Sigma x$, $\Sigma Y$ etc. are calculated as in the table below :

| $x$ | $y$ | $Y$ | $x^2$ | $xY$ |
|---|---|---|---|---|
| 0 | 1.05 | 0.0212 | 0 | 0 |
| 1 | 2.10 | 0.3222 | 1 | 0.3222 |
| 2 | 3.85 | 0.5855 | 4 | 1.1710 |
| 3 | 8.30 | 0.9191 | 9 | 2.7573 |
| $\Sigma x = 6$ | | $\Sigma Y = 1.8480$ | $\Sigma x^2 = 14$ | $\Sigma xY = 4.2505$ |

Substituting these values in the normal equations, we get

$4A + 6B = 1.848;\ 6A + 14B = 4.2505.$

Solving these equations, $A = 0.0185,\ B = 0.2956$

$\therefore$ $a = \text{antilog } A = 1.0186,\ b = B/\log_{10} e = 0.6806$

Hence the required curve of best fit is $y = 1.0186\, e^{0.6806x}$.

**Example 5.14.** *The pressure and volume of a gas are related by the equation $pV^{\gamma} = k$, $\gamma$ and $k$ being constants. Fit this equation to the following set of observations :*

| | | | | | | |
|---|---|---|---|---|---|---|
| $p$ (kg/cm$^2$) : | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
| $V$ (litres) : | 1.62 | 1.00 | 0.75 | 0.62 | 0.52 | 0.46 |

(Madras B.E., 2000 S)

**Sol.** We have $\log_{10} p + \gamma \log_{10} V = \log_{10} k$

or $$\log_{10} V = \frac{1}{\gamma}\log_{10} k - \frac{1}{\gamma}\log_{10} p$$

or $$Y = A + BX \quad \text{where } X = \log_{10} p,\ Y = \log_{10} V,\ A = \frac{1}{\gamma}\log_{10} k,\ B = -\frac{1}{\gamma}.$$



$\therefore$ The normal equations are

$$\Sigma Y = 6A + B\Sigma X \qquad \text{...(i)}$$
$$\Sigma XY = A\Sigma X + B\Sigma X^2 \qquad \text{...(ii)}$$

Now $\Sigma X$ etc. are calculated as follows :

| $p$ | $V$ | $X = \log_{10} p$ | $Y = \log_{10} V$ | $XY$ | $X^2$ |
|---|---|---|---|---|---|
| 0.5 | 1.62 | − 0.3010 | 0.2095 | − 0.0630 | 0.0906 |
| 1.0 | 1.00 | 0.0000 | 0.0000 | − 0.0000 | 0.0000 |
| 1.5 | 0.75 | 0.1761 | − 0.1249 | − 0.0220 | 0.0310 |
| 2.0 | 0.62 | 0.3010 | − 0.2076 | − 0.0625 | 0.0906 |
| 2.5 | 0.52 | 0.3979 | − 0.2840 | − 0.1130 | 0.1583 |
| 3.0 | 0.46 | 0.4771 | − 0.3372 | − 0.1609 | 0.2276 |
| Total | | 1.0511 | − 0.7442 | − 0.4214 | 0.5981 |

$\therefore$ Equations (i) and (ii) become

$$6A + 1.0511B = -\,0.7442\ ;\quad 1.0511A + 0.5981B = -\,0.4214$$

Solving these, we get $A = 0.0132,\ \ B = -\,0.7836.$

$\therefore$ $\gamma = -\,1/B = 1.276$

and $k = \text{antilog } (A\gamma) = \text{antilog } (0.0168) = 1.039.$

Hence the equation of best fit is $pV^{1.276} = 1.039.$

## 5.8 MOST PLAUSIBLE VALUES OF UNKNOWNS

Consider $m$ linear equations in $n$ unknowns :

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n &= k_1 \\ a_{12}x_1 + a_{22}x_2 + \ldots\ldots + x_{2n}x_n &= k_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n &= k_m \end{aligned}\right\} \qquad \text{...(1)}$$

When $m = n$, we can find a set of values of the unknowns uniquely.

When $m > n$, *i.e.*, the number of equations is greater than the number of unknowns, it may not be possible to find these values uniquely. Then we find those values of $x_1, x_2, \ldots\ldots x_n$ which satisfy (1) as nearly as possible. Applying the principle of least squares, these values can be obtained by minimising $E = \sum_{x=1}^{m} (a_{i1}x_1 + a_{i2}x_2 + \ldots\ldots + a_{in}x_n - k_i)^2$ using the conditions of minima *i.e.*,

$$\frac{\partial E}{\partial x_1} = 0,\ \frac{\partial E}{\partial x_2} = 0,\ \ldots\ldots,\ \frac{\partial E}{\partial x_n} = 0.$$

we get $n$ equations. Solving these equations, we get most plausible values of $x_1, x_2, \ldots\ldots x_n$.

**Example 5.15.** *Find the most plausible values of x, y and z from the equations $x - 3y - 3z = -\,14,\ 4x + y + 4z = 21,\ 3x + 2y - 5z = 5$ and $x - y + 2z = 3$, by forming the normal equations.*

**Sol.** Let $E = (x - 3y - 3z + 14)^2 + (4x + y + 4z - 21)^2 + (3x + 2y - 5z - 5)^2 + (x - y + 2z - 3)^2$

Scanned by CamScanner

The most plausible values of $x, y, z$ will be those which make $E$ minimum. These will be given by

$$\frac{\partial E}{\partial x}=0,\ \frac{\partial E}{\partial y}=0,\ \frac{\partial E}{\partial z}=0$$

$$\therefore\ \frac{\partial E}{\partial x}=2(x-3y-3z+14)+2(4x+y+4z-21)4+2(3x+2y-5z-5)3$$

$$+2(x-y+2z-3)=0 \text{ i.e., } 27x+6y=88 \quad \text{...(i)}$$

Similarly $\frac{\partial E}{\partial y}=0$ gives $6x+15y+z=70$ ...(ii)

and $\frac{\partial E}{\partial z}=0$ gives $y+54z=107$ ...(iii)

Solving (i), (ii) and (iii) we get the desired values $x = 2.47, y = 3.55, z = 1.92$.

## PROBLEMS 5.3

1. If V (km/hr) and R(kg/ton) are related by a relation of the type $R = a + bV^2$, find by the method of least squares $a$ and $b$ with the help of the following table :

| | | | | | |
|---|---|---|---|---|---|
| V = | 10 | 20 | 30 | 40 | 50 |
| R = | 8 | 10 | 15 | 21 | 30 |

(Indore, B.E., 2008)

2. Using the method of least squares fit the curve $y = ax + bx^2$ to following observations :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 |
| $y$ : | 1.8 | 5.1 | 8.9 | 14.1 | 19.8 |

3. Fit the curve $y = ax + b/x$ to the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $y$ : | 5.4 | 6.3 | 8.2 | 10.3 | 12.6 | 14.9 | 17.3 | 19.5 |

4. Estimate $y$ at $x = 2.25$ by fitting the *indifference curve* of the form $xy = Ax + B$ to the following data :

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 3 | 1.5 | 6 | 7.5 |

(J.N.T.U., B. Tech., 2003)

5. Fit a least square *geometric curve* $y = ax^b$ to the following data :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 |
| $y$ : | 0.5 | 2 | 4.5 | 8 | 12.5 |

6. Predict $y$ at $x = 3.75$, by fitting a *power curve* $y = ax^b$ to the given data :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 |
| $y$ : | 2.98 | 4.26 | 5.21 | 6.10 | 6.80 | 7.50 |

(J.N.T.U., B.E., 2003)

7. Fit the *exponential curve* $y = ae^{bx}$ to the following data :

| | | | | |
|---|---|---|---|---|
| $x$ : | 2 | 4 | 6 | 8 |
| $y$ : | 25 | 38 | 56 | 84 |

(Madras, B.E., 2003)

8. Fit the curve of the form $y = ae^{bx}$ to the following data :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 77 | 100 | 185 | 239 | 285 |
| $y$ : | 2.4 | 3.4 | 7.0 | 11.1 | 19.6 |

(J.N.T.U., B.Tech., 2006)



9. Obtain a relation of the form $y = kx^m$ for the following data by the method of least squares :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 |
| $y$ : | 7.1 | 27.8 | 62.1 | 110 | 161 |

10. Using method of least squares, fit a relation of the form $y = ab^x$ to the following data :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 2 | 3 | 4 | 5 | 6 |
| $y$ : | 144 | 172.8 | 207.4 | 248.8 | 298.5 |

(Tirchirapalli, B.E., 2001)

11. Growth of bacteria ($N$) in a culture after $t$ hrs. is given in the following table :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $t$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| $N$ : | 32 | 47 | 65 | 92 | 132 | 190 | 275 |

Fit a curve of the form $N = ab^t$ and estimate $N$ when $t = 7$.

12. The voltage $v$ across a capacitor at time $t$ seconds is given by the following table :

| | | | | | |
|---|---|---|---|---|---|
| $t$ : | 0 | 2 | 4 | 6 | 8 |
| $v$ : | 150 | 63 | 28 | 12 | 5.6 |

Use the method of least squares to fit a curve of the form $v = ae^{kt}$ to this data.

13. Find the most plausible values of $x$ and $y$ from the equations $x + 3y = 7.03$, $x + y = 3.01$, $2x - y = 0.03$, $3x + y = 4.97$, by forming the normal equations.

14. Obtain the most plausible values of $x, y$ and $z$ from the equations : $x + 2y + z = 1, -x + y + 2z = 3, 2x + y + z = 4, 4x + 2y - 5z = -7$

## 5.9. METHOD OF GROUP AVERAGES

Let the straight line $\quad y = a + bx$ ...(1)

fit the set of $n$ observations $(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$ quite closely. (Fig. 5.5).

When $x = x_1$, the *observed (or experimental) value of* $y = y_1 = L_1P_1$

and from (1), $\quad y = a + bx_1 = L_1M_1$,

which is known as the *expected* (or *calculated*) *value* of $y$ at $L_1$.

Then $\quad e_1$ = observed value at $L_1$ – expected value at $L_1$

$$= y_1 - (a + bx_1) = M_1P_1,$$

which is called the *error* (or *residual*) at $x_1$. Similarly the errors for the other observations are

$$e_2 = y_2 - (a + bx_2) = M_2P_2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$e_n = y_n - (a + bx_n) = M_nP_n$$

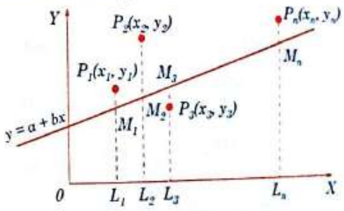


Fig. 5.5

Some of these errors may be positive and others negative.

The method of group averages is based on the assumption that the sum of the residuals is zero. To find the constants $a$ and $b$ in (1), we require two equations. As such we divide the data into two groups : the first containing $k$ observations $(x_1, y_1), (x_2, y_2) \ldots (x_k, y_k)$ ; and the second group having the remaining $n - k$ observations $(x_{k+1}, y_{k+1}), (x_{k+2}, y_{k+2}), \ldots\ldots, (x_n, y_n)$.

Assuming that the sum of the errors in each group is zero, we get

$$\{y_1-(a+bx_1)\}+\{y_2-(a+bx_2)\}+\ldots+\{y_k-(a+bx_k)\}=0$$
$$\{y_{k+1}-(a+bx_{k+1})\}+\{y_{k+2}-(a+bx_{k+2})\}+\ldots+\{y_n-(a+bx_n)\}=0$$

On simplification, we obtain

$$\frac{y_1+y_2+\ldots+y_k}{k}=a+b\,\frac{x_1+x_2+\ldots+x_k}{k} \quad \ldots(2)$$

$$\frac{y_{k+1}+y_{k+2}+\ldots+y_n}{n-k}=a+b\,\frac{x_{k+1}+x_{k+2}+\ldots+x_n}{n-k} \quad \ldots(3)$$

In (2), $\frac{1}{k}(x_1+x_2+\ldots+x_k)$ and $\frac{1}{k}(y_1+y_2+\ldots+y_k)$ are simply the average values of $x$'s and $y$'s of the first group. Hence the equations (2) and (3) are obtained from (1) by replacing $x$ and $y$ by their respective averages of the two groups. Solving (2) and (3), we get $a$ and $b$.

**Obs.** The main drawback of this method is that a different grouping of the observations will give different values of $a$ and $b$. In practice, we divide the data in such a way that each group contains almost an equal number of observations.

**Example 5.16.** *The latent heat of vaporisation of steam r, is given in the following table at different temperatures t :*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $t$: | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 |
| $r$: | 1069.1 | 1063.6 | 1058.2 | 1052.7 | 1049.3 | 1041.8 | 1036.3 | 1030.8 |

*For this range of temperature, a relation of the form $r = a + bt$ is known to fit the data. Find the values of a and b by the method of group averages.*

**Sol.** Let us divide the data into two groups each containing four readings. Then we have

| $t$ | $r$ | $t$ | $r$ |
|---|---|---|---|
| 40 | 1069.1 | 80 | 1049.3 |
| 50 | 1063.6 | 90 | 1041.8 |
| 60 | 1058.2 | 100 | 1036.3 |
| 70 | 1052.7 | 110 | 1030.8 |
| $\Sigma t = 220$ | $\Sigma r = 4243.6$ | $\Sigma t = 380$ | $\Sigma r = 4158.2$ |

Substituting the averages of $t$'s and $r$'s of the two groups in the given relation, we get

$$\frac{4243.6}{4}=a+b\,\frac{220}{4} \quad i.e. \quad 1060.9=a+55b \quad \ldots(i)$$

$$\frac{4158.2}{4}=a+b\,\frac{380}{4} \quad i.e. \quad 1039.55=a+95b \quad \ldots(ii)$$

Solving (i) and (ii), we obtain

$$a = 1090.26,\ b = -0.534.$$

**Example 5.17.** *The observations in the following table fit a law of the form $y = ax^n$. Estimate a and n by the method of group averages.*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$: | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| $y$: | 1.06 | 1.33 | 1.52 | 1.68 | 1.81 | 1.91 | 2.01 | 2.11 |



**Sol.** We have $\quad y = ax^n$

Taking logarithms, $\log_{10} y = \log_{10} a + n \log_{10} x$

*i.e.* $\quad Y = A + nX \quad \ldots(i)$

where $\quad X = \log_{10} x,\ Y = \log_{10} y,\ A = \log_{10} a$

Divide the data into two groups each containing four pairs of values, so that

| $x$ | $y$ | $X = \log_{10} x$ | $Y = \log_{10} y$ |
|---|---|---|---|
| 10 | 1.06 | 1.0253 | 0.0253 |
| 20 | 1.33 | 1.3010 | 0.1238 |
| 30 | 1.52 | 1.4771 | 0.1818 |
| 40 | 1.68 | 1.6021 | 0.2253 |
| | | $\Sigma X = 5.4055$ | $\Sigma Y = 0.5562$ |

| $x$ | $y$ | $X = \log_{10} x$ | $Y = \log_{10} y$ |
|---|---|---|---|
| 50 | 1.81 | 1.6990 | 0.2577 |
| 60 | 1.91 | 1.7782 | 0.2810 |
| 70 | 2.01 | 1.8451 | 0.3032 |
| 80 | 2.11 | 1.9031 | 0.3243 |
| | | $\Sigma X = 7.2254$ | $\Sigma Y = 1.1662$ |

Substituting the averages of $X$'s and $Y$'s of the two groups in (i), we get

$$\frac{0.5562}{4}=A+n\,\frac{5.4055}{4} \quad i.e. \quad 0.1390 = A + 1.3514\,n \quad \ldots(ii)$$

$$\frac{1.1662}{4}=A+n\,\frac{7.2254}{4} \quad i.e. \quad 0.2916 = A + 1.8064\,n \quad \ldots(iii)$$

(iii) – (ii) gives $0.1526 = 0.455\,n$ *i.e.* $n = 0.3354$

From (ii), $\quad A = -0.3142$ *i.e.* $\log_{10} a = -0.3142$

whence $\quad a = \text{antilog}(-0.3142) = 0.4851$

## 5.10. LAWS CONTAINING THREE CONSTANTS

We have so far applied the above method to fit the data to laws involving two constants only. But at times we come across laws of the form

$$y = a + bx + cx^2,\ y = a + bx^c \text{ and } y = a + be^{cx}$$

each of which contains three constants. To fit such laws to a set of observations, we device the following procedures to reduce these to laws hitherto discussed.

**(1) Equation $y = a + bx + cx^2$**

Let $(x_1, y_1)$ be a point on the curve satisfying the given data so that

$$y_1 = a + bx_1 + cx_1^2$$

Then $\quad y - y_1 = b(x - x_1) + c(x^2 - x_1^2)$

or $\quad \frac{y-y_1}{x-x_1} = b + c(x + x_1)$

Putting $x + x_1 = X$ and $(y - y_1)/(x - x_1) = Y$, it takes the linear form $Y = b + cX$.

Now $b$ and $c$ can be found by the graphical method or the method of averages.

Scanned by CamScanner

**(2) Equation $y = a + bx^c$**

It can be rewritten as $y - a = bx^c$ ...(1)

To find $a$, let $(x_1, y_1)$, $(x_2, y_2)$, $(x_3, y_3)$ be three particular points on the curve (1) such that $x_1, x_2, x_3$ are in geometric progression

i.e. $$x_1 x_3 = x_2^2 \quad ...(2)$$

Then $$y_1 - a = bx_1^c$$
$$y_2 - a = bx_2^c$$
and $$y_3 - a = bx_3^c$$

$$\therefore \quad (y_1 - a)(y_3 - a) = b^2(x_1x_3)^c = b^2(x_2^2)^c \quad \text{[by (2)]}$$
$$= (bx_2^c)^2 = (y_2 - a)^2$$

or $$a(y_1 + y_3 - 2y_2) = y_1y_3 - y_2^2$$

which gives $a$. Now (1) reduces to a law containing two constants $b$ and $c$ only.

Taking logarithms, (1) becomes
$$\log_{10}(y - a) = \log_{10} b + c \log_{10} x$$
or $$Y = B + cX \quad ...(3)$$
where $X = \log_{10} x$, $Y = \log_{10}(y - a)$, $B = \log_{10} b$.

Hence we can find $b$ and $c$ as before from (3).

**(3) Equation $y = a + be^{cx}$**

It can be written as $y - a = be^{cx}$ ...(1)

To find $a$, let $(x_1, y_1)$, $(x_2, y_2)$, $(x_3, y_3)$ be three particular points on the curve (1) such that $x_1, x_2, x_3$ are in arithmetic progression i.e. $x_1 + x_3 = 2x_2$ ...(2)

Then $$y_1 - a = be^{cx_1}, \quad y_2 - a = be^{cx_2} \quad \text{and} \quad y_3 - a = be^{cx_3}$$

$$\therefore \quad (y_1 - a)(y_3 - a) = b^2 e^{c(x_1 + x_3)} = (be^{cx_2})^2 = (y_2 - a)^2 \quad \text{[by (2)]}$$

or $$a(y_1 + y_3 - 2y_2) = y_1y_3 - y_2^2$$

which gives $a$. Now (1) reduces to a law containing two constants $b$ and $c$ only.

Taking logarithms, (1) becomes
$$\log_{10}(y - a) = \log_{10} b + cx \log_{10} e$$
or $$Y = B + Cx \quad ...(3)$$
where $Y = \log_{10}(y - a)$, $B = \log_{10} b$, $C = c \log_{10} e$.

Hence we can find $b$ and $c$ as before from (3).

**Example 5.18.** *The corresponding values of x and y are given by the following table:*

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$: | 87.5 | 84.0 | 77.8 | 63.7 | 46.7 | 36.9 |
| $y$: | 292 | 283 | 270 | 235 | 197 | 181 |

*Fit a parabola of the form $y = a + bx + cx^2$, by the method of group averages.*

**Sol.** Taking $x = 84$, $y = 283$ as a particular point on $y = a + bx + cx^2$, we get
$$283 = a + b(84) + c(84)^2 \quad ...(i)$$
$$y - 283 = b(x - 84) + c[x^2 - (84)^2]$$
or $$\frac{y - 283}{x - 84} = b + c(x + 84) \quad \text{i.e.} \quad Y = b + cX \quad ...(ii)$$
where $X = x + 84$, $Y = (y - 283)/(x - 84)$.



Now we have the following table of values :

| $x$ | $y$ | $X = x + 84$ | $Y = (y - 283)/(x - 84)$ |
|---|---|---|---|
| 87.5 | 292 | 171.5 | 2.571 |
| 84.0 | 283 | — | — |
| 77.8 | 270 | 161.8 | 2.097 |
| | | $\Sigma X = 333.3$ | $\Sigma Y = 4.668$ |
| 63.7 | 235 | 147.7 | 2.364 |
| 46.7 | 197 | 130.7 | 2.306 |
| 36.9 | 181 | 120.9 | 2.166 |
| | | $\Sigma X = 399.3$ | $\Sigma Y = 6.836$ |

Substituting the averages of $X$ and $Y$ in (ii), we get
$$\frac{4.668}{2} = b + c\,\frac{333.3}{2} \quad \text{i.e.} \quad 2.33 = b + 166.65\,c \quad ...(iii)$$
$$\frac{6.836}{3} = b + c\,\frac{399.3}{3} \quad \text{i.e.} \quad 2.28 = b + 131.1\,c \quad ...(iv)$$

(iv) – (iii) gives $c = 0.0014$

and (iii) gives $b = 2.0967$ i.e. 2.1 nearly

From (i), we get $a = 96.9988$ i.e. 97 nearly.

Hence the parabola of fit is $y = 97 + 2.1x + 0.0014x^2$.

**Example 5.19.** *The train resistance R (lbs/ton) is measured for the following values of its velocity V (km/hr) :*

| | | | | | |
|---|---|---|---|---|---|
| $V$: | 20 | 40 | 60 | 80 | 100 |
| $R$: | 5 | 9 | 14 | 25 | 36 |

*If R is related to V by the formula $R = a + bV^n$, find a, b and n.*

**Sol.** To find $a$, we take the following three values of $v$ which are in G.P. :

Then $$v_1 = 20, \quad v_2 = 40, \quad v_3 = 80$$
$$R_1 = 5, \quad R_2 = 9, \; R_3 = 25$$
$\therefore$ $$(R_1 - a)(R_3 - a) = (R_2 - a)^2$$
whence $$a = \frac{R_1R_3 - R_2^2}{R_1 + R_3 - 2R_2} = 3.67$$

Thus $$R - 3.67 = bV^n \quad \text{or} \quad \log_{10}(R - 3.67) = \log_{10} b + n \log_{10} V$$
i.e. $$Y = k + nX \quad ...(i)$$
where $X = \log_{10} V$, $Y = \log_{10}(R - 3.67)$, $k = \log_{10} b$.

Now we have the following table of values :

| $V$ | $R$ | $X = \log_{10} V$ | $Y = \log_{10}(R - 3.67)$ |
|---|---|---|---|
| 20 | 5 | 1.3010 | 0.1238 |
| 40 | 9 | 1.6021 | 0.7267 |
| 60 | 14 | 1.7782 | 1.0141 |
| | | $\Sigma X = 4.6813$ | $\Sigma Y = 1.8646$ |
| 80 | 25 | 1.9031 | 1.3290 |
| 100 | 36 | 2.0000 | 1.5096 |
| | | $\Sigma X = 3.9031$ | $\Sigma Y = 2.8386$ |

Substituting the averages of $X$'s and $Y$'s in (1), we obtain

$$\frac{1.8646}{2} = k + n\,\frac{4.6813}{2} \quad i.e. \quad 0.6215 = k + 1.5604\,n \qquad ...(ii)$$

$$\frac{2.8386}{2} = k + n\,\frac{3.9031}{2} \quad i.e. \quad 1.4193 = k + 1.9516\,n \qquad ...(iii)$$

Solving (*ii*) and (*iii*), we get $n = 2.04$, $k = -2.56$ approx.

$\therefore$ $b = \text{antilog } k = \text{antilog } (-2.56) = 0.0028.$

## PROBLEMS 5.4

1. Fit a straight line of the form $y = a + bx$ to the following data by the method of group averages :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 0 | 5 | 10 | 15 | 20 | 25 |
| $y$ : | 12 | 15 | 17 | 22 | 24 | 30 |

(*Tiruchirapalli, B.E., 2001*)

2. Apply the method of group averages to work out Prob. 1 of page 122.
3. The weights of a calf taken at weekly intervals are given below :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Age : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Weight : | 52.5 | 58.7 | 65.0 | 70.2 | 75.4 | 81.1 | 87.2 | 95.5 | 102.2 | 108.4 |

Find a straight line of best fit.

4. Work out Example 5.1 page 119, by the method of group averages.
5. The head of water $H$ (*ft*) and the quantity of water $Q(\text{ft}^3)$ flowing per second are related by the law $Q = CH^n$. Find the best values of $C$ and $n$ by the method of group averages for the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| $H$ : | 1.2 | 1.4 | 1.6 | 1.8 | 2.0 | 2.4 |
| $Q$ : | 4.2 | 6.1 | 8.5 | 11.5 | 14.9 | 23.5. |

6. Using the method of averages, fit a parabola $y = ax^2 + bx + c$ to the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 20 | 40 | 60 | 80 | 100 | 120 |
| $y$ : | 5.5 | 9.1 | 14.9 | 22.8 | 33.3 | 46.0 |

7. While testing a centrifugal pump, the following data is obtained. It is assumed to fit the equation $y = a + bx + cx^2$, where $x$ is the discharge in litre/sec and $y$, head in metres of water. Find the values of the constants $a$, $b$, $c$ by the method of group averages.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $x$ : | 2 | 2.5 | 3 | 3.5 | 4 | 4.5 | 5 | 5.5 | 6 |
| $y$ : | 18 | 17.8 | 17.5 | 17 | 15.8 | 14.8 | 13.3 | 11.7 | 9 |

8. By the method of averages, fit a curve of the form $y = ae^{bx}$ to the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 5 | 15 | 20 | 30 | 35 | 40 |
| $y$ : | 10 | 14 | 25 | 40 | 50 | 62 |

(*Madras, B.E., 2002*)

9. In an experiment, the voltage $v$ is observed for the following values of the current $i$ :

| | | | | | | |
|---|---|---|---|---|---|---|
| $i$ : | 0.5 | 1 | 2 | 4 | 8 | 12 |
| $v$ : | 160 | 120 | 94 | 75 | 62 | 56 |

If $v$ and $i$ are connected by the relation $v = a + bi^k$, find $a$, $b$ and $k$.

10. The variables $s$ and $t$ are connected by the relation $s = a + be^{nt}$ and their corresponding values are given in the following table :

| | | | | | |
|---|---|---|---|---|---|
| $t$ : | 1 | 2 | 6 | 8 | 11 |
| $s$ : | 12.7 | 12.5 | 11.6 | 11.3 | 11 |

Find the best possible values of $a$, $b$ and $n$.



## 5.11. METHOD OF MOMENTS

Let $(x_1, y_1), (x_2, y_2), \ldots\ldots, (x_n, y_n)$ be the set of $n$ observations such that

$$x_2 - x_1 = x_3 - x_2 = \ldots\ldots = x_n - x_{n-1} = h \text{ (say)}$$

We define the *moments of the observed values* of $y$ as follows :

$m_1$, the 1st moment $= h\Sigma y$

$m_2$, the 2nd moment $= h\Sigma xy$

$m_3$, the third moment $= h\Sigma x^2 y$ and so on.

Let the curve fitting the given data be $y = f(x)$. Then the *moments of the calculated values* of $y$ are

$\mu_1$, the 1st moment $= \int y dx$

$\mu_2$, the 2nd moment $= \int xy\, dx$

$\mu_3$, the 3rd moment $= \int x^2 y\, dx$ and so on.

*This method is based on the assumption that the moments of the observed values of v are respectively equal to the moments of the calculated values of y i.e.,* $m_1 = \mu_1$, $m_2 = \mu_2$, $m_3 = \mu_3$ etc. These equations (known as *observation equations*) are used to determine the constants in $f(x)$.

$m$'s are calculated from the tabulated values of $x$ and $y$ while $\mu$'s are computed as follows :

In Fig. 5.6, $y_1$ the ordinate of $P_1(x = x_1)$, can be taken as the value of $y$ at the mid-point of the interval $(x_1 - h/2, x_1 + h/2)$. Similarly $y_n$, the ordinate of $P_n(x = x_n)$, can be taken as the value of $y$ at the mid-point of the interval $(x_n - h/2, x_n + h/2)$. If $A$ and $B$ be the points such that

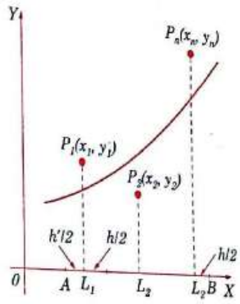


Fig. 5.6

$$OA = x_1 - h/2 \quad \text{and} \quad OB = x_n + h/2,$$

then
$$\mu_1 = \int y\, dx = \int_{x_1 - h/2}^{x_n + h/2} f(x)\, dx$$

$$\mu_2 = \int_{x_1 - h/2}^{x_n + h/2} xf(x)\, dx\ ; \qquad \mu_3 = \int_{x_1 - h/2}^{x_n + h/2} x^2 f(x)\, dx \text{ and so on.}$$

**Example 5.20.** *Fit a straight line $y = a + bx$ to the following data by the method of moments :*

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 16 | 19 | 23 | 26 |

(*Madras, B.E., 2001 S*)

**Sol.** Since only two constants $a$ and $b$ are to be found, it is sufficient to calculate the first two moments in each case. Here $h = 1$.

$$m_1 = h\Sigma y = 1(16 + 19 + 23 + 26) = 84$$

$$m_2 = h\Sigma xy = 1(1 \times 16 + 2 \times 19 + 3 \times 22 + 4 \times 26) = 227.$$

To compute the moments of calculated values of $y = a + bx$, the limits of integration will be $1 - h/2$ and $4 + h/2$ *i.e.*, 0.5 and 4.5.

$$\therefore \quad \mu_1 = \int_{0.5}^{4.5} (a + bx)\, dx = \left| ax + b\frac{x^2}{2} \right|_{0.5}^{4.5} = 4a + 10b$$

$$\mu_2 = \int_{0.5}^{4.5} x(a + bx)\, dx = 10a + \frac{91}{3} b.$$

Thus, the observation equations $m_r = \eta_r$ $(r = 1, 2)$ are

$$4a + 10b = 84 \ ; \quad 10a + \frac{91}{3} b = 227$$

Solving these, $a = 13.02$ and $b = 3.19$.

Hence the required equation is $y = 13.02 + 3.19x$.

**Example 5.21.** *Given the following data :*

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 |
| $y$ : | 1 | 5 | 10 | 22 | 38 |

*find the parabola of best fit by the method of moments.*

**Sol.** Let the parabola of best fit be $y = a + bx + cx^2$ ...(i)

Since three constants are to be found, we calculate the first three moments in each case. Here $h = 1$.

$$m_1 = h\Sigma y = 1(1 + 5 + 10 + 22 + 38) = 76$$
$$m_2 = h\Sigma xy = 1(0 + 5 + 20 + 66 + 152) = 243$$
$$m_3 = h\Sigma x^2 y = 1(0 + 5 + 40 + 198 + 608) = 851$$

For computing the moments of calculated values of (i), the limits of integration will be $0 - h/2$ and $4 + h/2$ *i.e.* $-0.5$ and $4.5$.

$$\therefore \quad \mu_1 = \int_{-0.5}^{4.5} (a + bx + cx^2)\, dx = 5a + 10b + 30.4c$$

$$\mu_2 = \int_{-0.5}^{4.5} x(a + bx + cx^2)\, dx = 10a + 30.4b + 102.5c$$

$$\mu_3 = \int_{-0.5}^{4.5} x^2(a + bx + cx^2)\, dx = 30.4a + 102.5b + 369.1c$$

Thus the observation equations $m_r = \mu_r$ $(r = 1, 2, 3)$ are

$$5a + 10b + 30.4c = 76$$
$$10a + 30.4b + 102.5c = 243$$
$$30.4a + 102.5b + 369.1c = 851$$

Solving these equations, we get $a = 0.4, b = 3.15, c = 1.4$.

Hence the parabola of best fit is $y = 0.4 + 3.15x + 1.4x^2$.

## PROBLEMS 5.5

1. Use the method of moments to fit the straight line $y = a + bx$ to the data :

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 0.17 | 0.18 | 0.23 | 0.32 |

2. Fit a straight line to the following data, using the method of moments :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 3 | 5 | 7 | 9 |
| $y$ : | 1.5 | 2.8 | 4.0 | 4.7 | 6.0 |

(*Madras, B.E.*, 2001)



3. Fit a parabola of the form $y = a + bx + cx^2$ to the data :

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 1.7 | 1.8 | 2.3 | 3.2 |

by the method of moments.

4. By using the method of moments, fit a parabola to the following data .

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 0.30 | 0.64 | 1.32 | 5.40 |

(*Madras*, B.E., 2000 S)

## 5.12. OBJECTIVE TYPE OF QUESTIONS

## PROBLEMS 5.6

*Select the correct answer or fill up the blanks in the following questions :*

1. The *method of group averages* is based on the assumption that the sum of the residuals is ......
2. $y = ax^b + c$ in linear form is ........
3. To fit the straight line $y = mx + c$ to $n$ observations, the normal equations are
   (*i*) $\Sigma y = n\Sigma x + \Sigma cm$, $\Sigma xy = c\,\Sigma x^2 + c\Sigma n$.
   (*ii*) $\Sigma y = m\Sigma x + nc$, $\Sigma xy = m\Sigma x^2 + c\Sigma x$.
   (*iii*) $\Sigma y = c\Sigma x + m\Sigma n$, $\Sigma xy = c\Sigma x^2 + m\Sigma x$.
4. To fit $y = ab^x$ by least square method, normal equations are ......
5. The observation equations for fitting a straight line by *method of moments* are ......
6. The principle of '*least squares*' states that ......
7. $y = ax^2 + b\log_{10} x$ reduced to linear law takes the form ...... .
8. Given $\begin{bmatrix} x: & 0 & 1 & 2 \\ y: & 0 & 1.1 & 2.1 \end{bmatrix}$, then the straight line of best fit is ...... .
9. The *method of moments* is based on the assumption that ...... .
10. In $y = a + bx$, $\Sigma x = 50$, $\Sigma y = 80$, $\Sigma xy = 1030$, $\Sigma x^2 = 750$ and $n = 10$, then $a =$ ......, $b =$ ....... .
11. The gas equation $pv^r = k$ can be reduced to $y = a + bx$ where $a =$ ....... $b =$ .......
12. $y = \dfrac{x}{ax + b}$ in linear form is ..... .
13. If $y = ke^{mx}$, then the first normal equation is $\Sigma\log_{10} y =$ ......
    (*a*) $kn + m\Sigma x$ (*b*) $k\Sigma x + m\Sigma x^2$
    (*c*) $n\log_{10} k + m\log_{10} e\Sigma x$ (*d*) $k\Sigma\log_{10} y + m\Sigma x$.
14. If $y = a + bx + cx^2$ and

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 |
| $y$ : | 1 | 1.8 | 1.3 | 2.5 | 7.3 |

, then the first normal equation is
    (*a*) $15 = 5a + 10b + 29c$ (*b*) $15 = 5a + 10b + 31c$
    (*c*) $12.9 = 5a + 10b + 30c$ (*d*) $34 = 5a + 10b + 27c$.

**15.** If $y = 2x + 5$ is the best fit for 8 pairs of values of $(x, y)$ by the method of least squares, and $\Sigma y = 120$, then $\Sigma x = \ldots\ldots$

(a) 35 (b) 40
(c) 45 (d) 30.

**16.** If $y = a + bx^2$ and $n$ is the number of observations, then the first normal equations is $\Sigma y = \ldots\ldots$

(a) $na + b\Sigma x^2$ (b) $na\ \Sigma x + b\Sigma x^2$
(c) $na + b\Sigma x^3$ (d) $na\ \Sigma y + b\Sigma yx^2$.

# 6

# Finite Differences



## 6.1. INTRODUCTION

*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 between the values assumed by the function, whenever the independent variable changes by finite jumps whether equal or unequal. On the other hand, in infinitesimal calculus, we study those changes of the function which occur when the independent variable changes continuously in a given interval. In this chapter, we shall study the variations in the function when the independent variable changes by equal intervals.

## 6.2. 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 $y = y_0, y_1, y_2, \ldots\ldots, y_n$. To determine the values of $f(x)$ or $f'(x)$ for some intermediate values of $x$, the following three types of differences are found useful :

**(1) Forward differences.** The differences $y_1 - y_0, y_2 - y_1, \ldots\ldots, y_n - y_{n-1}$ when denoted by $\Delta y_0, \Delta y_1, \ldots\ldots, \Delta y_{n-1}$ respectively are called the *first forward differences* where $\Delta$ is the *forward difference operator*. Thus the first forward differences are $\Delta y_r = y_{r+1} - y_r$.

Similarly the second forward differences are defined by $\Delta^2 y_r = \Delta y_{r+1} - \Delta y_r$

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 in Table 6.1.

In a difference table, $x$ is called the *argument* and $y$ the *function* or the *entry*. $y_0$, the first entry is called the *leading term* and $\Delta y_0, \Delta^2 y_0, \Delta^3 y_0$ etc. are called the *leading differences*.

**Table 6.1. Forward Difference Table**

| *Value of x* | *Value of y* | *1st diff.* | *2nd diff.* | *3rd diff.* | *4th diff.* | *5th diff.* |
|---|---|---|---|---|---|---|
| $x_0$ | $y_0$ | | | | | |
| | | $\Delta y_0$ | | | | |
| $x_0+h$ | $y_1$ | | $\Delta^2 y_0$ | | | |
| | | $\Delta y_1$ | | $\Delta^3 y_0$ | | |
| $x_0+2h$ | $y_2$ | | $\Delta^2 y_1$ | | $\Delta^4 y_0$ | |
| | | $\Delta y_2$ | | $\Delta^3 y_1$ | | $\Delta^5 y_0$ |
| $x_0+3h$ | $y_3$ | | $\Delta^2 y_2$ | | $\Delta^4 y_1$ | |
| | | $\Delta y_3$ | | $\Delta^3 y_2$ | | |
| $x_0+4h$ | $y_4$ | | $\Delta^2 y_3$ | | | |
| | | $\Delta y_4$ | | | | |
| $x_0+5h$ | $y_5$ | | | | | |

**Obs. 1.** *Any higher order forward difference can be expressed in terms of the entries.*

We have
$$\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = (y_2 - y_1) - (y_1 - y_0) = y_2 - 2y_1 + y_0$$
$$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = (y_3 - 2y_2 + y_1) - (y_2 - 2y_1 + y_0) = y_3 - 3y_2 + 3y_1 - y_0$$
$$\Delta^4 y_0 = \Delta^3 y_1 - \Delta^3 y_0 = (y_4 - 3y_3 + 3y_2 - y_1) - (y_3 - 3y_2 + 3y_1 - y_0)$$
$$= y_4 - 4y_3 + 6y_2 - 4y_1 + y_0$$

The coefficients occurring on the right hand side being the binomial coefficients, we have in general, $\Delta^n y_0 = y_n - {}^n c_1 y_{n-1} + {}^n c_2 y_{n-2} - \ldots\ldots + (-1)^n y_0$.

**Obs. 2.** *The operator $\Delta$ obeys the distributive, commutative and index laws*

*i.e.* (*i*) $\Delta[f(x) \pm \phi(x)] = \Delta f(x) \pm \Delta\phi(x)$

(*ii*) $\Delta[cf(x)] = c\,\Delta f(x)$, $c$ being a constant.

(*iii*) $\Delta^m \Delta^n f(x) = \Delta^{m+n} f(x)$, $m$ and $n$ being positive integers. In view of (*i*) and (*ii*), $\Delta$ is a *linear operator*.

But $\Delta[f(x) \,.\, \phi(x)] \neq f(x) \,.\, \Delta\phi(x)$.

(2) **Backward differences.** The differences $y_1 - y_0, y_2 - y_1, \ldots\ldots, y_n - y_{n-1}$ when denoted by $\nabla y_1, \nabla y_2, \ldots\ldots, \nabla y_n$ respectively, are called the *first backward differences* where $\nabla$

**Table 6.2. Backward Difference Table**

| *Value of x* | *Value of y* | *1st diff.* | *2nd diff.* | *3rd diff.* | *4th diff.* | *5th diff.* |
|---|---|---|---|---|---|---|
| $x_0$ | $y_0$ | | | | | |
| | | $\nabla y_1$ | | | | |
| $x_0+h$ | $y_1$ | | $\nabla^2 y_2$ | | | |
| | | $\nabla y_2$ | | $\nabla^3 y_3$ | | |
| $x_0+2h$ | $y_2$ | | $\nabla^2 y_3$ | | $\nabla^4 y_4$ | |
| | | $\nabla y_3$ | | $\nabla^3 y_4$ | | $\nabla^5 y_5$ |
| $x_0+3h$ | $y_3$ | | $\nabla^2 y_4$ | | $\nabla^4 y_5$ | |
| | | $\nabla y_4$ | | $\nabla^3 y_5$ | | |
| $x_0+4h$ | $y_4$ | | $\nabla^2 y_5$ | | | |
| | | $\nabla y_5$ | | | | |
| $x_0+5h$ | $y_5$ | | | | | |



is the *backward difference operator*. Similarly we define higher order backward differences. Thus we have $\nabla y_r = y_r - y_{r-1}$, $\nabla^2 y_r = \nabla y_r - \nabla y_{r-1}$, $\nabla^3 y_r = \nabla^2 y_r - \nabla^2{}_{r-1}$, etc.

These differences are exhibited in the Table 6.2.

(3) **Central differences.** Sometimes it is convenient to employ another system of differences known as *central differences*. In this system, the *central difference operator* $\delta$ is defined by the relations : $y_1 - y_0 = \delta y_{1/2}, y_2 - y_1 = \delta y_{3/2}, \ldots\ldots, y_n - y_{n-1} = \delta y_{n-1/2}$

Similarly, higher order central differences are defined as
$$\delta y_{3/2} - \delta y_{1/2} = \delta^2 y_1,\ \delta y_{5/2} - \delta y_{3/2} = \delta^2 y_2,\ \ldots\ldots,\ \delta^2 y_2 - \delta^2 y_1 = \delta^3 y_{3/2} \text{ and so on.}$$

These differences are shown in Table 6.3.

**Table 6.3. Central Difference Table**

| *Value of x* | *Value of y* | *1st diff.* | *2nd diff.* | *3rd diff.* | *4th diff.* | *5th diff.* |
|---|---|---|---|---|---|---|
| $x_0$ | $y_0$ | | | | | |
| | | $\delta y_{1/2}$ | | | | |
| $x_0+h$ | $y_1$ | | $\delta^2 y_1$ | | | |
| | | $\delta y_{3/2}$ | | $\delta^3 y_{3/2}$ | | |
| $x_0+2h$ | $y_2$ | | $\delta^2 y_2$ | | $\delta^4 y_2$ | |
| | | $\delta y_{5/2}$ | | $\delta^3 y_{5/2}$ | | $\delta^5 y_{5/2}$ |
| $x_0+3h$ | $y_3$ | | $\delta^2 y_3$ | | $\delta^4 y_3$ | |
| | | $\delta y_{7/2}$ | | $\delta^3 y_{7/2}$ | | |
| $x_0+4h$ | $y_4$ | | $\delta^2 y_4$ | | | |
| | | $\delta y_{9/2}$ | | | | |
| $x_0+5h$ | $y_5$ | | | | | |

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 and those of even order for only integral values of the suffix.

It is often required to find the mean of adjacent values in the same column of differences. We denote this mean by $\mu$. Thus $\mu\delta y_1 = \frac{1}{2}(\delta y_{1/2} + \delta y_{3/2})$, $\mu\delta^2 y_{3/2} = \frac{1}{2}(\delta^2 y_1 + \delta^2 y_2)$ etc.

**Obs.** *The reader should note that it is only the notation which changes and not the differences.*

*e.g.*
$$y_1 - y_0 = \nabla y_0 = \Delta y_1 = \delta y_{1/2}.$$

Of all the formulae, those involving central differences are most useful in practice as the coefficients in such formulae decrease much more rapidly.

**Example 6.1.** *Evaluate* (*i*) $\Delta \tan^{-1} x$ (*ii*) $\Delta (e^x \log 2x)$ (*iii*) $\Delta (x^2 / \cos 2x)$ (*iv*) $\Delta^2 \cos 2x$.

(*P.T.U., B. Tech., 2001*)

**Sol.** (*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*) $\Delta (e^x \log 2x) = e^{x+h} \log 2(x+h) - e^x \log 2x$

$$= e^{x+h} \log 2(x+h) - e^{x+h} \log 2x + e^{x+h} \log 2x - e^x \log 2x$$

$$= e^{x+h}\log\frac{x+h}{x} + (e^{x+h} - e^x)\log 2x$$

$$= e^x\left[e^h\log\left(1+\frac{h}{x}\right) + (e^h - 1)\log 2x\right]$$

(*iii*) $$\Delta\left(\frac{x^2}{\cos 2x}\right) = \frac{(x+h)^2}{\cos 2(x+h)} - \frac{x^2}{\cos 2x} = \frac{(x+h)^2\cos 2x - x^2\cos 2(x+h)}{\cos 2(x+h)\cos 2x}$$

$$= \frac{[(x+h)^2 - x^2]\cos 2x + x^2[\cos 2x - \cos 2(x+h)]}{\cos 2(x+h)\cos 2x}$$

$$= \frac{(2hx + h^2)\cos 2x + 2x^2\sin(h)\sin(2x+h)}{\cos 2(x+h)\cos 2x}$$

(*iv*) $$\Delta^2\cos 2x = \Delta[\cos 2(x+h) - \cos 2x]$$
$$= \Delta\cos 2(x+h) - \Delta\cos 2x$$
$$= [\cos 2(x+2h) - \cos 2(x+h)] - [\cos 2(x+h) - \cos 2x]$$
$$= -2\sin(2x+3h)\sin h + 2\sin(2x+h)\sin h$$
$$= -2\sin h[\sin(2x+3h) - \sin(2x+h)]$$
$$= -2\sin h[2\cos(2x+2h)\sin h] = -4\sin^2 h\cos(2x+2h).$$

**Example 6.2.** *Evaluate* (*i*) $\Delta^2\left(\dfrac{5x+12}{x^2+5x+16}\right)$ *(Mumbai, B. Tech., 2003)*

(*ii*) $\Delta^2(ab^x)$ (*iii*) $\Delta^n(e^x)$ *(Rohtak, B. Tech., 2003)*

*interval of differencing being unity.*

**Sol.** (*i*) $$\Delta^2\left(\frac{5x+12}{x^2+5x+6}\right) = \Delta^2\left\{\frac{5x+12}{(x+2)(x+3)}\right\} = \Delta^2\left\{\frac{2}{x+2} + \frac{3}{x+3}\right\}$$

$$= \Delta\left\{\Delta\left(\frac{2}{x+2}\right) + \Delta\left(\frac{3}{x+3}\right)\right\} = \Delta\left\{2\left(\frac{1}{x+3} - \frac{1}{x+2}\right) + 3\left(\frac{1}{x+4} - \frac{1}{x+3}\right)\right\}$$

$$= -2\Delta\left\{\frac{1}{(x+2)(x+3)}\right\} - 3\Delta\left\{\frac{1}{(x+3)(x+4)}\right\}$$

$$= -2\left\{\frac{1}{(x+3)(x+4)} - \frac{1}{(x+2)(x+3)}\right\} - 3\left\{\frac{1}{(x+4)(x+5)} - \frac{1}{(x+3)(x+4)}\right\}$$

$$= \frac{4}{(x+2)(x+3)(x+4)} + \frac{6}{(x+3)(x+4)(x+5)} = \frac{2(5x+16)}{(x+2)(x+3)(x+4)(x+5)}.$$

(*ii*) $$\Delta(ab^x) = a\,\Delta(b^x) = a(b^{x+1} - b^x) = ab^x(b-1)$$
$$\Delta^2(ab^x) = \Delta[\Delta(ab^x)] = a(b-1)\,\Delta(b^x) = a(b-1)(b^{x+1} - b^x) = a(b-1)^2 b^x.$$

(*iii*) $$\Delta e^x = e^{x+1} - e^x = (e-1)e^x$$
$$\Delta^2 e^x = \Delta(\Delta e^x) = \Delta[(e-1)e^x] = (e-1)\,\Delta e^x = (e-1)(e-1)e^x = (e-1)^2 e^x$$

Similarly $\Delta^3 e^x = (e-1)^3 e^x,\ \Delta^4 e^x = (e-1)^4 e^x, \ldots\ldots$

and $\Delta^n e^x = (e-1)^n e^x.$



**Example 6.3.** *If* $y = a(3)^x + b(-2)^x$ *and* $h = 1$, *prove that* $(\Delta^2 + \Delta - 6)y = 0$ *(Mumbai, B.Tech., 2003)*

**Sol.** We have $y = a(3)^x + b(-2)^x$

$\therefore$ $$\Delta y = [a(3)^{x+1} + b(-2)^{x+1}] - [a(3)^x + b(-2)^x] = 2a(3)^x - 3b(-2)^x$$

and $$\Delta^2 y = [2a(3)^{x+1} - 3b(-2)^{x+1}] - [2a(3)^x - 3b(-2)^x] = 4a(3)^x + 9b(-2)^x$$

Hence $(\Delta^2 + \Delta - 6)y = [4a(3)^x + 9b(-2)^x] + (2a(3)^x - 3b(-2)^x] - 6[a(3)^x + b(-2)^x] = 0$

**Example 6.4.** *Find the missing* $y_x$ *values from the first differences provided :*

| $y_x$ | 0 | — | — | — | — | — |
|---|---|---|---|---|---|---|
| $\Delta y_x$ | 0 | 1 | 2 | 4 | 7 | 11 |

**Sol.** Let the missing values be $y_1, y_2, y_3, y_4, y_5$. Then we have

| $y_x$ | 0 | $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ |
|---|---|---|---|---|---|---|
| $\Delta y_x$ | 0 | 1 | 2 | 4 | 7 | 11 |

$\therefore$ $$y_1 - 0 = 1,\ y_2 - y_1 = 2,\ y_3 - y_2 = 4,\ y_4 - y_3 = 7,\ y_5 - y_4 = 11$$

*i.e.,* $$y_1 = 1,\ y_2 = 2 + y_1 = 3,\ y_3 = 4 + y_2 = 7,\ y_4 = 7 + y_3 = 14,\ y_5 = 11 + y_4 = 25.$$

## 6.3. DIFFERENCES OF A POLYNOMIAL

*The nth differences of a polynomial of the nth degree are constant and all higher order differences are zero.*

Let the polynomial of the $n$th degree in $x$, be $f(x) = ax^n + bx^{n-1} + cx^{n-2} + \ldots\ldots + kx + l$

$\therefore$ $$\Delta f(x) = f(x+h) - f(x)$$
$$= a[(x+h)^n - x^n] + b[(x+h)^{n-1} - x^{n-1}] + \ldots\ldots + kh$$
$$= anhx^{n-1} + b'x^{n-2} + c'x^{n-3} + \ldots\ldots + k'x + l' \quad \ldots(1)$$

where $b', c', \ldots\ldots, l'$ are the new constant co-efficients.

Thus the first differences of a polynomial of the $n$th degree is a polynomial of degree $(n-1)$.

Similarly $$\Delta^2 f(x) = \Delta[f(x+h) - f(x)] = \Delta f(x+h) - \Delta f(x)$$
$$= anh[(x+h)^{n-1} - x^{n-1}] + b'[(x+h)^{n-2} - x^{n-2}] + \ldots\ldots + k'h$$
$$= an(n-1)h^2x^{n-2} + b''x^{n-3} + c''x^{n-4} + \ldots\ldots + k'', \quad \text{by (1)}$$

$\therefore$ The second differences represent a polynomial of degree $(n-2)$.

Continuing this process, for the $n$th differences we get a polynomial of degree zero *i.e.*

$$\Delta^n f(x) = an(n-1)(n-2)\ldots 1.h^n = a\,n!\,h^n \quad \ldots(2)$$

which is a constant. Hence the $(n+1)$th and higher differences of a polynomial of $n$th degree will be zero.

**Obs.** The converse of this theorem is also true *i.e. if the nth differences of a function tabulated at equally spaced intervals are constant, the function is a polynomial of degree n.* This fact is important in numerical analysis as it enables us to approximate a function by a polynomial of $n$th degree, if its $n$th order differences become nearly constant.

**Example 6.5.** *Evaluate* $\Delta^{10}[(1-ax)(1-bx^2)(1-cx^3)(1-dx^4)]$.

**Sol.** $\Delta^{10}[(1-ax)(1-bx^2)(1-cx^3)(1-dx^4)] = \Delta^{10}[abcd\,x^{10} + (\ )x^9 + (\ )x^8 + \ldots + 1]$

$= abcd\,\Delta^{10}(x^{10})$ $\qquad [\because \Delta^{10}(x^n) = 0$ for $n < 10$

$= abcd\,(10\,!)$ $\qquad$ [by (2) above.

## PROBLEMS 6.1

1. Write forward difference table if

| | | | | |
|---|---|---|---|---|
| $x$ : | 10 | 20 | 30 | 40 |
| $y$ : | 1.1 | 2.0 | 4.4 | 7.9 |

2. Construct the table of differences for the data below :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 |
| $f(x)$ : | 1.0 | 1.5 | 2.2 | 3.1 | 4.6 |

Evaluate $\Delta^3 f(2)$.

3. If $u_0 = 3$, $u_1 = 12$, $u_2 = 81$, $u_3 = 2000$, $u_4 = 100$, calculate $\Delta^4 u_0$.
4. Show that $\Delta^3 y_i = y_{i+3} - 3y_{i+2} + 3y_{i+1} - y_i$.
5. Form the table of backward differences of the function $f(x) = x^3 - 3x^2 - 5x - 7$ for $x = -1, 0, 1, 2, 3, 4, 5$.
6. Form a table of differences for the function $f(x) = x^3 + 5x - 7$ for $x = -1, 0, 1, 2, 3, 4, 5$. Continue the table to obtain $f(6)$.
7. Extend the following table to two more terms on either side by constructing the difference table :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | − 0.2 | 0.0 | 0.2 | 0.4 | 0.6 | 0.8 | 1.0 |
| $y$ : | 2.6 | 3.0 | 3.4 | 4.28 | 7.08 | 14.2 | 29.0. |

8. Show that

(i) $\Delta\left[\dfrac{1}{f(x)}\right] = \dfrac{-\Delta f(x)}{f(x)f(x+1)}$ (J.N.T.U., B. Tech., 2009)

(ii) $\Delta \log f(x) = \log\left\{1 + \dfrac{\Delta f(x)}{f(x)}\right\}$.

(iii) $\Delta[x(x+1)(x+2)(x+3)] = 4(x+1)(x+2)(x+3)$ (J.N.T.U., B. Tech., 2009)

9. Evaluate (taking interval of differencing as unity) :

(i) $\Delta(x + \cos x)$ (ii) $\Delta \tan^{-1}\left(\dfrac{n-1}{n}\right)$ (J.N.T.U., B. Tech., 2008)

(iii) $\Delta(e^{3x} \log 2x)$ (iv) $\Delta(2^x/x\,!)$

(v) $\Delta\left\{\dfrac{1}{x(x+4)(x+8)}\right\}$ (Madras, B.E., 2001)

10. Evaluate :

(i) $\Delta^2 \cos 3x$ (Mumbai, B. Tech., 2003) (ii) $\Delta^2\left(\dfrac{1}{x^2+5x+6}\right)$ (P.T.U., B. Tech., 2001)

(iii) $\Delta^n (e^{2x+3})$ (iv) $\Delta^n\left(\dfrac{1}{x}\right)$ (v) $\Delta^n \sin(ax+b)$.



11. If $f(x) = e^{ax+b}$, show that its leading differences form a geometric progression. (Mumbai, B. Tech., 2003)
12. Prove that

(i) $y_3 = y_2 + \Delta y_1 + \Delta^2 y_0 + \Delta^3 y_0$ (ii) $\nabla^2 y_8 = y_8 - 2y_7 + y_6$ (iii) $\delta^2 y_5 = y_6 - 2y_5 + y_4$

13. Evaluate :

(i) $\Delta^4[(1-x)(1-2x)(1-3x)(1-4x)]$, $(h = 1)$. (Madras, B.E., 2001)

(ii) $\Delta^{10}[(1-x)(1-2x^2)(1-3x^3)(1-4x^4)]$, if the interval of differencing is 2. (J.N.T.U., B.Tech., 2009)

## 6.4. FACTORIAL NOTIATON

*A product of the form $x(x-1)(x-2)\ldots\ldots(x-r+1)$ is denoted by $[x]^r$ and is called a* **factorial.**

In particular $[x] = x$, $[x]^2 = x(x-1)$, $[x]^3 = x(x-1)(x-2)$, etc.

In general $\quad [x]^n = x(x-1)(x-2)\ldots\ldots(x-n+1)$

*If the interval of differencing is h, then $[x]^n = x(x-h)(x-2h)\ldots\ldots(x-\overline{n-1}h)$ which is called a* **Factorial polynomial** *or* **function.**

*The factorial notation is of special utility in the theory of finite differences. It helps in finding the successive differences of a polynomial directly by simple rule of differentiation.* Similarly given any difference of a function in factorial notation, we can find the corresponding function by simple integration.

*The result of differencing $[x]^r$ is analogous to that of differentiating $x^r$.*

**(2) To show that $\Delta^n [x]^n = n\,!$ and $\Delta^{n+1} [x]^n = 0$**

We have $\quad \Delta[x]^n = [x+h]^n - [x]^n$

$$= (x+h)(x+h-h)(x+h-2h)\ldots\ldots(x+h-\overline{n-1}h)$$

$$- x(x-h)(x-2h)\ldots\ldots(x-\overline{n-1}h)$$

$$= x(x-h)\ldots(x-\overline{n-2}h)\,[x+h-(x-nh+h)]$$

$$= nh[x]^{n-1} \qquad \ldots(i)$$

Similarly $\quad \Delta^2[x]^n = \Delta[nh\,[x]^{n-1}] = nh\,\Delta[x]^{n-1}$

Replacing $n$ by $n-1$ in $(i)$, we get $\Delta^2[x]^n = nh.(n-1)h[x]^{n-2} = n(n-1)h^2[x]^{n-2}$

Proceeding in this way, we obtain $\Delta^{n-1}[x]^n = n(n-1)\ldots\ldots 2\,h^{n-1}\,x$

$\therefore$

$$\Delta^n[x]^n = n(n-1)\ldots\ldots 2.h^{n-1}\,\Delta x$$

$$= n(n-1)\ldots\ldots 2.1.h^{n-1}(x+h-x)$$

$$= n\,!\,h^n \qquad \ldots(ii)$$

Also $\quad \Delta^{n+1}[x]^n = n\,!\,h^n - n\,!\,h^n = 0$

In particular, when $h = 1$, we have $\qquad \ldots(iii)$

$$\Delta[x]^n = n[x]^{n-1} \quad \text{and} \quad \Delta^n[x]^n = n\,!$$

*Thus we arrive at the following* **important conclusion :**

*The result of differencing $[x]^r$ is analogous to that of differentiating $x^r$ when $h = 1$.*

**Obs.** Every polynomial of degree $n$ can be expressed as a factorial polynomial of the same degree and *vice versa.*

**Example 6.6.** *Express $y = 2x^3 - 3x^2 + 3x - 10$ in factorial notation and hence show that $\Delta^3 y = 12$.* (Bhopal, B.E., 2007)

**Sol. First method :** Let $y = A[x]^3 + B[x]^2 + C[x] + D$.

Using the method of synthetic division (p. 18), we divide by $x, x-1, x-2$, etc. successively. Then

| | $x^3$ | $x^2$ | $x$ | |
|---|---|---|---|---|
| 1 | 2 | −3 | 3 | $-10 = D$ |
| | – | 2 | −1 | |
| 2 | 2 | −1 | $2 = C$ | |
| | — | 4 | | |
| 3 | 2 | $3 = B$ | | |
| | — | | | |
| | $2 = A$ | | | |

Hence $y = 2[x]^3 + 3[x]^2 + 2[x] - 10$

$\therefore$ $\Delta y = 2 \times 3[x]^2 + 3 \times 2[x] + 2$

$\Delta^2 y = 6 \times 2[x] + 6$

$\Delta^3 y = 12$, which shows that the third differences of $y$ are constant, as they should be.

**Obs.** *The coefficient of the highest power of $x$ remains unchanged while transforming a polynomial to factorial notation.*

**Second method** (*Direct method*) :

Let $y = 2x^3 - 3x^2 + 3x - 10$
$= 2x(x-1)(x-2) + Bx(x-1) + Cx + D$

Putting $x = 0$, $-10 = D$.

Putting $x = 1$, $2 - 3 + 3 - 10 = C + D$

$\therefore$ $C = -8 - D = -8 + 10 = 2$

Putting $x = 2$, $16 - 12 + 6 - 10 = 2B + 2C + D$

$\therefore$ $B = \frac{1}{2}(-2C - D) = \frac{1}{2}(-4 + 10) = 3$.

Hence $y = 2x(x-1)(x-2) + 3x(x-1) + 2x - 10 = 2[x]^3 + 3[x]^2 + 2[x] - 10$

$\therefore$ $\Delta y = 2 \times 3[x]^2 + 3 \times 2[x] + 2$, $\Delta^2 y = 6 \times 2[x] + 6$, $\Delta^3 y = 12$.

**Example 6.7.** *Obtain the function whose first difference is $9x^2 + 11x + 5$.*

**Sol.** Let $f(x)$ be the required function so that $\Delta f(x) = 9x^2 + 11x + 5$.

Let $9x^2 + 11x + 5 = 9[x]^2 + A[x] + B = 9x(x-1) + Ax + B$

Putting $x = 0$, $B = 5$.

Putting $x = 1$, $A = 20$.

$\therefore$ $\Delta f(x) = 9[x]^2 + 20[x] + 5$

Integrating, we get

$$f(x) = 9\frac{[x]^3}{3} + 20\frac{[x]^2}{2} + 5[x] + c$$
$$= 3x(x-1)(x-2) + 10x(x-1) + 5x + c = 3x^3 + x^2 + x + c$$

where c is the constant of integration.



**Example 6.8.** *Express $u = x^4 - 12x^3 + 24x^2 - 30x + 9$ and its successive differences in factorial notation. Hence show that $\Delta^5 u = 0$.* (Mumbai, B. Tech., 2004)

**Sol.** Let $u = A[x]^4 + B[x]^3 + C[x]^2 + D[x] + E$.

Using the method of synthetic division, we divide by $x, x-1, x-2, x-3$ successively.

Then

| | $x^4$ | $x^3$ | $x^2$ | $x$ | |
|---|---|---|---|---|---|
| 1 | 1 | −12 | 24 | −30 | $9 (= E)$ |
| | 0 | 1 | −11 | 13 | |
| 2 | 1 | −11 | 13 | $-17 (= D)$ | |
| | 0 | 2 | −18 | | |
| 3 | 1 | −9 | $-5 (= C)$ | | |
| | 0 | 3 | | | |
| | $1 (= A)$ | $-6 (= B)$ | | | |

Hence $u = [x]^4 - 6[x]^3 - 5[x]^2 - 17[x] + 9$

$\therefore$ $\Delta u = 4[x]^3 - 18[x]^2 - 10[x] - 17$

$\Delta^2 u = 12[x]^2 - 36[x] - 10$

$\Delta^3 u = 24[x] - 36$

$\Delta^4 u = 24$ and $\Delta^5 u = 0$.

## 6.5. EFFECT OF AN ERROR ON A DIFFERENCE TABLE

Suppose there is an error $\varepsilon$ in the entry $y_5$ of a table. As higher differences are formed, this error spreads out and is considerably magnified. Let us see, how it effects the difference table.

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|---|---|---|---|---|---|
| $x_0$ | $y_0$ | | | | |
| | | $\Delta y_0$ | | | |
| $x_1$ | $y_1$ | | $\Delta^2 y_0$ | | |
| | | $\Delta y_1$ | | $\Delta^3 y_0$ | |
| $x_2$ | $y_2$ | | $\Delta^2 y_1$ | | $\Delta^4 y_0$ |
| | | $\Delta y_2$ | | $\Delta^3 y_1$ | |
| $x_3$ | $y_3$ | | $\Delta^2 y_2$ | | $\Delta^4 y_1 + \varepsilon$ |
| | | $\Delta y_3$ | | $\Delta^3 y_2 + \varepsilon$ | |
| $x_4$ | $y_4$ | | $\Delta^2 y_3 + \varepsilon$ | | $\Delta^4 y_2 - 4\varepsilon$ |
| | | $\Delta y_4 + \varepsilon$ | | $\Delta^3 y_3 - 3\varepsilon$ | |
| $x_5$ | $y_5 + \varepsilon$ | | $\Delta^2 y_4 - 2\varepsilon$ | | $\Delta^4 y_3 + 6\varepsilon$ |
| | | $\Delta y_5 - \varepsilon$ | | $\Delta^3 y_4 + 3\varepsilon$ | |
| $x_6$ | $y_6$ | | $\Delta^2 y_5 + \varepsilon$ | | $\Delta^4 y_4 - 4\varepsilon$ |
| | | $\Delta y_6$ | | $\Delta^3 y_5 - \varepsilon$ | |
| $x_7$ | $y_7$ | | $\Delta^2 y_6$ | | $\Delta^4 y_5 + \varepsilon$ |
| | | $\Delta y_7$ | | $\Delta^3 y_6$ | |
| $x_8$ | $y_8$ | | $\Delta^2 y_7$ | | |
| | | $\Delta y_8$ | | | |
| $x_9$ | $y_9$ | | | | |

The above table shows that :

(*i*) The error increases with the order of differences.

(*ii*) The coefficients of $\varepsilon$'s in any column are the binomial coefficients of $(1-\varepsilon)^n$. Thus the errors in the fourth difference column are $\varepsilon, -4\varepsilon, 6\varepsilon, -4\varepsilon, \varepsilon$.

(*iii*) The algebraic sum of the errors in any difference column is zero.

(*iv*) The maximum error in each column, occurs opposite to the entry containing the error *i.e.* $y_5$.

The above facts enable us to detect errors in a difference table.

**Example 6.9.** *One entry in the following table is incorrect and y is a cubic polynomial in x. Use the difference table to locate and correct the error.*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $y$ : | 25 | 21 | 18 | 18 | 27 | 45 | 76 | 123 |

**Sol.** The difference table is as under :

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ |
|---|---|---|---|---|
| 0 | 25 | | | |
| | | − 4 | | |
| 1 | 21 | | 1 | |
| | | − 3 | | 2 |
| 2 | 18 | | 3 | |
| | | 0 | | 6 |
| 3 | 18 | | 9 | |
| | | 9 | | 0 |
| 4 | 27 | | 9 | |
| | | 18 | | 4 |
| 5 | 45 | | 13 | |
| | | 31 | | 3 |
| 6 | 76 | | 16 | |
| | | 47 | | |
| 7 | 123 | | | |

$y$ being a polynomial of the third degree, $\Delta^3 y$ must be constant *i.e.* the same. The sum of the third differences being 15, each entry under $\Delta^3 y$ must be 15/5 *i.e.* 3. Thus the four entries under $\Delta^3 y$ are in error which can be written as

$$3 + (-1),\ 3 - 3(-1),\ 3 + 3(-1),\ 3 - (-1)$$

Taking $\varepsilon = -1$, we find that the entry corresponding to $x = 3$ is in error.

[Comparing with table on p. 130]

$$\therefore \qquad y + \varepsilon = 18$$

Thus the true value of $y = 18 - \varepsilon = 18 - (-1) = 19$.

**Example 6.10.** *Assuming that the following values of y belong to a polynomial of degree 4, compute the next three values :*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| $y$ : | 1 | − 1 | 1 | − 1 | 1 | — | — | — |



**Sol.** We construct the difference table from the given data.

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|---|---|---|---|---|---|
| 0 | $y_0 = 1$ | | | | |
| | | − 2 | | | |
| 1 | $y_1 = -1$ | | 4 | | |
| | | 2 | | | |
| 2 | $y_2 = 1$ | | − 4 | − 8 | |
| | | − 2 | | | |
| 3 | $y_3 = -1$ | | 4 | 8 | 16 |
| | | 2 | | | |
| 4 | $y_4 = 1$ | | $\Delta^2 y_3$ | $\Delta^3 y_2$ | 16 |
| | | $\Delta y_4$ | | | |
| 5 | $y_5$ | | $\Delta^2 y_4$ | $\Delta^3 y_3$ | 16 |
| | | $\Delta y_5$ | | | |
| 6 | $y_6$ | | $\Delta^2 y_5$ | $\Delta^3 y_4$ | 16 |
| | | $\Delta y_6$ | | | |
| 7 | $y_7$ | | | | |

Since the values of $y$ belong to a polynomial of degree 4, the fourth differences must be constant. But $\Delta^4 y = 16$.

$\therefore$ The other fourth order differences must also be 16. Thus

$$\Delta^4 y_1 = 16 = \Delta^3 y_2 - \Delta^3 y_1$$

*i.e.*

$$\Delta^3 y_2 = \Delta^3 y_1 + \Delta^4 y_1 = 8 + 16 = 24$$
$$\Delta^2 y_3 = \Delta^2 y_2 + \Delta^3 y_2 = 4 + 24 = 28$$
$$\Delta y_4 = \Delta y_3 + \Delta^2 y_3 = 2 + 28 = 30$$

and

$$y_5 = y_4 + \Delta y_4 = 1 + 30 = 31$$

Similarly starting with $\Delta^4 y_2 = 16$, we get $\Delta^3 y_3 = 40, \Delta^2 y_4 = 68, \Delta y_5 = 98, y_6 = 129$.

Starting with $\Delta^4 y_3 = 16$, we obtain $\Delta^3 y_4 = 56, \Delta^2 y_5 = 124, \Delta y_6 = 222, y_7 = 351$.

## PROBLEMS 6.2

1. Express $x^3 - 2x^2 + x - 1$ into factorial polynomial. Hence show that $\Delta^4 f(x) = 0$. *(P.T.U., B. Tech., 2001)*
2. Express $3x^4 - 4x^3 + 6x^2 + 2x + 1$ as a factorial polynomial and find differences of all orders.
3. Find the first and second differences of $x^4 - 6x^3 + 11x^2 - 5x + 8$ with $h = 1$. Show that the fourth difference is constant.
4. Obtain the function whose first difference is $2x^3 + 3x^2 - 5x + 4$.
5. Given $\log 100 = 2$, $\log 101 = 2.0043$, $\log 103 = 2.0128$, $\log 104 = 2.0170$, find $\log 102$.
6. Find the first term of the series whose second and subsequent terms are 8, 3, 0, − 1, 0.
7. Write down the polynomial of lowest degree which satisfies the following set of numbers :

   0, 7, 26, 63, 124, 215, 342, 511.
8. Find and correct by means of differences, the error in the data :

   20736, 28561, 38416, 50625, 65540, 83521, 104976, 130321, 160000.

## 6.6. OTHER DIFFERENCE OPERATORS

We have already introduced the operators $\Delta$, $\nabla$ and $\delta$. Besides these, there are the operators $E$ and $\mu$, which we define below :

**(1) Shift operator E** *is the operation of increasing the argument $x$ by $h$ so that* $Ef(x) = f(x + h)$, $E^2 f(x) = f(x + 2h)$, $E^3 f(x) = f(x + 3h)$ etc.

The inverse operator $E^{-1}$ is defined by $E^{-1} f(x) = f(x - h)$

If $y_x$ is the function $f(x)$, then $Ey_x = y_{x+h}$, $E^{-1}y_x = y_{x-h}$, $E^n y_x = y_{x+nh}$, where $n$ may be any real number.

**(2) Averaging operator** $\mu$ is defined by the equation $\mu y_x = \frac{1}{2}(y_{x+\frac{1}{2}h} + y_{x-\frac{1}{2}h})$

**Obs.** *In the difference calculus $E$ is regarded as the fundamental operator and $\Delta, \nabla, \delta, \mu$ can be expressed in terms of $E$.*

## 6.7. RELATIONS BETWEEN THE OPERATORS

(1) We shall now establish the following identities :

(*i*) $\Delta = \mathbf{E} - 1$ (*ii*) $\nabla = 1 - \mathbf{E}^{-1}$

(*iii*) $\delta = \mathbf{E}^{1/2} - \mathbf{E}^{-1/2}$ (*iv*) $\mu = \frac{1}{2}(\mathbf{E}^{1/2} + \mathbf{E}^{-1/2})$

(*v*) $\Delta = \mathbf{E}\nabla = \nabla\mathbf{E} = \delta\mathbf{E}^{1/2}$ (*vi*) $\mathbf{E} = \mathbf{e}^{\mathbf{hD}}$.

*Proofs.* (*i*) $\Delta y_x = y_{x+h} - y_x = Ey_x - y_x = (E - 1)y_x$

This shows that the operators $\Delta$ and $E$ are connected by the symbolic relation

$$\Delta = E - 1 \quad \text{or} \quad \mathbf{E} = 1 + \Delta.$$

**Obs.** *These relations imply that the effect of operator $E$ on $y_x$ is the same as that of the operator $(1 + \Delta)$ on $y_x$. The operator's E and $\Delta$ do not have any existence as separate entities.*

(*ii*) $\nabla y_x = y_x - y_{x-h} = y_x - E^{-1} y_x = (1 - E^{-1})y_x$

$\therefore$ $\nabla = 1 - E^{-1}$

(*iii*) $\delta y_x = y_{x+\frac{1}{2}h} - y_{x-\frac{1}{2}h} = E^{1/2} y_x - E^{-1/2} y_x = (E^{1/2} - E^{-1/2})y_x$

$\delta = E^{1/2} - E^{-1/2}$.

(*iv*) $\mu y_x = \frac{1}{2}(y_{x+\frac{1}{2}h} + y_{x-\frac{1}{2}h}) = \frac{1}{2}(E^{1/2}y_x + E^{-1/2}y_x) = \frac{1}{2}(E^{1/2} + E^{-1/2})y_x$

$\therefore$ $\mu = \frac{1}{2}(E^{1/2} + E^{-1/2})$.

(*v*) $E\nabla y_x = E(y_x - y_{x-h}) = Ey_x - Ey_{x-h} = y_{x+h} - y_x = \Delta y_x$

$\therefore$ $E\nabla = \Delta$

$\nabla E y_x = \nabla y_{x+h} = y_{x+h} - y_x = \Delta y_x$

$\therefore$ $\nabla E = \Delta$

$\delta E^{1/2} y_x = \delta y_{x+\frac{1}{2}h} = y_{x+\frac{1}{2}h+\frac{1}{2}h} - y_{x+\frac{1}{2}h-\frac{1}{2}h} = y_{x+h} - y_x = \Delta y_x$

$\therefore$ $\delta E^{1/2} = \Delta$

Hence $\Delta = E\nabla = \nabla E = \delta E^{1/2}$.

(*vi*) $Ef(x) = f(x + h) = f(x) + hf'(x) + \frac{h^2}{2!} f''(x) + \ldots$ [by Taylor's series]

$= f(x) + hDf(x) + \frac{h^2}{2!} D^2 f(x) + \ldots$



$$= \left(1 + hD + \frac{h^2D^2}{2!} + \frac{h^3D^3}{3!} + \ldots\right) f(x) = e^{hD} f(x)$$

$E = e^{hD}$

$\therefore$ $E = 1 + \Delta = e^{hD}$.

**Cor.**

**Note.** *A table showing the symbolic relations between the various operators is given below for ready reference. To prove such relations between the operators, always express each operator in terms of the fundamental operator E.*

**(2) Relations between the various operators**

| *In terms of* | $E$ | $\Delta$ | $\nabla$ | $\delta$ | $hD$ |
|---|---|---|---|---|---|
| $E$ | — | $\Delta + 1$ | $(1 - \nabla)^{-1}$ | $1 + \frac{1}{2}\delta^2 + \delta\sqrt{(1 + \delta^2/4)}$ | $e^{hD}$ |
| $\Delta$ | $E - 1$ | — | $(1 - \nabla)^{-1} - 1$ | $\frac{1}{2}\delta^2 + \delta\sqrt{(1 + \delta^2/4)}$ | $e^{hD} - 1$ |
| $\nabla$ | $1 - E^{-1}$ | $1 - (1 + \Delta)^{-1}$ | — | $-\frac{1}{2}\delta^2 + \delta\sqrt{(1 + \delta^2/4)}$ | $1 - e^{-hD}$ |
| $\delta$ | $E^{1/2} - E^{-1/2}$ | $\Delta(1 + \Delta)^{-1/2}$ | $\nabla(1 - \nabla)^{-1/2}$ | — | $2 \sinh (hD/2)$ |
| $\mu$ | $\frac{1}{2}(E^{1/2} + E^{-1/2})$ | $(1 + \Delta/2)(1 + \Delta)^{-1/2}$ | $(1 + \nabla/2)(1 + \nabla)^{-1/2}$ | $\sqrt{(1 + \delta^2/4)}$ | $\cosh (hD/2)$ |
| $hD$ | $\log E$ | $\log (1 + \Delta)$ | $\log (1 - \nabla)^{-1}$ | $2 \sinh^{-1} (\delta/2)$ | — |

**Example 6.11.** *Prove that* $e^x = \left(\frac{\Delta^2}{E}\right) e^x \cdot \frac{Ee^x}{\Delta^2 e^x}$, *the interval of differencing being h.* (*Bhopal, B.E., 2009*)

**Sol.** Since $\left(\frac{\Delta^2}{E}\right) e^x = \Delta^2 . E^{-1} e^x = \Delta^2 e^{x-h} = \Delta^2 e^x . e^{-h} = e^{-h} \Delta^2 e^x$

$\therefore$ R.H.S. $= e^{-h} \Delta^2 e^x . \frac{Ee^x}{\Delta^2 e^x} = e^{-h} Ee^x = e^{-h} . e^{x+h} = e^x$.

**Example 6.12.** *Prove with the usual notations, that* (*M.D.U., B. Tech., 2005*)

(*i*) $hD = \log(1 + \Delta) = -\log(1 - \nabla) = \sinh^{-1}(\mu\delta)$ (*Bhopal, B. Tech., 2009*)

(*ii*) $(E^{1/2} + E^{-1/2})(1 + \Delta)^{1/2} = 2 + \Delta$ (*Mumbai, B.E., 2005*)

(*iii*) $\Delta - \nabla = \Delta\nabla = \delta^2$

(*iv*) $\Delta^3 y_2 = \nabla^3 y_5$.

**Sol.** (*i*) We know that $e^{hD} = E = 1 + \Delta$

$\therefore$ $hD = \log(1 + \Delta)$

Also $hD = \log E = -\log(E^{-1}) = -\log(1 - \nabla)$ [$\because$ $E^{-1} = 1 - \nabla$]

We have proved that $\mu = \frac{1}{2}(E^{1/2} + E^{-1/2})$

and $\delta = E^{1/2} - E^{-1/2}$

$\therefore$ $\mu\delta = \frac{1}{2}(E^{1/2} + E^{-1/2})(E^{1/2} - E^{-1/2})$

$= \frac{1}{2}(E - E^{-1}) = \frac{1}{2}(e^{hD} - e^{-hD}) = \sinh(hD)$

*i.e.* $hD = \sinh^{-1}(\mu\delta)$.

Hence $hD = \log(1 + \Delta) = -\log(1 - \Delta) = \sinh^{-1}(\mu\delta)$.

(*ii*) $(E^{1/2} + E^{-1/2})(1 + \Delta)^{1/2}$

$= (E^{1/2} + E^{-1/2})E^{1/2} = E + 1 = 1 + \Delta + 1 = 2 + \Delta$.

(*iii*) We know that $\Delta = E - 1$, $\nabla = 1 - E^{-1}$ and $\delta = E^{1/2} - E^{-1/2}$

$\therefore$ $\Delta - \nabla = E - 2 + E^{-1} = (E^{1/2} - E^{-1/2})^2 = \delta^2$

Also $\Delta\nabla = (E - 1)(1 - E^{-1}) = E + E^{-1} - 2$

$= (E^{1/2} - E^{-1/2})^2 = \delta^2$.

Hence $\Delta - \nabla = \Delta\nabla = \delta^2$.

(*iv*) $\Delta^3 y_2 = (E - 1)^3 y_2$ $[\because \Delta = E - 1$

$= (E^3 - 3E^2 + 3E - 1)y_2$

$= y_5 - 3y_4 + 3y_3 - y_2$ ...(1)

$\nabla^3 y_5 = (1 - E^{-1})^3 y_5$ $[\because \nabla = 1 - E^{-1}$

$= (1 - 3E^{-1} + 3E^{-2} - E^{-3})y_5$

$= y_5 - 3y_4 + 3y_3 - y_2$ ...(2)

From (1) and (2), $\Delta^3 y_2 = \nabla^3 y_5$.

**Example 6.13.** *Prove that*

(*i*) $\Delta = \frac{1}{2}\delta^2 + \delta\sqrt{\left(1 + \frac{\delta^2}{4}\right)}$ (*Mumbai, B.E., 2004*)

(*ii*) $1 + \delta^2\mu^2 = \left(1 + \frac{1}{2}\delta^2\right)^2$

(*iii*) $\mu = \frac{2 + \Delta}{2\sqrt{(1 + \Delta)}} = \sqrt{\left(1 + \frac{1}{4}\delta^2\right)}$

**Sol.** (*i*) $\frac{1}{2}\delta^2 + \delta\sqrt{\left(1 + \frac{\delta^2}{4}\right)}$ $[\because \delta = E^{1/2} - E^{-1/2}]$

$$= \frac{1}{2}(E^{1/2} - E^{-1/2})^2 + (E^{1/2} - E^{-1/2})\sqrt{[1 + (E^{1/2} - E^{-1/2})^2/4]}$$

$$= \frac{1}{2}(E + E^{-1} - 2) + (E^{1/2} - E^{-1/2})\sqrt{[(E + E^{-1} + 2)/4]}$$

$$= \frac{1}{2}(E + E^{-1} - 2) + (E^{1/2} - E^{-1/2})(E^{1/2} + E^{-1/2})/2$$

$$= \frac{1}{2}[(E + E^{-1} - 2) + (E - E^{-1})] = E - 1 = \Delta.$$



(*ii*) We know that $\delta = E^{1/2} - E^{-1/2}$ and $\mu = (E^{1/2} + E^{-1/2})/2$.

$\therefore$ L.H.S. $= 1 + \delta^2\mu^2 = 1 + (E^{1/2} - E^{-1/2})^2(E^{1/2} + E^{-1/2})^2/4$

$$= \frac{1}{4}[4 + (E - E^{-1})^2] = \frac{1}{4}(E^2 + E^{-2} + 2) = \frac{1}{4}(E + E^{-1})^2$$

R.H.S. $= (1 + \frac{1}{2}\delta^2)^2 = [1 + \frac{1}{2}(E^{1/2} - E^{-1/2})^2]^2 = [1 + \frac{1}{2}(E + E^{-1} - 2)]^2$

$$= \frac{1}{4}(E + E^{-1})^2$$

Hence $1 + \delta^2\mu^2 = \left(1 + \frac{1}{2}\delta^2\right)^2$.

(*iii*) Since $\Delta = E - 1$, $\delta = E^{1/2} - E^{-1/2}$ and $\mu = \frac{1}{2}(E^{1/2} + E^{-1/2})$

$$\therefore \quad \frac{2 + \Delta}{2\sqrt{(1 + \Delta)}} = \frac{2 + E - 1}{2\sqrt{(1 + E - 1)}} = \frac{E + 1}{2\sqrt{E}}$$

$$= \frac{1}{2}(E^{1/2} + E^{-1/2}) = \mu \quad ...(1)$$

Also $\sqrt{\left(1 + \frac{1}{4}\delta^2\right)} = \sqrt{\left[1 + \frac{1}{4}(E^{1/2} - E^{-1/2})^2\right]} = \sqrt{\left[1 + \frac{1}{4}(E + E^{-1} - 2)\right]}$

$$= \frac{1}{2}\sqrt{(E + E^{-1} + 2)} = \frac{1}{2}(E^{1/2} + E^{-1/2}) = \mu \quad ...(2)$$

Hence from (1) and (2), we get

$$\mu = \frac{2 + \Delta}{2\sqrt{(1 + \Delta)}} = \sqrt{\left(1 + \frac{1}{4}\delta^2\right)}.$$

**Example 6.14.** *Prove that* $\nabla y_{n+1} = h\left(1 + \frac{1}{2}\nabla + \frac{5}{12}\nabla^2 + \ldots\right)y_n'$

**Sol.** We have $\nabla_{y+1} = y_{n+1} - y_n = (E - 1)y_n$

$$= (e^{hD} - 1)y_n = \left(1 + hD + \frac{h^2D^2}{2!} + \frac{h^3D^3}{3!} + \ldots - 1\right)y_n$$

$$= hD\left(1 + \frac{hD}{2!} + \frac{h^2D^2}{3!} + \ldots\right)y_n$$

$$= h\left(1 + \frac{hD}{2!} + \frac{h^2D^2}{3!} + \ldots\right)Dy_n$$

Since $E^{-1} = 1 - \nabla = e^{-hD}$, $\therefore$ $hD = -\log(1 - \nabla) = \nabla + \frac{1}{2}\nabla^2 + \frac{1}{3}\nabla^3 + \ldots$

$$\therefore \quad \nabla y_{n+1} = h\left\{1+\frac{1}{2}\left(\nabla+\frac{1}{2}\nabla^2+\frac{1}{3}\nabla^3+\ldots\ldots\right)+\frac{1}{6}\left(\nabla+\frac{1}{2}\nabla^2+\frac{1}{3}\nabla^3+\ldots\ldots\right)^2+\ldots\ldots\right\}y'_n$$

Hence $$\nabla y_{n+1} = h\left(1+\frac{1}{2}\nabla+\frac{5}{12}\nabla^2+\ldots\ldots\right)y'_n .$$

## 6.8 TO FIND ONE OR MORE MISSING TERMS

When one or more values of $y = f(x)$ corresponding to the equidistant values of $x$ are missing, we can find these using any of the following two methods :

***First method:*** We assume the missing term or terms as $a, b$ etc. and form the difference table. Assuming the last difference as zero, we solve these equations for $a, b$. These give the missing term/terms.

***Second method:*** If $n$ entries of $y$ are given, $f(x)$ can be represented by a $(n-1)$th degree polynomial *i.e.*, $\Delta^n y = 0$. Since $\Delta = E - 1$, therefore $(E-1)^n y = 0$. Now expanding $(E-1)^n$ and substituting the given values, we obtain the missing term/terms.

**Example 6.15.** *Find the missing term is the table :*

| $x$ : | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $y$ : | 45.0 | 49.2 | 54.1 | ... | 67.4 |

*(U.P.T.C., B. Tech., 2008)*

**Sol.** Let the missing value be $a$. Then the difference table is as follows :

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|---|---|---|---|---|---|
| 2 | $45.0\ (= y_0)$ | | | | |
| | | 4.2 | | | |
| 3 | $49.2\ (= y_1)$ | | 0.7 | | |
| | | 4.9 | | $a - 59.7$ | |
| 4 | $54.1\ (= y_2)$ | | $a - 59.0$ | | $240.2 - 4a$ |
| | | $a - 54.1$ | | $180.5 - 3a$ | |
| 5 | $a\ (= y_3)$ | | $121.5 - 2a$ | | |
| | | $67.4 - \alpha$ | | | |
| 6 | $67.4\ (= y_4)$ | | | | |

We know that $\Delta^4 y = 0$ *i.e.*, $240.2 - 4a = 0$.

Hence $\quad a = 60.05.$

**Otherwise.** As only four entries $y_0, y_1, y_2, y_3$ are given, therefore $y = f(x)$ can be represented by a third degree polynomial.

$\therefore \quad \Delta^3 y = \text{constant} \quad \text{or} \quad \Delta^4 y = 0$ *i.e.*, $(E-1)^4 y = 0$

*i.e.*, $\quad (E^4 - 4E^3 + 6E^2 - 4E + 1)\,y = 0 \quad$ or $\quad y_4 - 4y_3 + 6y_2 - 4y_1 + y_0 = 0$

Let the missing entry $y_3$ be $a$ so that

$$67.4 - 4a + 6(54.1) - 4(49.2) + 45 = 0 \quad \text{or} \quad -4a = -240.2$$

Hence $\quad a = 60.05.$

**Example 6.16.** *Find the missing values in the following data :*

| $x$ : | 45 | 50 | 55 | 60 | 65 |
|---|---|---|---|---|---|
| $y$ : | 3.0 | ... | 2.0 | ... | $-2.4$ |

*(Bhopal, B.E., 2007)*

**Sol.** Let the missing values be $a, b$. Then the difference table is as follows :



| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ |
|---|---|---|---|---|
| 45 | $3\ (= y_0)$ | | | |
| | | $a - 3$ | | |
| 50 | $a\ (= y_1)$ | | $5 - 2a$ | |
| | | $2 - a$ | | $3a + b - 9$ |
| 55 | $2\ (= y_2)$ | | $b + a - 4$ | |
| | | $b - 2$ | | $3.6 - a - 36$ |
| 60 | $b\ (= y_3)$ | | $-0.4 - 2b$ | |
| | | $-2.4 - b$ | | |
| 65 | $-2.4\ (= y_4)$ | | | |

As only three entries $y_0, y_2, y_4$ are given, $y$ can be represented by a second degree polynomial having third differences as zero.

$\therefore \quad \Delta^3 y_0 = 0$ and $\Delta^3 y_1 = 0$

*i.e.*, $\quad 3a + b = 9,\ a + 3b = 3.6$

Solving these, we get $a = 2.925$, $b = 0.225$.

**Otherwise.** As only three entries $y_0 = 3$, $y_2 = 2$, $y_4 = -2.4$ are given, $y$ can be represented by a second degree polynomial having third differences as zero.

$\therefore \quad \Delta^3 y_0 = 0$ and $\Delta^3 y_1 = 0$

*i.e.*, $\quad (E-1)^3 y_0 = 0$ and $(E-1)^3 y_1 = 0$

*i.e.*, $\quad (E^3 - 3E^2 + 3E - 1)\,y_0 = 0\ ;\ (E^3 - 3E^2 + 3E - 1)\,.\,y_1 = 0$

or $\quad y_3 - 3y_2 + 3y_1 - y_0 = 0\ ;\ y_4 - 3y_3 + 3y_2 - y_1 = 0$

or $\quad y_3 + 3y_1 = 9\ ;\ 3y_3 + y_1 = 3.6$

Solving three, we get $y_1 = 2.925$, $y_2 = 0.225$.

**Example 6.17.** *The following table gives the values of y which is a polynomial of degree five. It is known that f (3) is in error. Correct the error.*

| $x$ : | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $y$ : | 1 | 2 | 33 | 254 | 1025 | 3126 | 7777 |

*(Mumbai, B.E., 2004)*

**Sol.** Let the correct value of $y$ when $x = 3$ be $a$. Then the difference table is as follows :

| $x$ : | $y$ : | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ | $\Delta^5 y$ | $\Delta^6 y$ |
|---|---|---|---|---|---|---|---|
| 0 | 1 | | | | | | |
| | | 1 | | | | | |
| 1 | 2 | | 30 | | | | |
| | | 31 | | $a-94$ | | | |
| | | | | | $1216-4a$ | | |
| 2 | 33 | | $a-64$ | | | $-2320-10a$ | |
| | | $a-33$ | | $1122-3a$ | | | |
| | | | | | $-1104+6a$ | | $4880-20a$ |
| 3 | $a$ | | $1058-2a$ | | | $2560-10a$ | |
| | | $1025-a$ | | $18+3a$ | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 4 | 1025 | | 1076 + $a$ | | 1456 – 4$a$ | | |
| | | 2101 | | 1474 – $a$ | | | |
| 5 | 3126 | | 2550 | | | | |
| | | 4651 | | | | | |
| 6 | 7777 | | | | | | |

Since $y$ is a polynomial of fifth degree, the sixth difference $\Delta^6 y = 0$

*i.e.,* $\quad 4880 - 20a = 0$

Hence $a = 244$.

**Otherwise.** As $y$ is a polynomial of fifth degree, the sixth difference $\Delta^6 y = 0$

*i.e.,* $\quad (E-1)^6 y = 0$

or $\quad (E^6 - 6E^5 + 15E^4 - 20E^3 + 15E^2 - 6E + 1)\, y_0 = 0$

or $\quad y_6 - 6y_5 + 15y_4 - 20y_3 + 15y_2 - 6y_1 + y_0 = 0$

*i.e.,* $\quad 7777 - 6\,(3126) + 15\,(1025) + 20y_3 + 15\,(33) - 6\,(2) + 1 = 0$

$\therefore \quad 4880 = 20y_3 \quad \therefore \quad y_3 = 244.$

Hence the error $= 254 - 244 = 10$.

**Example 6.18.** *If $y_{10} = 3$, $y_{11} = 6$, $y_{12} = 11$, $y_{13} = 18$, $y_{14} = 27$, find $y_4$.* *(Mumbai, B. Tech., 2005)*

**Sol.** Taking $y_{14}$ as $u_0$, we are required to find $y_4$ *i.e.* $u_{-10}$. Then the difference table is

| $x$ | $u$ | $\Delta u$ | $\Delta^2 u$ | $\Delta^3 u$ |
|---|---|---|---|---|
| $x_{-4}$ | $y_{10} = u_{-4} = 3$ | | | |
| | | 3 | | |
| $x_{-3}$ | $y_{11} = u_{-3} = 6$ | | 2 | |
| | | 5 | | 0 |
| $x_{-2}$ | $y_{12} = u_{-2} = 11$ | | 2 | |
| | | 7 | | 0 |
| $x_{-1}$ | $y_{13} = u_{-1} = 18$ | | 2 | |
| | | 9 | | |
| $x_0$ | $y_{14} = u_0 = 27$ | | | |

Then $\quad y_4 = u_{-10} = (E^{-1})^{10}\, u_0 = (1 - \nabla)^{10}\, u_0$

$$= \left(1 - 10\nabla + \frac{10.9}{2}\nabla^2 - \frac{10.9.8}{1.2.3}\nabla^3 + \ldots\right) u_0$$

$$= u_0 - 10\,\nabla u_0 + 45\,\nabla^2 u_0 - 120\,\nabla^3 u_0$$

$$= 27 - 10 \times 9 + 45 \times 2 - 120 \times 0 = 27.$$

**Example 6.19.** *If $y_x$ is a polynomial for which fifth difference is constant and $y_1 + y_7 = -784$, $y_2 + y_6 = 686$, $y_3 + y_5 = 1088$, find $y_4$.* *(Mumbai, B. Tech., 2004)*

**Sol.** Starting with $y_1$ instead of $y_0$, we note that $\Delta^6 y_1 = 0$ $\quad [\because \Delta^5 y_1$ is constant.

*i.e.,* $\quad (E-1)^6 y_1 = (E^6 - 6E^5 + 15E^4 - 20E^3 + 15E^2 - 6E + 1)\, y_1 = 0$



$\therefore \quad y_7 - 6y_6 + 15y_5 - 20y_4 + 15y_3 - 6y_2 + y_1 = 0$

or $\quad (y_7 + y_1) - 6(y_6 + y_2) + 15\,(y_5 + y_3) - 20\,y_4 = 0$

*i.e.,* $\quad y_4 = \frac{1}{20}\,[(y_1 + y_7) - 6(y_2 + y_6) + 15(y_3 + y_5)]$

$$= \frac{1}{20}\,[-784 - 6(686) + 15(1088)] = 571.$$

**Example 6.20.** *Using the method of separation of symbols, prove that*

*(i)* $u_1 x + u_2 x^2 + u_3 x^3 + \ldots\ldots = \frac{x}{1-x}\,u_1 + \left(\frac{x}{1-x}\right)^2 \Delta u_1 + \left(\frac{x}{1-x}\right)^3 \Delta^2 u_1 + \ldots$

*(ii)* $u_0 + \frac{u_1 x}{1!} + \frac{u_2 x^2}{2!} + \frac{u_3 x^3}{3!} + \ldots\ldots = e^x\left(u_0 + x\Delta u_0 + \frac{x^2}{2!}\Delta^2 u_0 + \frac{x^3}{3!}\Delta^3 u_0 + \ldots\ldots\right)$

**Sol.** *(i)* $\quad$ L.H.S. $= xu_1 + x^2 Eu_1 + x^3 E^2 u_1 + \ldots\ldots$ $\quad [\because u_{x+n} = E^n u_x$

$$= x(1 + xE + x^2E^2 + \ldots)u_1 = x \cdot \frac{1}{1 - xE}\,u_1, \text{ taking sum of infinite G.P.}$$

$$= x\left[\frac{1}{1 - x(1+\Delta)}\right] u_1 \qquad [\because E = 1 + \Delta$$

$$= x\left(\frac{1}{1 - x - x\Delta}\right) u_1 = \frac{x}{1-x}\left(1 - \frac{x\Delta}{1-x}\right)^{-1} u_1$$

$$= \frac{x}{1-x}\left(1 + \frac{x\Delta}{1-x} + \frac{x^2\Delta^2}{(1-x)^2} + \ldots\ldots\right) u_1$$

$$= \frac{x}{1-x}\,u_1 + \frac{x^2}{(1-x)^2}\,\Delta u_1 + \frac{x^3}{(1-x)^3}\,\Delta^2 u_1 + \ldots\ldots = \text{R.H.S.}$$

*(ii)* $\quad$ L.H.S. $= u_0 + \frac{x}{1!}\,Eu_0 + \frac{x^2}{2!}\,E^2 u_0 + \frac{x^3}{3!}\,E^3 u_0 + \ldots\ldots$

$$= \left(1 + \frac{xE}{1!} + \frac{x^2E^2}{2!} + \frac{x^3E^3}{3!} + \ldots\ldots\right) u_0 = e^{xE}\,u_0 = e^{x(1+\Delta)}u_0 = e^x \cdot e^{x\Delta} u_0$$

$$= e^x\left(1 + \frac{x\Delta}{1!} + \frac{x^2\Delta^2}{2!} + \frac{x^3\Delta^3}{3!} + \ldots\ldots\right) u_0$$

$$= e^x\left(u_0 + \frac{x}{1!}\,\Delta u_0 + \frac{x^2}{2!}\,\Delta^2 u_0 + \frac{x^3}{3!}\,\Delta^3 u_0 + \ldots\ldots\right) = \text{R.H.S.}$$

## PROBLEMS 6.3

1. Explain the difference between $\left(\dfrac{\Delta^2}{E}\right)u_x$ and $\left(\dfrac{\Delta^2 u_x}{Eu_x}\right)$. (*Madras B. Tech., 2003*)

2. Evaluate taking $h$ as the interval of differencing :

 (*i*) $\dfrac{\Delta^2}{E}\sin x$  (*ii*) $(\Delta + \nabla)^2(x^2 + x)$, $(h = 1)$.

 (*iii*) $\dfrac{\Delta^2 x^3}{Ex^3}$  (*iv*) $\dfrac{\Delta^2}{E}\sin(x+h) + \dfrac{\Delta^2 \sin(x+h)}{E\sin(x+h)}$.

3. With the usual notations, show that

 (*i*) $\nabla = 1 - e^{-hD}$  (*ii*) $D = \dfrac{2}{h}\sinh^{-1}\left(\dfrac{\delta}{2}\right)$ (*Madras, B.Tech., 2001*)

 (*iii*) $(1 + \Delta)(1 - \nabla) = 1$  (*iv*) $\Delta\nabla = \nabla\Delta = \delta^2$.

4. Prove that

 (*i*) $\delta = \Delta(1 + \Delta)^{-1/2} = \nabla(1 - \nabla)^{-1/2}$ (*Madras, B.E., 2001 S*)

 (*ii*) $\mu^2 = 1 + \dfrac{\delta^2}{2}$  (*iii*) $\delta(E^{1/2} + E^{-1/2}) = \Delta E^{-1} + \Delta$.

5. Show that

 (*i*) $\delta = \Delta E^{-1/2} = \nabla E^{1/2}$  (*ii*) $\mu\delta = \frac{1}{2}(\Delta + \nabla)$

 (*iii*) $1 + \delta^2/2 = \sqrt{(1+\delta^2\mu^2)}$ (*Madras, B. Tech., 2001*)

6. Show that

 (*i*) $\Delta = \mu\delta + \dfrac{\delta^2}{2}$  (*ii*) $E^{1/2} = \left(1 + \dfrac{\delta^2}{4}\right)^{1/2} + \dfrac{\delta^2}{2}$

 (*iii*) $E^r = (\mu + \frac{1}{2}\,\delta)^{2r}$  (*iv*) $\mu = \dfrac{2+\Delta}{2\lambda(1+\Delta)} = \dfrac{2-\nabla}{2\lambda(1-\nabla)}$

7. Prove that

 (*i*) $\Delta + \nabla = \dfrac{\Delta}{\nabla} - \dfrac{\nabla}{\Delta}$  (*ii*) $\nabla = \Delta E^{-1} = E^{-1}\Delta = 1 - E^{-1}$

 (*iii*) $E = \sum\limits_{i=0}^{\infty}\nabla_i$  (*iv*) $\nabla^2 = h^2D^2 - h^3D^3 + \dfrac{7}{12}h^4D^4 - \ldots$

8. Prove that $\delta^2 y_5 = y_6 - 2y_5 + y_4$.

9. Prove with usual notations, that

 (*i*) $\nabla^r f_k = \Delta^r f_{k-r}$  (*ii*) $\Delta(f_k^2) = (f_k + f_{k+1})\Delta f_k$ (*J.N.T.U., B. Tech., 2006*)

 (*iii*) $\sum\limits_{k=0}^{n-1}\Delta^2 f_k = \Delta f_n - \Delta f_0$.



10. Estimate the missing term in the following table :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 | 4 |
| $f(x)$ : | 1 | 3 | 9 | – | 81 |

(*S.V.T.U., B. Tech., 2007*)

11. Obtain the estimate of the missing figures in the following table :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $y$ : | 2 | 4 | 8 | ... | 32 | ... | 128 | 256 |

(*Mumbai, B.E., 2004*)

12. Find the missing values in the following table :

| | | | | | | |
|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 5 | 11 | 22 | 40 | ... | 140 | ... |

(*V.T.U., B.E., 2003*)

13. Estimate the production for 2004 and 2006 from the following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Year* : | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
| *Production* : | 200 | 200 | 260 | ... | 350 | ... | 430 |

14. If $u_{13} = 1$, $u_{14} = -3$, $u_{15} = -1$, $u_{16} = 13$, find $u_8$. (*Mumbai, B. Tech., 2004*)

15. Evaluate $y_4$ from the following data (stating the assumptions you make) :
 $y_0 + y_8 = 1.9243$, $y_1 + y_7 = 1.9590$
 $y_2 + y_6 = 1.9823$, $y_3 + y_5 = 1.9956$ (*Mumbai, B. Tech., 2003*)

Using the method of separation of symbols, prove that

16. $u_0 + u_1 + u_2 + \ldots + u_n = {}^{n+1}C_1u_0 + {}^{n+1}C_2\Delta u_0 + {}^{n+1}C_3\Delta^2u_0 + \ldots + {}^{n+1}C_{n+1}\Delta^n u_0$

17. $\Delta^n u_x = u_{x+n} - {}^nC_1u_{x+n-1} + {}^nC_2u_{x+n-2} + \ldots + (-1)^n u_x$.

18. $y_x = y_n - {}^{n-x}C_1\Delta y_{n-1} + {}^{n-x}C_2\Delta^2 y_{n-2} - \ldots + (-1)^{n-x}\Delta^{n-x}y_{n-(n-x)}$.

## 6.9. APPLICATION TO SUMMATION OF SERIES

The calculus of finite differences is very useful for finding the sum of a given series. The method is best illustrated by the following examples :

**Example 6.21.** *Sum the following series* $1^3 + 2^3 + 3^3 + \ldots + n^3$.

**Sol.** Denoting $1^3, 2^3, 3^3, \ldots$, by $u_0, u_1, u_2, \ldots$ respectively, the required sum

$$S = u_0 + u_1 + u_2 + \ldots + u_{n-1}$$
$$= (1 + E + E^2 + \ldots + E^{n-1})u_0 \qquad [\because\ u_1 = Eu_0,\ u_2 = E^2u_0 \text{ etc.}]$$
$$= \frac{E^n - 1}{E - 1}u_0 = \frac{(1+\Delta)^n - 1}{\Delta}u_0$$
$$= \frac{1}{\Delta}\left[1 + n\Delta + \frac{n(n-1)}{2!}\Delta^2 + \frac{n(n-1)(n-2)}{3!}\Delta^3 + \ldots + \Delta^n - 1\right]u_0$$
$$= n + \frac{n(n-1)}{2!}\Delta u_0 + \frac{n(n-1)(n-2)}{3!}\Delta^2 u_0 + \ldots$$

Now $\Delta u_0 = u_1 - u_0 = 2^3 - 1^3 = 7$, $\Delta^2 u_0 = u^2 - 2u_1 + u_0 = 3^3 - 2.2^3 + 1^3 = 12$,
$\Delta^3 u_0 = u_3 - 3u_2 + 3u_1 - u_0 = 4^3 - 3.3^3 + 3.2^3 - 1^3 = 6$
and $\Delta^4 u_0, \Delta^5 u_0, \ldots$ are all zero as $u_r = r^3$ is a polynomial of third degree.

Hence $S = n + \dfrac{n(n-1)}{2}\,.\,7 + \dfrac{n(n-1)(n-2)}{6}\,.\,12 + \dfrac{n(n-1)(n-2)(n-3)}{24}\,.\,6$

$$= \frac{n^2}{4}(n^2 + 2n + 1) = \left[\frac{n(n+1)}{2}\right]^2.$$

**Example 6.22.** *Prove that* $u_0 + u_1x + u_2x^2 + \ldots\ldots \infty = \dfrac{u_0}{1-x} + \dfrac{x\Delta u_0}{(1-x)^2} + \dfrac{x^2\Delta^2 u_0}{(1-x)^3} + \ldots\ldots \infty.$ *Hence find the sum of the series* $1.2 + 2.3x + 3.4x^2 + \ldots\ldots \infty.$

**Sol.** $u_0 + u_1x + u_2x^2 + \ldots\ldots \infty = (1 + xE + x^2E^2 + x^3E^3 + \ldots\ldots \infty)u_0$

$$= \frac{1}{1-xE}\, u_0 = \frac{1}{1-x(1+\Delta)}\, u_0$$

$$= \frac{1}{(1-x)-x\Delta}\, u_0 = \frac{1}{(1-x)}\left(1-\frac{x}{1-x}\Delta\right)^{-1} u_0$$

$$= \frac{1}{1-x}\left\{1-\frac{x\Delta}{1-x}+\frac{x^2\Delta^2}{(1-x^2)}+\ldots\ldots\right\} u_0$$

$$= \frac{u_0}{1-x} + \frac{x}{(1-x)^2}\,\Delta u_0 + \frac{x^2}{(1-x)^3}\,\Delta^2 u_0 + \ldots\ldots \infty$$

Now let us construct the difference table for the coefficients of the given series :

| $u$ | $\Delta u$ | $\Delta^2 u$ | $\Delta^3 u$ |
|---|---|---|---|
| $u_0 = 2$ | | | |
| | 4 | | |
| $u_1 = 6$ | | 2 | |
| | 6 | | 0 |
| $u_2 = 12$ | | 2 | |
| | 8 | | 0 |
| $u_3 = 20$ | | 2 | |
| | 10 | | |
| $u_4 = 30$ | | | |

This shows that $u_0 = 2$, $\Delta u_0 = 4$, $\Delta^2 u_0 = 2$, $\Delta^3 u_0 = \Delta^4 u_0$ etc. all $= 0$.

Thus $\quad 1.2 + 2.3x + 3.4x^2 + \ldots\ldots \infty$

$$= u_0 + u_1x + u_2x^2 + \ldots\ldots \infty$$

$$= \frac{u_0}{1-x} + \frac{x}{(1-x)^2}\,\Delta u_0 + \frac{x^2}{(1-x)^3}\,\Delta^2 u_0 + \ldots\ldots \infty$$

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

## PROBLEMS 6.4

Using the method of finite differences, sum the following series :

1. $1^2 + 2^2 + 3^2 + \ldots\ldots + n^2$.
2. $2.5 + 5.8 + 8.11 + 11.14 + \ldots\ldots$ $n$ terms.
3. $1.2.3 + 2.3.4 + 3.4.5 + \ldots\ldots$ to $n$ terms.



4. $\sum_{x=1}^{n} x(x+2)(x+4).$
5. Show that $u_0 + \dfrac{u_1x}{1!} + \dfrac{u_2x^2}{2!} + \ldots\ldots = e^x\left(u_0 + \dfrac{x}{1!}\Delta u_0 + \dfrac{x^2}{2!}\Delta^2 u_0 + \ldots\ldots\right).$

   Hence sum the series

   (*i*) $1^3 + \dfrac{2^3}{1!}x + \dfrac{3^3}{2!}x^2 + \dfrac{4^3}{3!}x^3 + \ldots\ldots\ldots \infty$

   (*ii*) $1 + \dfrac{4x}{1!} + \dfrac{10x^2}{2!} + \dfrac{20x^3}{3!} + \dfrac{35x^4}{4!} + \dfrac{56x^5}{5!} + \ldots\ldots \infty.$
6. Using the identity of Example 6.13, find the sum of the series $1.3 + 3.5x + 5.7x^2 + 7.9x^3 + \ldots\ldots \infty.$
7. Show that $\sum_{r=1}^{n} u_r = {}^nC_1\, u_1 + {}^nC_2\Delta u_1 + {}^nC_3\Delta^2 u_1 + \ldots\ldots + \Delta^{n-1}u_1$

   Hence evaluate $1^4 + 2^4 + 3^4 + \ldots\ldots + n^4$.
8. Sum the series $1.2\,\Delta x^n - 2.3\Delta^2 x^n + 3.4\Delta^3 x^n - 4.5\Delta^4 x^n + \ldots\ldots$ to $n$ terms.
9. Show that $\Delta x^n - \frac{1}{2}\Delta^2 x^n + \dfrac{1.3}{2.4}\Delta^3 x^n - \dfrac{1.3.5}{2.4.6}\Delta^4 x^n + \ldots\ldots$ to $n$ terms $= (x + 1/2)^n - (x - 1/2)^n$.

## 6.10. OBJECTIVE TYPE OF QUESTIONS

## PROBLEMS 6.5

*Select the correct answer or fill up the blanks in the following questions :*

1. $\Delta\nabla =$

   (*a*) $\nabla\Delta$ (*b*) $\nabla + \Delta$ (*c*) $\nabla - \Delta$.
2. Which one of the following results is correct :

   (*a*) $\Delta x^n = nx^{n-1}$ (*b*) $\Delta x^{(n)} = nx^{(n-1)}$

   (*c*) $\Delta^n e^x = e^x$ (*d*) $\Delta \cos x = -\sin x$.
3. If $f(x) = 3x^3 - 2x^2 + 1$, then $\Delta^3 f(x) = \ldots\ldots$ .
4. The relationship between the operators $E$ and $D$ is ...... .
5. The $(n + 1)$th order difference of the $n$th degree polynomial is ...... .
6. If $y(x) = x(x - 1)(x - 2)$, then $\Delta y(x) = \ldots\ldots$ .
7. $x^3 - 2x^2 + x - 1$ in factorial form $= \ldots\ldots$ .
8. Taking $h$ as the interval of differencing, $\Delta^2 x^3 = \ldots\ldots$ .
9. In terms of $E$, $\nabla = \ldots\ldots$ .
10. The form of the function tabulated at equally spaced intervals with sixth differences constant, is ...... .
11. If the interval of differencing is unity, then $\Delta^4[(1-x)(1-2x)(1-3x)] = \ldots\ldots$
12. Taking the interval of differencing as unity, the first difference of $x^4 - 3x^3 + 2x - 1$ is ...... .

13. The missing values of $y$ in the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| $y_x$ | : | 0 | ... | ... | ... | 25 |
| $\Delta y_x$ | : | 1 | 2 | 4 | 7 | 11, are ...... . |

14. $\Delta^3 [(1-x)(1-3x)(1-5x)]$ = ........ . (interval of differencing being 1)
15. $\Delta \tan^{-1} x$ = ....... .
16. If $y = x^2 - 2x + 2$, taking interval of differencing as unity, $\Delta^2 y$ = ...... .
17. Relation between $\Delta$ and $E$ is given by ....... .
18. The $k$th difference of a polynomial of degree $k$ is ..... .
19. $\Delta^r y_k$ in terms of backward differences = ...... .
20. The value of $(\Delta^2/E)e^x$ = ...... .
21. The relation between the shift operator $E$ and second order backward difference operator $\nabla^2$ is ..... .
22. The value of $\Delta^n(e^x)$ = ........ (interval of differencing being 1).
23. Relationship between $E$, $\Delta$ and $\nabla$ is ..... .
24. If the 5th and higher order differences of a function vanish, then the function represents a polynomial of degree ...... .
25. The value of $E^{-1} \nabla$ = ...... .
26. If $E^2 u_x = x^2$ and $h = 1$, then $u_x$ = ...... .
27. Given $y_0 = 2, y_1 = 4, y_2 = 8, y_4 = 32$, then $y_3$ = ...... .
28. $y_0 = 1, y_1 = 5, y_2 = 8, y_3 = 3, y_4 = 7, y_5 = 0$, then $\Delta^5 y_0$ =
    (*a*) 61 (*b*) $-62$
    (*c*) 62 (*d*) $-61$.
29. Given $x = 1 \quad 2 \quad 3$
    $f(x) = 3 \quad 8 \quad 15$, then $\Delta^2 f(1)$ =
    (*a*) 3 (*b*) 4
    (*c*) 2 (*d*) 1.
30. $(E^{1/2} + E^{-1/2})(1 + \Delta)^{1/2}$ =
    (*a*) $\Delta + 1$ (*b*) $\Delta - 1$
    (*c*) $\Delta + 2$ (*d*) $\Delta - 2$.
31. Which one is incorrect ?
    (*a*) $E = 1 + \Delta$ (*b*) $\Delta(5) = 0$
    (*c*) $\Delta(f_1 + f_2) = \Delta f_1 + \Delta f_2$ (*d*) $\Delta(f_1 \cdot f_2) = \Delta f_1 + \Delta f_2$.
32. $\Delta - \nabla = \delta^2$. (True or False)
33. $\Delta + \nabla = E + E^{-1}$. (True or False)
34. $E = e^{-hD}$. (True or False)
35. If $f(x) = e^x$, then $\Delta^6 e^x = (e^h - 1)^6 e^x$. (True or False)
36. $\Delta^n = \delta^n E^{n/2}$. (True or False)
37. $(1 + \Delta)(1 - \nabla) = 1$. (True or False)
38. With the usual notations, match the items on right hand side with those in left hand side :
    (*i*) $E\nabla$ (*a*) $\frac{1}{2}(\Delta + \nabla)$
    (*ii*) $hD$ (*b*) $\Delta - \nabla$
    (*iii*) $\nabla\Delta$ (*c*) $\Delta$
    (*iv*) $\mu\delta$ (*d*) $-\log(1 - \nabla)$ (*Bhopal, B.E., 2007*)

# 7

# INTERPOLATION



## 7.1. INTRODUCTION

Suppose we are given the following values of $y = f(x)$ for a set of values of $x$ :

| | | | |
|---|---|---|---|
| $x$ : | $x_0$ | $x_1$ | $x_2$ ........ $x_n$ |
| $y$ : | $y_0$ | $y_1$ | $y_2$ ........ $y_n$. |

Then the process of finding the value of $y$ corresponding to any value of $x = x_i$ between $x_0$ and $x_n$ is called *interpolation*. Thus *interpolation is the technique of estimating the value of a function for any intermediate value of the independent variable* while the process of computing the value of the function outside the given range is called *extrapolation*. The term interpolation however, is taken to include extrapolation.

If the function $f(x)$ is known explicitly, then the value of $y$ corresponding to any value of $x$ can easily be found. Conversely, if the form of $f(x)$ is not known (as is the case in most of the applications), it is very difficult to determine the exact form of $f(x)$ with the help of tabulated set of values $(x_i, y_i)$. In such cases, $f(x)$ is replaced by a simpler function $\phi(x)$ which assumes the same values as those of $f(x)$ at the tabulated set of points. Any other value may be calculated from $\phi(x)$ which is known as the *interpolating function* or *smoothing function*. If $\phi(x)$ is a polynomial, then it called the *interpolating polynomial* and the process is called the *polynomial interpolation*. Similarly when $\phi(x)$ is a finite trigonometric series, we have trigonometric interpolation. But we shall confine ourselves to polynomial interpolation only.

The study of interpolation is based on the calculus of finite differences. We begin by deriving two important *interpolation formulae* by means of forward and backward differences of a function. These formulae are often employed in engineering and scientific investigations.

## 7.2. NEWTON'S FORWARD INTERPOLATION FORMULA

Let the function $y = f(x)$ take the values $y_0, y_1, \ldots\ldots, y_n$ corresponding to the values $x_0, x_1, \ldots\ldots, x_n$ of $x$. Let these values of $x$ be equi-spaced such that $x_i = x_0 + ih$ $(i = 0, 1, \ldots\ldots)$. Assuming $y(x)$ to be a polynomial of the $n$th degree in $x$ such that $y(x_0) = y_0, y(x_1) = y_1, \ldots\ldots, y(x_n) = y_n$. We can write

$$y(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2) + \ldots\ldots + a_n(x - x_0)(x - x_1) \ldots (x - x_{n-1}) \quad \ldots(1)$$

Putting $x = x_0, x_1, \ldots\ldots, x_n$ successively in (1), we get

$$y_0 = a_0,\ y_1 = a_0 + a_1(x_1 - x_0),\ y_2 = a_0 + a_1(x_2 - x_0) + a_2(x_2 - x_0)(x_2 - x_1)$$

and so on.

From these, we find that $a_0 = y_0$, $\Delta y_0 = y_1 - y_0 = a_1(x_1 - x_0) = a_1 h$

$$\therefore \quad a_1 = \frac{1}{h} \Delta y_0$$

Also $\quad \Delta y_1 = y_2 - y_1 = a_1(x_2 - x_1) + a_2(x_2 - x_0)(x_2 - x_1) = a_1 h + a_2 2h.h = \Delta y_0 + 2h^2 a_2$

$$\therefore \quad \boxed{a_2 = \frac{1}{2h^2}(\Delta y_1 - \Delta y_0)} = \frac{1}{2!\,h^2} \Delta^2 y_0$$

Similarly $a_3 = \dfrac{1}{3!\,h^3} \Delta^3 y_0$ and so on.

Substituting these values in (1), we obtain

$$y(x) = y_0 + \frac{\Delta y_0}{h}(x - x_0) + \frac{\Delta^2 y_0}{2!\,h^2}(x - x_0)(x - x_1) + \frac{\Delta^3 y_0}{3!\,h^3}(x - x_0)(x - x_1)(x - x_2) + \ldots\ldots \quad \ldots(2)$$

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

$$\mathbf{y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \ldots\ldots + \frac{p(p-1)\ldots(p-\overline{n-1})}{n!}\Delta^n y_0} \quad \ldots(3)$$

It is called *Newton's forward interpolation formula* as (3) contains $y_0$ and the forward differences of $y_0$.

**Otherwise :** Let the function $y = f(x)$ take the values $y_0, y_1, y_2, \ldots\ldots$ corresponding to the values $x_0, x_0 + h, x_0 + 2h, \ldots\ldots$ of $x$. Suppose it is required to evaluate $f(x)$ for $x = x_0 + ph$, where $p$ is any real number.

For any real number $p$, we have defined $E$ such that

$$E^p f(x) = f(x + ph)$$
$$y_p = f(x_0 + ph) = E^p f(x_0) = (1 + \Delta)^p y_0 \qquad [\because E = 1 + \Delta$$



$$= \left\{1 + p\Delta + \frac{p(p-1)}{2!}\Delta^2 + \frac{p(p-1)(p-2)}{3!}\Delta^3 + \ldots\ldots\right\} y_0 \quad \ldots(4)$$

[using Binomial theorem

i.e. $$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \ldots\ldots$$

If $y = f(x)$ is a polynomial of the $n$th degree, then $\Delta^{n+1} y_0$ and higher differences will be zero. Hence (4) will become

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \ldots\ldots + \frac{p(p-1)(p-\overline{n-1})}{n!}\Delta^n y_0$$

which is same as (3).

**Obs.** *This formula is used for interpolating the values of y near the beginning of a set of tabulated values and extrapolating values of y a little backward (i.e. to the left) of $y_0$.*

## 7.3. NEWTON'S BACKWARD INTERPOLATION FORMULA

Let the function $y = f(x)$ take the values $y_0, y_1, y_2, \ldots\ldots$ corresponding to the values $x_0, x_0 + h, x_0 + 2h, \ldots\ldots$ of $x$. Suppose it is required to evaluate $f(x)$ for $x = x_n + ph$, where $p$ is any real number. Then we have

$$y_p = f(x_n + ph) = E^p f(x_n) = (1 - \nabla)^{-p} y_n \qquad [\because E^{-1} = 1 - \nabla$$
$$= \left[1 + p\nabla + \frac{p(p+1)}{2!}\nabla^2 + \frac{p(p+1)(p+2)}{3!}\nabla^3 + \ldots\ldots\right] y_n$$

[using Binomial theorem

i.e. $$\mathbf{y_p = y_n + p\nabla y_n + \frac{p(p+1)}{2!}\nabla^2 y_n + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_n + \ldots} \quad \ldots(1)$$

It is called *Newton's backward interpolation formula* as (1) contains $y_n$ and backward differences of $y_n$.

**Obs.** *This formula is used for interpolating the values of y near the end of a set of tabulated values and also for extrapolating values of y a little ahead* (to the right) *of $y_n$.*

**Example 7.1.** *The table gives the distance in nautical miles of the visible horizon for the given heights in feet above the earth's surface :*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *x = height :* | 100 | 150 | 200 | 250 | 300 | 350 | 400 |
| *y = distance :* | 10.63 | 13.03 | 15.04 | 16.81 | 18.42 | 19.90 | 21.27 |

*Find the values of y when*

*(i) x = 218 ft*

*(ii) x = 410.* (Madras, B.E., 2003 S)

(V.T.U., B.E., 2002)

**Sol.** The difference table is as under :

$x$: $x_0$

| $x$ | $y$ | $\Delta$ | $\Delta^2$ | $\Delta^3$ | $\Delta^4$ |
|---|---|---|---|---|---|
| 100 | 10.63 | | | | |
| | | 2.40 | | | |
| 150 | 13.03 | | −0.39 | | |
| | | 2.01 | | 0.15 | |
| 200 | **15.04** | | −0.24 | | −0.07 |
| | | **1.77** | | 0.08 | |
| 250 | 16.81 | | **−0.16** | | −0.05 |
| | | 1.61 | | **0.03** | |
| 300 | 18.42 | | −0.13 | | **−0.01** |
| | | 1.48 | | **0.02** | |
| 350 | 19.90 | | **−0.11** | | |
| | | **1.37** | | | |
| 400 | **21.27** | | | | |

(*i*) If we take $x_0 = 200$, then $y_0 = 15.04$, $\Delta y_0 = 1.77$, $\Delta^2 y_0 = -0.16$, $\Delta^3 = 0.03$ etc.

Since $x = 218$ and $h = 50$, $\therefore\quad p = \dfrac{x - x_0}{h} = \dfrac{18}{50} = 0.36$

$\therefore$ Using Newton's forward interpolation formula, we get

$$y_{218} = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \frac{p(p-1)(p-2)(p-3)}{4!}\Delta^4 y_0 + \ldots$$

$\left(\frac{x - x_n}{h}\right)$

$$f(218) = 15.04 + 0.36(1.77) + \frac{0.36(-0.64)}{2}(-0.16) + \frac{0.36(-0.64)(-1.64)}{6}(0.03) + \frac{0.36(-0.64)(-1.64)(-2.64)}{24}(-0.01)$$

$$= 15.04 + 0.637 + 0.018 + 0.002 + 0.0004$$

$$= 15.697 \text{ i.e. } 15.7 \text{ nautical miles.}$$

(*ii*) Since $x = 410$ is near the end of the table, we use Newton's backward interpolation formula.

$\therefore$ Taking $x_n = 400$, $p = \dfrac{x - x_n}{h} = \dfrac{10}{50} = 0.2$

Using the line of backward difference

$$y_n = 21.27,\ \nabla y_n = 1.37,\ \nabla^2 y_n = -0.11,\ \nabla^3 y_n = 0.02 \text{ etc.}$$

$\therefore$ Newton's backward formula gives

$$y_{410} = y_{400} + p\nabla y_{400} + \frac{p(p+1)}{2!}\nabla^2 y_{400} + \frac{p(p+1)(p+2)}{3!}\nabla^3 y_{400} + \frac{p(p+1)(p+2)(p+3)}{4!}\nabla^4 y_{400}$$



$$= 21.27 + 0.2(1.37) + \frac{0.2(1.2)}{2!}(-0.11) + \frac{0.2(1.2)(2.2)}{3!}(0.02) + \frac{0.2(1.2)(2.2)(3.2)}{4!}(-0.01)$$

$$= 21.27 + 0.274 - 0.0132 + 0.0018 - 0.0007$$

$$= 21.53 \text{ nautical miles.}$$

**Example 7.2.** *From the following table, estimate the number of students who obtained marks between 40 and 45 :*

| *Marks* : | *30—40* | *40—50* | *50—60* | *60—70* | *70—80* |
|---|---|---|---|---|---|
| *No. of students* : | *31* | *42* | *51* | *35* | *31* |

(V.T.U., B. Tech., 2007)

**Sol.** First we prepare the cumulative frequency table, as follows :

| Marks less than ($x$) : | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|
| No. of students ($y_x$) : | 31 | 73 | 124 | 159 | 190 |

Now the difference table is

| $x$ | $y_x$ | $\Delta y_x$ | $\Delta^2 y_x$ | $\Delta^3 y_x$ | $\Delta^4 y_x$ |
|---|---|---|---|---|---|
| 40 | **31** | | | | |
| | | **42** | | | |
| 50 | 73 | | **9** | | |
| | | 51 | | **−25** | |
| 60 | 124 | | −16 | | **37** |
| | | 35 | | 12 | |
| 70 | 159 | | −4 | | |
| | | 31 | | | |
| 80 | 190 | | | | |

We shall find $y_{45}$ *i.e.* number of students with marks less than 45. Taking $x_0 = 40$, $x = 45$, we have

$$p = \frac{x - x_0}{h} = \frac{5}{10} = 0.5 \qquad [\because h = 10$$

$\therefore$ Using Newton's forward interpolation formula, we get

$$y_{45} = y_{40} + p\,\Delta y_{40} + \frac{p(p-1)}{2!}\Delta^2 y_{40} + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_{40} + \frac{p(p-1)(p-2)(p-3)}{3!}\Delta^4 y_{40}$$

$$= 31 + 0.5 \times 42 + \frac{0.5(-0.5)}{2} \times 9 + \frac{0.5(-0.5)(-1.5)}{6} \times (-25) + \frac{0.5(-0.5)(-1.5)(-2.5)}{24} \times 37$$

$$= 31 + 21 - 1.125 - 1.5625 - 1.4453$$

$$= 47.87, \text{ on simplification.}$$

The number of students with marks less than 45 is 47.87 *i.e.*, 48.

But the number of students with marks less than 40 is 31.

Hence the number of students getting marks between 40 and 45 = 48 – 31 = 17.

**Example 7.3.** *Find the cubic polynomial which takes the following values :*

| | | | | |
|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 |
| $f(x)$ : | 1 | 2 | 1 | 10 |

*Hence or otherwise evaluate f(4).* *(Bhopal, B.E., 2009)*

**Sol.** The difference table is

| $x$ | $f(x)$ | $\Delta f(x)$ | $\Delta^2 f(x)$ | $\Delta^3 f(x)$ |
|---|---|---|---|---|
| 0 | 1 | | | |
| | | 1 | | |
| 1 | 2 | | – 2 | |
| | | – 1 | | 12 |
| 2 | 1 | | 10 | |
| | | 9 | | |
| 3 | 10 | | | |

We take $x_0 = 0$ and $p = \dfrac{x - 0}{h} = x$ $[\because h = 1$

$\therefore$ Using Newton's forward interpolation formula, we get

$$f(x) = f(0) + \frac{x}{1} \Delta f(0) + \frac{x(x-1)}{1.2} \Delta^2 f(0) + \frac{x(x-1)(x-2)}{1.2.3} \Delta^3 f(0)$$

$$= 1 + x(1) + \frac{x(x-1)}{2}(-2) + \frac{x(x-1)(x-2)}{6}(12)$$

$$= 2x^3 - 7x^2 + 6x + 1,$$

which is the required polynomial.

To compute $f(4)$, we take $x_n = 3$, $x = 4$ so that $p = \dfrac{x - x_n}{h} = 1$ $[\because h = 1$

**Obs.** Using Newton's backward interpolation formula, we get

$$f(4) = f(3) + p\nabla f(3) + \frac{p(p+1)}{1.2} \nabla^2 f(3) + \frac{p(p+1)(p+2)}{1.2.3} \nabla^3 f(3)$$

$$= 10 + 9 + 10 + 12 = 41$$

which is the same value as that obtained by substituting $x = 4$ in the cubic polynomial above.

The above example shows that *if a tabulated function is a polynomial, then interpolation and extrapolation give the same values.*

**Example 7.4.** *Using Newton's backward difference formula, construct an interpolating polynomial of degree 3 for the data : f(– 0.75) = – 0.0718125, f(– 0.5) = – 0.02475, f(– 0.25) = 0.3349375, f(0) = 1.10100. Hence find f(– 1/3).* *(Anna, B.Tech., 2003)*



**Sol.** The difference table is

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ |
|---|---|---|---|---|
| – 0.75 | – 0.0718125 | | | |
| | | 0.0470625 | | |
| – 0.5 | – 0.02475 | | 0.312625 | |
| | | 0.3596875 | | |
| – 0.25 | 0.3349375 | | 0.400375 | 0.09375 |
| | | 0.7660625 | | |
| 0 | 1.10100 | | | |

We use Newton's backward difference formula

$$y(x) = y_3 + \frac{p}{1!} \nabla y_3 + \frac{p(p+1)}{2!} \nabla^2 y_3 + \frac{p(p+1)(p+2)}{3!} \nabla^3 y_3$$

taking $x_3 = 0, \; p = \dfrac{x - 0}{h} = \dfrac{x}{0.25} = 4x,$ $[\because h = 0.25]$

$$\therefore \quad y(x) = 1.10100 + 4x(0.7660625) + \frac{4x(4x+1)}{2}(0.400375) + \frac{4x(4x+1)(4x+2)}{6}(0.09375)$$

$$= 1.101 + 3.06425x + 3.251x^2 + 0.81275x + x^3 + 0.75x^2 + 0.125x$$

$$= x^3 + 4.001x^2 + 4.002x + 1.101.$$

Put $x = -\dfrac{1}{3}$, so that

$$y\left(-\frac{1}{3}\right) = \left(-\frac{1}{3}\right)^3 + 4.001\left(-\frac{1}{3}\right)^2 + 4.002\left(-\frac{1}{3}\right) + 1.101$$

$$= 0.1745$$

**Example 7.5.** *In the table below, the values of y are consecutive terms of a series of which 23.6 is the 6th term. Find the first and tenth terms of the series :*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $y$ : | 4.8 | 8.4 | 14.5 | 23.6 | 36.2 | 52.8 | 73.9 |

*(Anna, B.E., 2007)*

**Sol.** The difference table is

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|---|---|---|---|---|---|
| 3 | 4.8 | | | | |
| | | 3.6 | | | |
| 4 | 8.4 | | 2.5 | | |
| | | 6.1 | | 0.5 | |
| 5 | 14.5 | | 3.0 | | 0 |
| | | 9.1 | | 0.5 | |

| | | | | | |
|---|---|---|---|---|---|
| 6 | 23.6 | | 3.5 | | |
| | | 12.6 | | 0.5 | 0 |
| 7 | 36.2 | | 4.0 | | |
| | | 16.6 | | 0.5 | 0 |
| 8 | 52.8 | | 4.5 | | |
| | | 21.1 | | | |
| 9 | 73.9 | | | | |

To find the first term, use Newton's forward interpolation formula with $x_0 = 3$, $x = 1$, $h = 1$ and $p = -2$. We have

$$y(1) = 4.8 + \frac{(-2)}{1} \times 3.6 + \frac{(-2)(-3)}{1.2} \times 2.5 + \frac{(-2)(-3)(-4)}{1.2.3} \times 0.5 = 3.1$$

To obtain the tenth term, use Newton's backward interpolation formula with $x_n = 9$, $x = 10$, $h = 1$ and $p = 1$. This gives

$$y(10) = 73.9 + \frac{1}{1} \times 21.1 + \frac{1(2)}{1.2} \times 4.5 + \frac{1(2)(3)}{1.2.3} \times 0.5 = 100.$$

**Example 7.6.** *Using Newton's forward interpolation formula, show that*

$$\Sigma n^3 = \left[\frac{n(n+1)}{2}\right]^2.$$

**Sol.** If $s_n = \sum n^3$, then $s_{n+1} = \Sigma(n+1)^3$

$\therefore$ $\Delta s_n = s_{n+1} - s_n = \Sigma(n+1)^3 - \Sigma n^3 = (n+1)^3$.

Then $\Delta^2 s_n = \Delta s_{n+1} - \Delta s_n = (n+2)^3 - (n+1)^3 = 3n^2 + 9n + 7$

$\Delta^3 s_n = \Delta^2 s_{n+1} - \Delta^2 s_n$

$= [3(n+1)^2 + 9(n+1) + 7] - (3n^2 + 9n + 7) = 6n + 12.$

$\Delta^4 s_n = \Delta^3 s_{n+1} - \Delta^3 s_n = [6(n+1) + 12] - [6n + 12] = 6$

and $\Delta^5 s_n = \Delta^5 \Delta_n = \ldots\ldots = 0.$

Since the first term of the given series is 1, therefore taking $n = 1$, $s_1 = 1$, $\Delta s_1 = 8$, $\Delta^2 s_1 = 19$, $\Delta^3 s_1 = 18$, $\Delta^4 s_1 = 6$.

Substituting these in the Newton's forward interpolation formula *i.e.*

$$s_n = s_1 + (n-1)\Delta s_1 + \frac{(n-1)(n-2)}{2!}\Delta^2 s_1 + \frac{(n-1)(n-2)(n-3)}{3!}\Delta^3 s_1 + \frac{(n-1)(n-2)(n-3)(n-4)}{4!}\Delta^4 s_1$$

$$s_n = 1 + 8(n-1) + \frac{19}{2}(n-1)(n-2) + 3(n-1)(n-2)(n-3) + \frac{1}{4}(n-1)(n-2)(n-3)(n-4) = \frac{1}{4}(n^4 + 2n^3 + n^2) = \left\{\frac{n(n+1)}{2}\right\}^2.$$



## PROBLEMS 7.1

1. 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 |

(J.N.T.U., B.Tech., 2008)

2. From the following table, find $y$ when $x = 1.85$ and $2.4$ by Newton's interpolation formulae :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 1.7 | 1.8 | 1.9 | 2.0 | 2.1 | 2.2 | 2.3 |
| $y = e^x$ : | 5.474 | 6.050 | 6.686 | 7.389 | 8.166 | 9.025 | 9.974 |

(Kottayam, B.E., 2005)

3. If $f(1.15) = 1.0723$, $f(1.20) = 1.0954$, $f(1.25) = 1.1180$ and $f(1.30) = 1.1401$, find $f(1.28)$.

4. Given $\sin 45° = 0.7071$, $\sin 50° = 0.7660$, $\sin 55° = 0.8192$, $\sin 60° = 0.8660$, find $\sin 52°$ using Newton's forward formula. (J.N.T.U., B.Tech., 2006)

5. From the following table :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 |
| $f(x)$ : | 2.68 | 3.04 | 3.38 | 3.68 | 3.96 | 4.21 |

find $f(0.7)$ approximately. (V.T.U., B.E., 2001)

6. 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 the area of a circle of diameter 105. (V.T.U., B.E., 2004)

7. From the following table :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $x°$ : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| $\cos x$ : | .9848 | .9397 | .8660 | .7660 | .6428 | .5000 | .3420 | .1737 |

Calculate $\cos 25°$ and $\cos 73°$ using Gregory-Newton formula. (U.P.T.U., B.Tech., 2006)

8. A test performed on $NPN$ transistor gives the following result :

| Base current $I$ (mA) | 0 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 |
|---|---|---|---|---|---|---|
| Collector current $I_c$ (mA) | 0 | 1.2 | 2.5 | 3.6 | 4.3 | 5.34 |

Calculate (*i*) the value of the collector current for the base current of 0.005 mA.
(*ii*) the value of base current required for a collector current of 4.0 mA.
(Pune, B.Tech., 2004)

9. Find $f(22)$ from the following data using Newton's backward formulae.

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 20 | 25 | 30 | 35 | 40 | 45 |
| $f(x)$ : | 354 | 332 | 291 | 260 | 231 | 204 |

(J.N.T.U., B.Tech., 2007)

10. Find the number of men getting wages between Rs. 10 and 15 from the following data :

| | | | | |
|---|---|---|---|---|
| Wages in Rs. : | 0—10 | 10—20 | 20—30 | 30—40 |
| Frequency : | 9 | 30 | 35 | 42 |

(Nagarjuna, B.E., 2001)

11. From the following data, estimate the number of persons having incomes between 2000 and 2500 :

| Income | Below 500 | 500–1000 | 1000–2000 | 2000–3000 | 3000–4000 |
|---|---|---|---|---|---|
| No. of persons | 6000 | 4250 | 3600 | 1500 | 650 |

**12.** Construct Newton's forward interpolation polynomial for the following data :

| | | | | |
|---|---|---|---|---|
| $x$ : | 4 | 6 | 8 | 10 |
| $y$ : | 1 | 3 | 8 | 16 |

Hence evaluate $y$ for $x = 5$. (*Madras, B. Tech., 2006*)

**13.** Find the cubic polynomial which takes the following values :

$y(0) = 1, y(1) = 0, y(2) = 1$ and $y(3) = 10$.

Hence or otherwise, obtain $y(4)$. (*U.P.T.U., B. Tech., 2005*)

**14.** Construct the difference table for the following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 0.1 | 0.3 | 0.5 | 0.7 | 0.9 | 1.1 | 1.3 |
| $f(x)$ : | 0.003 | 0.067 | 0.148 | 0.248 | 0.370 | 0.518 | 0.697 |

Evaluate $f(0.6)$ (*J.N.T.U., B. Tech., 2007*)

**15.** Apply Newton's backward difference formula to the data below, to obtain a polynomial of degree 4 in $x$ :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 |
| $y$ : | 1 | $-1$ | 1 | $-1$ | 1 |

**16.** The following table gives the population of a town during the last six censuses. Estimate the increase in the population during the period from 1976 to 1978 :

| | | | | | | |
|---|---|---|---|---|---|---|
| Year : | 1941 | 1951 | 1961 | 1971 | 1981 | 1991 |
| Population : (in thousands) | 12 | 15 | 20 | 27 | 39 | 52 |

(*U.P.T.C., B.E., 2009*)

**17.** In the following table, the values of $y$ are consecutive terms of a series of which 12.5 is the 5th term. Find the first and tenth terms of the series.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $y$ : | 2.7 | 6.4 | 12.5 | 21.6 | 34.3 | 51.2 | 72.9 |

(*P.T.U., B. Tech., 2001*)

**18.** Using a polynomial of the third degree, complete the record given below of the export of a certain commodity during five years :

| | | | | | |
|---|---|---|---|---|---|
| Year : | 1989 | 1990 | 1991 | 1992 | 1993 |
| Export : (in tons) | 443 | 384 | — | 397 | 467 |

**19.** Given $u_1 = 40, u_3 = 45, u_5 = 54$, find $u_2$ and $u_4$. (*Nagarjuna, B.E., 2003 S*)

**20.** If $u_{-1} = 10, u_1 = 8, u_2 = 10, u_4 = 50$, find $u_0$ and $u_3$.

**21.** Given $y_0 = 3, y_1 = 12, y_2 = 81, y_3 = 200, y_4 = 100, y_5 = 8$, without forming the difference table, find $\Delta^5 y_0$.

## 7.4. CENTRAL DIFFERENCE INTERPOLATION FORMULAE

In the preceding sections, we derived Newton's forward and backward interpolation formulae which are applicable for interpolation near the beginning and end of tabulated values. Now we shall develop central difference formulae which are best suited for interpolation near the middle of the table.

If $x$ takes the values $x_0 - 2h, x_0 - h, x_0, x_0 + h, x_0 + 2h$ and the corresponding values of $y = f(x)$ are $y_{-2}, y_{-1}, y_0, y_1, y_2$, then we can write the difference table in the two notations as follows :



| $x$ | $y$ | 1st diff. | 2nd diff. | 3rd diff. | 4th diff. |
|---|---|---|---|---|---|
| $x_0 - 2h$ | $y_{-2}$ | | | | |
| | | $\Delta y_{-2} (= \delta y_{-3/2})$ | | | |
| $x_0 - h$ | $y_{-1}$ | | $\Delta^2 y_{-2} (= \delta^2 y_{-1})$ | | |
| | | $\Delta y_{-1} (= \delta y_{-1/2})$ | | $\Delta^3 y_{-2} (= \delta^3 y_{-1/2})$ | |
| $x_0$ | $y_0$ | | $\Delta^2 y_{-1} (= \delta^2 y_0)$ | | $\Delta^4 y_{-2} (= \delta^4 y_0)$ |
| | | $\Delta y_0 (= \delta y_{1/2})$ | | $\Delta^3 y_{-1} (= \delta^3 y_{1/2})$ | |
| $x_0 + h$ | $y_1$ | | $\Delta^2 y_0 (= \delta^2 y_1)$ | | |
| | | $\Delta y_1 (= \delta y_{3/2})$ | | | |
| $x_0 + 2h$ | $y_2$ | | | | |

## 7.5. GAUSS'S FORWARD INTERPOLATION FORMULA

The Newton's forward interpolation formula is

$$y_p = y_0 + p\,\Delta y_0 + \frac{p(p-1)}{1.2}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{1.2.3}\Delta^3 y_0 + \ldots \qquad \ldots(1)$$

We have $\Delta^2 y_0 - \Delta^2 y_{-1} = \Delta^3 y_{-1}$

*i.e.* $\qquad \Delta^2 y_0 = \Delta^2 y_{-1} + \Delta^3 y_{-1}$ $\qquad \ldots(2)$

Similarly $\qquad \Delta^3 y_0 = \Delta^3 y_{-1} + \Delta^4 y_{-1}$ $\qquad \ldots(3)$

$\Delta^4 y_0 = \Delta^4 y_{-1} + \Delta^5 y_{-1}$ etc. $\qquad \ldots(4)$

Also $\quad \Delta^3 y_{-1} - \Delta^3 y_{-2} = \Delta^4 y_{-2}$

*i.e.* $\qquad \Delta^3 y_{-1} = \Delta^3 y_{-2} + \Delta^4 y_{-2}$

Similarly $\qquad \Delta^4 y_{-1} = \Delta^4 y_{-2} + \Delta^5 y_{-2}$ etc. $\qquad \ldots(5)$

Substituting for $\Delta^2 y_0, \Delta^3 y_0, \Delta^4 y_0 \ldots$ from (2), (3), (4) ..... in (1), we get

$$y_p = y_0 + p\,\Delta y_0 + \frac{p(p-1)}{1.2}(\Delta^2 y_{-1} + \Delta^3 y_{-1}) + \frac{p(p-1)(p-2)}{1.2.3}(\Delta^3 y_{-1} + \Delta^4 y_{-1})$$
$$+ \frac{p(p-1)(p-2)(p-3)}{1.2.3.4} \times (\Delta^4 y_{-1} + \Delta^5 y_{-1}) + \ldots$$

Hence
$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-1}$$
$$+ \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \ldots \quad [\text{using } (5)$$

which is called *Gauss's forward interpolation formula.*

**Cor.** *In the central differences notation, this formula will be*

$$y_p = y_0 + p\delta y_{1/2} + \frac{p(p-1)}{2!}\delta^2 y_0 + \frac{(p+1)p(p-1)}{3!}\delta^3 y_{1/2} + \frac{(p+1)p(p-1)(p-2)}{4!}\delta^4 y_0 + \ldots$$

**Obs. 1.** *It employs odd differences just below the central line and even difference on the central line as shown below :*

$$y_0 \;\cdots\cdots\; \Delta^2 y_{-1} \;\cdots\cdots\; \Delta^4 y_{-2} \;\cdots\cdots\; \Delta^6 y_{-3} \;\cdots\cdots \text{ Central line}$$
$$\Delta y_0 \qquad \Delta^3 y_{-1} \qquad \Delta^5 y_{-2} \qquad \Delta^7 y_{-3}$$

**Obs. 2.** This formula is used to interpolate the values of $y$ for $p$ $(0 < p < 1)$ measured forwardly from the origin.

## 7.6. GAUSS'S BACKWARD INTERPOLATION FORMULA

The Newton's forward interpolation formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{1.2}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{1.2.3}\Delta^3 y_0 + \dots \quad \dots(1)$$

We have $\Delta y_0 - \Delta y_{-1} = \Delta^2 y_{-1}$

*i.e.* $\Delta y_0 = \Delta y_{-1} + \Delta^2 y_{-1}$ ...(2)

Similarly $\Delta^2 y_0 = \Delta^2 y_{-1} + \Delta^3 y_{-1}$ ...(3)

$\Delta^3 y_0 = \Delta^3 y_{-1} + \Delta^4 y_{-1}$ etc. ...(4)

Also $\Delta^3 y_{-1} - \Delta^3 y_{-2} = \Delta^4 y_{-2}$

*i.e.* $\Delta^3 y_{-1} = \Delta^3 y_{-2} + \Delta^4 y_{-2}$ ...(5)

Similarly $\Delta^4 y_{-1} = \Delta^4 y_{-2} + \Delta^5 y_{-2}$ etc. ...(6)

Substituting for $\Delta y_0, \Delta^2 y_0, \Delta^3 y_0, \dots$ from (2), (3), (4) in (1), we get

$$y_p = y_0 + p(\Delta y_{-1} + \Delta^2 y_{-1}) + \frac{p(p-1)}{1.2}(\Delta^2 y_{-1} + \Delta^3 y_{-1})$$
$$+ \frac{p(p-1)(p-2)}{1.2.3}(\Delta^3 y_{-1} + \Delta^4 y_{-1}) + \frac{p(p-1)(p-2)(p-3)}{1.2.3.4} \times (\Delta^4 y_{-1} + \Delta^5 y_{-1}) + \dots$$
$$= y_0 + p\Delta y_{-1} + \frac{(p+1)p}{1.2}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{1.2.3}\Delta^3 y_{-1}$$
$$+ \frac{(p+1)p(p-1)(p-2)}{1.2.3.4}\Delta^4 y_{-1} + \frac{p(p-1)(p-2)(p-3)}{1.2.3.4}\Delta^5 y_{-1} + \dots$$
$$= y_0 + p\Delta y_{-1} + \frac{(p+1)p}{1.2}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{1.2.3}(\Delta^3 y_{-2} + \Delta^4 y_{-2})$$
$$+ \frac{(p+1)p(p-1)(p-2)}{1.2.3.4}(\Delta^4 y_{-2} + \Delta^5 y_{-2}) + \dots$$

[using (5) and (6)]

Hence
$$y_p = y_0 + p\Delta y_{-1} + \frac{(p+1)p}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-2}$$
$$+ \frac{(p+2)(p+1)p(p-1)}{4!}\Delta^4 y_{-2} + \dots$$

which is called *Gauss's backward interpolation formula*.



**Cor.** *In the central differences notation, this formula will be*

$$y_p = y_0 + p\delta y_{-1/2} + \frac{(p+1)p}{2!}\delta^2 y_0 + \frac{(p+1)p(p-1)}{3!}\delta^3 y_{-1/2} + \frac{(p+2)(p+1)p(p-1)}{4!}\delta^4 y_0 + \dots$$

**Obs. 1.** This formula contains odd differences above the central line and even differences on the central line as shown below :

$$\Delta y_{-1} \qquad \Delta^3 y_{-2} \qquad \Delta^5 y_{-3}$$
$$y_0 \;\cdots\cdots\; \Delta^2 y_{-1} \;\cdots\cdots\; \Delta^4 y_{-2} \;\cdots\cdots\; \Delta^6 y_{-3} \;\cdots\cdots \text{ Central line}$$

**Obs. 2.** It is used to interpolate the values of $y$ for a negative value of $p$ lying between $-1$ and $0$.

**Obs. 3.** Gauss's forward and backward formulae are not of much practical use. However, these serve as intermediate steps for obtaining the important formulae of the following sections.

## 7.7. STIRLING'S FORMULA

Gauss's forward interpolation formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-1}$$
$$+ \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \dots \quad \dots(1)$$

Gauss's backward interpolation formula is

$$y_p = y_0 + p\Delta y_{-1} + \frac{(p+1)p}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-2}$$
$$+ \frac{(p+2)(p+1)p(p-1)}{4!}\Delta^4 y_{-2} + \dots \quad \dots(2)$$

Taking the mean of (1) and (2), we obtain

$$y_p = y_0 + p\left(\frac{\Delta y_0 + \Delta y_{-1}}{2}\right) + \frac{p^2}{2!}\Delta^2 y_{-1} + \frac{p(p^2-1)}{3!}$$
$$\times \left(\frac{\Delta^3 y_{-1} + \Delta^3 y_{-2}}{2}\right) + \frac{p^2(p^2-1)}{4!}\Delta^4 y_{-2} + \dots \quad \dots(3)$$

$\frac{p(p^2-1)(p^2-2^2)}{5!}$

which is called *Stirling's formula*.

**Cor.** *In the central differences notation, (3) takes the form*

$$y_p = y_0 + p\,\mu\delta y_0 + \frac{p^2}{2!}\delta^2 y_0 + \frac{p(p^2-1^2)}{3!}\mu\delta^3 y_0 + \frac{p^2(p^2-1^2)}{4!}\delta^4 y_0 + \dots \quad \dots(4)$$

for
$$\tfrac{1}{2}(\Delta y_0 + \Delta y_{-1}) = \tfrac{1}{2}(\delta y_{1/2} + \delta y_{-1/2}) = \mu\,\delta y_0,$$
$$\tfrac{1}{2}(\Delta^3 y_{-1} + \Delta^3 y_{-2}) = \tfrac{1}{2}(\delta^3 y_{1/2} + \delta^3 y_{-1/2}) = \mu\,\delta^3 y_0 \text{ etc.}$$

**Obs.** *This formula involves means of the odd differences just above and below the central line and even differences on this line as shown below :*

$$\dots y_0 \cdots \begin{pmatrix}\Delta y_{-1}\\ \Delta y_0\end{pmatrix} \cdots \Delta^2 y_{-1} \cdots \begin{pmatrix}\Delta^3 y_{-2}\\ \Delta^3 y_{-1}\end{pmatrix} \cdots \Delta^4 y_{-2} \cdots \begin{pmatrix}\Delta^5 y_{-3}\\ \Delta^5 y_{-2}\end{pmatrix} \cdots \Delta^6 y_{-3} \dots \text{ Central line.}$$

## 7.8. BESSEL'S FORMULA

Gauss's forward interpolation formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-1} + \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \ldots \quad \ldots(1)$$

We have $\Delta^2 y_0 - \Delta^2 y_{-1} = \Delta^3 y_{-1}$

*i.e.* $\Delta^2 y_{-1} = \Delta^2 y_0 - \Delta^3 y_{-1}$ ...(2)

Similarly $\Delta^4 y_{-2} = \Delta^4 y_{-1} - \Delta^5 y_{-2}$ etc. ...(3)

Now (1) can be written as

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\left(\frac{1}{2}\Delta^2 y_{-1} + \frac{1}{2}\Delta^2 y_{-1}\right) + \frac{p(p^2-1)}{3!}\Delta^3 y_{-1} + \frac{p(p^2-1)(p-2)}{4!}\left(\frac{1}{2}\Delta^4 y_{-2} + \frac{1}{2}\Delta^4 y_{-2}\right) + \ldots$$

$$= y_0 + p\Delta y_0 + \frac{1}{2}\frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{1}{2}\frac{p(p-1)}{2!}(\Delta^2 y_0 - \Delta^3 y_{-1}) + \frac{p(p^2-1)}{3!}\Delta^3 y_{-1} + \frac{1}{2}\frac{p(p^2-1)(p-2)}{4!}\Delta^4 y_{-2} + \frac{1}{2}\frac{p(p^2-1)(p-2)}{4!} \times (\Delta^4 y_{-1} - \Delta^5 y_{-2}) + \ldots \quad \text{[using (2), (3) etc.}$$

$$= y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\cdot\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{p(p-1)}{2!} \times \left(\frac{p+1}{3} - \frac{1}{2}\right)\Delta^3 y_{-1} + \frac{p(p^2-1)(p-2)}{4!}\cdot\frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \ldots$$

Hence
$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\cdot\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{(p-\frac{1}{2})p(p-1)}{3!}\Delta^3 y_{-1} + \frac{(p+1)p(p-1)(p-2)}{4!}\cdot\frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2} + \ldots \quad \ldots(4)$$

which is known as the *Bessel's formula.*

**Cor.** *In the central differences notation, (4) becomes*

$$y_p = y_0 + p\delta y_{1/2} + \frac{p(p-1)}{2!}\mu\delta^2 y_{1/2} + \frac{(p-\frac{1}{2})p(p-1)}{3!}\delta^3 y_{1/2} + \frac{(p+1)p(p-1)(p-2)}{4!}\mu\delta^4 y_{1/2} + \ldots \quad \ldots(5)$$

for $\frac{1}{2}(\Delta^2 y_{-1} + \Delta^2 y_0) = \mu\delta^2 y_{1/2}, \frac{1}{2}(\Delta^4 y_{-2} + \Delta^4 y_{-1}) = \mu\delta^4 y_{1/2}$ etc.

**Obs.** *This is a very useful formula for practical purposes. It involves odd differences below the central line and means of even differences of and below his line* as shown below :

$$y_0 \cdots\cdots \begin{Bmatrix}\Delta^2 y_{-1} \\ \\ \Delta^2 y_0\end{Bmatrix} \cdots\cdots\cdots \begin{Bmatrix}\Delta^4 y_{-2} \\ \\ \Delta^4 y_{-1}\end{Bmatrix} \cdots\cdots\cdots \begin{Bmatrix}\Delta^6 y_{-3} \\ \\ \Delta^6 y_{-2}\end{Bmatrix} \cdots\cdots\cdots \text{Central line.}$$
$$\Delta y_0 \qquad \Delta^3 y_{-1} \qquad \Delta^5 y_{-2}$$



## 7.9. LAPLACE-EVERETT'S FORMULA

Gauss's forward interpolation formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-1} + \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \frac{(p+2)(p+1)p(p-1)(p-2)}{5!} \times \Delta^5 y_{-2} + \ldots \quad \ldots(1)$$

We eliminate the odd differences in (1) by using the relations
$\Delta y_0 = y_1 - y_0$, $\Delta^3 y_{-1} = \Delta^2 y_0 - \Delta^2 y_{-1}$, $\Delta^5 y_{-2} = \Delta^4 y_{-1} - \Delta^4 y_{-2}$ etc.
Then (1) becomes

$$y_p = y_0 + p(y_1 - y_0) + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}(\Delta^2 y_0 - \Delta^2 y_{-1}) + \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \frac{(p+2)(p+1)p(p-1)(p-2)}{5!} \times (\Delta^4 y_{-1} - \Delta^4 y_{-2}) + \ldots$$

$$= (1-p)y_0 + py_1 - \frac{p(p-1)(p-2)}{3!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^2 y_0 - \frac{(p+1)p(p-1)(p-2)(p-3)}{5!}\Delta^4 y_{-2} + \frac{(p+2)(p+1)p(p-1)(p-2)}{5!} \times \Delta^4 y_{-1} - \ldots$$

To change the terms with negative sign, putting $p = 1 - q$, we obtain

$$y_p = qy_0 + \frac{q(q^2-1^2)}{3!}\Delta^2 y_{-1} + \frac{q(q^2-1^2)(q^2-2^2)}{5!}\Delta^4 y_{-2} + \ldots + py_1 + \frac{p(p^2-1^2)}{3!}\Delta^2 y_0 + \frac{p(p^2-1^2)(p^2-2^2)}{5!}\Delta^4 y_{-1} + \ldots$$

This is known as *Laplace-Everett's formula.*

**Obs. 1.** *This formula is extensively used and involves only even differences on and below the central line* as shown below :

$$y_0 \cdots\cdots \Delta^2 y_{-1} \cdots\cdots \Delta^4 y_{-2} \cdots\cdots \Delta^6 y_{-3} \cdots\cdots \text{Central line}$$
$$y_1 \qquad \Delta^2 y_0 \qquad \Delta^4 y_{-1} \qquad \Delta^6 y_{-2}$$

**Obs. 2.** *There is a close relationship between Bessel's formula and Everett's formula and one can be deduced from the other by suitable rearrangements. It is also interesting to observe that Bessel's formula truncated after third differences is Everett's formula truncated after second differences.*

## 7.10. CHOICE OF AN INTERPOLATION FORMULA

So far we have derived several interpolation formulae for calculating $y_p$ from equispaced values. Now, we have to see which formula yields most accurate results in a particular problem.

The coefficients in the central difference formulae are smaller and converge faster than those in Newton's formulae. After a few terms, the coefficients in the Stirling's formula decrease more rapidly than those of the Bessel's formula and the coefficients of Bessel's formula decrease more rapidly than those of Newton's formula. As such, whenever possible, central difference formulae should be used in preference to Newton's formulae.

The right choice of an interpolation formula however, depends on the position of the interpolated value in the given data.

*The following rules will be found useful :*

*1. To find a tabulated value near the beginning of the table, use Newton's forward formula.*

*2. To find a value near the end of the table, use Newton's backward formula.*

*3. To find an interpolated value near the centre of the table, use either Stirling's or Bessel's or Everett's formula.*

If interpolation is required for $p$ lying between $-\frac{1}{4}$ and $\frac{1}{4}$, prefer Stirling's formula.

If interpolation is desired for $p$ lying between $\frac{1}{4}$ and $\frac{3}{4}$, use Bessel's or Everett's formula.

**Example 7.7.** *Find f (22) from the Gauss forward formula :*

| $x$ : | 20 | 25 | 30 | 35 | 40 | 45 |
|---|---|---|---|---|---|---|
| $f(x)$ : | 354 | 332 | 291 | 260 | 231 | 204 |

(J.N.T.U., B. Tech., 2007)

**Sol.** Taking $x_0 = 25$, $h = 5$, we have to find the value of $f(x)$ for $x = 22$.

*i.e.*, for $$p = \frac{x - x_0}{h} = \frac{22 - 25}{5} = -0.6.$$

The difference table is as follows :

| $x$ | $p$ | $y_p$ | $\Delta y_p$ | $\Delta^2 y_p$ | $\Delta^3 y_p$ | $\Delta^4 y_p$ | $\Delta^5 y_p$ |
|---|---|---|---|---|---|---|---|
| 20 | −1 | 354 $(= y_{-1})$ | | | | | |
| | | | −22 | | | | |
| 25 | 0 | 332 $(= y_0)$ | | −19 | | | |
| | | | −41 | | 29 | | |
| 30 | 1 | 291 $(= y_1)$ | | 10 | | −37 | |
| | | | −31 | | −8 | | 45 |
| 35 | 2 | 260 $(= y_2)$ | | 2 | | 8 | |
| | | | −29 | | 0 | | |
| 40 | 3 | 231 $(= y_3)$ | | 2 | | | |
| | | | −27 | | | | |
| 45 | 4 | 204 $(= y_4)$ | | | | | |

Gauss forward formula is

$$y_p = y_0 + p\,\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-1} + \frac{(p+1)p(p-1)(p-2)}{4!}\Delta^4 y_{-2} + \frac{(p+1)(p-1)(p-2)(p+2)}{5!}\Delta^5 y_{-2}$$

$$\therefore\ f(22) = 332 + (0.6)(-41) + \frac{(-0.6)(-0.6-1)}{2!}(-19) + \frac{(-0.6+1)(-0.6)(-0.6-1)}{3!}(-8)$$
$$+ \frac{(-0.6+1)(-0.6)(-0.6-1)(-0.6-2)}{4!}(-37)$$
$$+ \frac{(-0.6+1)(-0.6)(-0.6-1)(-0.6-2)(-0.6+2)}{5!}(45)$$
$$= 332 + 24.6 - 9.12 - 0.512 + 1.5392 - 0.5241$$

Hence $f(22) = 347.983.$



**Example 7.8.** *Use Gauss's forward formula to evaluate $y_{30}$, given that $y_{21} = 18.4708$, $y_{25} = 17.8144$, $y_{29} = 17.1070$, $y_{33} = 16.3432$ and $y_{37} = 15.5154$.*

**Sol.** Taking $x_0 = 29$, $h = 4$, we require the value of $y$ for $x = 30$

*i.e.* for $$p = \frac{x - x_0}{h} = \frac{30 - 29}{4} = 0.25.$$

The difference table is given below :

| $x$ | $p$ | $y_p$ | $\Delta y_p$ | $\Delta^2 y_p$ | $\Delta^3 y_p$ | $\Delta^4 y_p$ |
|---|---|---|---|---|---|---|
| 21 | −2 | 18.4708 | | | | |
| | | | −0.6564 | | | |
| 25 | −1 | 17.8144 | | −0.0510 | | |
| | | | −0.7074 | | −0.0054 | |
| 29 | 0 | 17.1070 | | −0.0564 | | −0.0022 |
| | | | −0.7638 | | −0.0076 | |
| 33 | 1 | 16.3432 | | −0.0640 | | |
| | | | −0.8278 | | | |
| 37 | 2 | 15.5154 | | | | |

Gauss's forward formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{1.2}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{1.2.3}\Delta^3 y_{-1} + \frac{(p+1)p(p-1)(p-2)}{1.2.3.4}\Delta^4 y_{-2} + \ldots$$

$$y_{30} = 17.1070 + (0.25)(-0.7638) + \frac{(0.25)(-0.75)}{2}(-0.0564)$$
$$+ \frac{(1.25)(0.25)(-0.75)}{6}(-0.0076) + \frac{(1.25)(0.25)(-0.75)(-1.75)}{24} \times (-0.0022)$$
$$= 17.1070 - 0.19095 + 0.00529 + 0.0003 - 0.00004 = 16.9216 \text{ approx.}$$

**Example 7.9.** *Using Gauss backward difference formula, find y (8) from the following table :*

| $x$ : | 0 | 5 | 10 | 15 | 20 | 25 |
|---|---|---|---|---|---|---|
| $y$ : | 7 | 11 | 14 | 18 | 24 | 32 |

(J.N.T.U., B. Tech., 2007)

**Sol.** Taking $x_0 = 10$, $h = 5$, we have to find $y$ for $x = 8$, *i.e.*, for $p = \frac{x - x_0}{h} = \frac{8 - 10}{5} = -0.4$.

The difference table is as follows :

| $x$ | $p$ | $y_p$ | $\Delta y_p$ | $\Delta^2 y_p$ | $\Delta^3 y_p$ | $\Delta^4 y_p$ | $\Delta^5 y_5$ |
|---|---|---|---|---|---|---|---|
| 0 | −2 | 7 | | | | | |
| | | | 4 | | | | |
| 5 | −1 | 11 | | −1 | | | |
| | | | 3 | | 2 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 10 | 0 | 14 | | 1 | | −1 | |
| | | | 4 | | 1 | | |
| 15 | 1 | 18 | | 2 | | | 0 |
| | | | 6 | | 0 | −1 | |
| 20 | 2 | 24 | | 2 | | | |
| | | | 8 | | | | |
| 25 | 3 | 32 | | | | | |

Gauss backward formula is

$$y_p = y_0 + p\Delta y_{-1} + \frac{(p+1)p}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-2} + \frac{(p+2)(p+1)p(p-1)}{4!}\Delta^4 y_{-2} + \ldots\ldots$$

$$y_{(8)} = 14 + (-0.4)(3) + \frac{(-0.4+1)(-0.4)}{2!}(1) + \frac{(-0.4+1)(-0.4)(-0.4-1)}{3!}(2) + \frac{(-0.4+2)(-0.4+1)(-0.4)(-0.4-1)}{4!}(-1)$$

$$= 14 - 1.2 - 0.12 + 0.112 + 0.034$$

Hence $y_{(8)} = 12.826$

**Example 7.10.** *Interpolate by means of Gauss's backward formula, the population of a town for the year 1974, given that :*

| *Year* | : | *1939* | *1949* | *1959* | *1969* | *1979* | *1989* |
|---|---|---|---|---|---|---|---|
| *Population (in thousands)* | : | *12* | *15* | *20* | *27* | *39* | *52* |

*(Kottayam, B. Tech., 2005)*

**Sol.** Taking $x_0 = 1969$, $h = 10$, the population of the town is to be found for

$$p = \frac{1974 - 1969}{10} = 0.5.$$

The central difference table is

| $x$ | $p$ | $y_p$ | $\Delta y_p$ | $\Delta^2 y_p$ | $\Delta^3 y_p$ | $\Delta^4 y_p$ | $\Delta^5 y_p$ |
|---|---|---|---|---|---|---|---|
| 1939 | −3 | 12 | | | | | |
| | | | 3 | | | | |
| 1949 | −2 | 15 | | 2 | | | |
| | | | 5 | | 0 | | |
| 1959 | −1 | 20 | | 2 | | 3 | |
| | | | 7 | | 3 | | −10 |
| 1969 | 0 | 27 | | 5 | | −7 | |
| | | | 12 | | −4 | | |
| 1979 | 1 | 39 | | 1 | | | |
| | | | 13 | | | | |
| 1989 | 2 | 52 | | | | | |



Gauss's backward formula is

$$y_p = y_0 + p\Delta y_{-1} + \frac{(p+1)p}{2!}\Delta^2 y_{-1} + \frac{(p+1)p(p-1)}{3!}\Delta^3 y_{-2} + \frac{(p+2)(p+1)p(p-1)}{4!}\Delta^4 y_{-2} + \frac{(p+2)(p+1)p(p-1)(p-2)}{5!}\Delta^5 y_{-3} + \ldots$$

*i.e.*

$$y_{0.5} = 27 + (0.5)(7) + \frac{(1.5)(0.5)}{2}(5) + \frac{(1.5)(0.5)(-0.5)}{6}(3) + \frac{(2.5)(1.5)(-0.5)}{24}(-7) + \frac{(2.5)(1.5)(0.5)(-0.5)(-1.5)}{120}(-10)$$

$$= 27 + 3.5 + 1.875 - 0.1875 + 0.2743 - 0.1172 = 32.532 \text{ thousands approx}$$

**Example 7.11.** *Employ Stirling's formula to compute $y_{12.2}$ from the following table ($y_x = 1 + \log_{10} \sin x$) :*

| $x°$ : | *10* | *11* | *12* | *13* | *14* |
|---|---|---|---|---|---|
| $10^5 u_x$ : | *23,967* | *28,060* | *31,788* | *35,209* | *38,368* |

*(V.T.U., B. Tech., 2004)*

**Sol.** Taking the origin at $x_0 = 12°$, $h = 1$ and $p = x - 12$, we have the following central difference table :

| $p$ | $y_x$ | $\Delta y_x$ | $\Delta^2 y_x$ | $\Delta^3 y_x$ | $\Delta^4 y_x$ |
|---|---|---|---|---|---|
| $-2 = x_{-2}$ | $0.23967 = y_{-2}$ | | | | |
| | | $0.04093 = \Delta y_{-2}$ | | | |
| $-1 = x_{-1}$ | $0.28060 = y_{-1}$ | | $-0.00365 = \Delta^2 y_{-2}$ | | |
| | | $0.03728 = \Delta y_{-1}$ | | $0.00058 = \Delta^3 y_{-2}$ | |
| $0 = x_0$ | $0.31788 = y_0$ | | $-0.00307 = \Delta^2 y_{-1}$ | | $-0.00013 = \Delta^4 y_{-2}$ |
| | | $0.03421 = \Delta y_0$ | | $-0.00045 = \Delta^3 y_{-1}$ | |
| $1 = x_1$ | $0.35209 = y_1$ | | $-0.00262 = \Delta^2 y_0$ | | |
| | | $0.03159 = \Delta y_1$ | | | |
| $2 = x_2$ | $0.38368 = y_2$ | | | | |

At $x = 12.2$, $p = 0.2$. (As $p$ lies between $-\frac{1}{4}$ and $\frac{1}{4}$, the use of Stirling's formula will be quite suitable.)

Stirling's formula is

$$y_p = y_0 + \frac{p}{1} \cdot \frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{p^2}{2!}\Delta^2 y_{-1} + \frac{p(p^2-1)}{3!}\frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \frac{p^2(p^2-1)}{4!}\Delta^4 y_{-2} + \ldots$$

When $p = 0.2$, we have

$$\therefore \quad y_{0.2} = 0.31788 + 0.2\left(\frac{0.03728 + 0.03421}{2}\right) + \frac{(0.2)^2}{2}(-0.00307) + \frac{(0.2)[(0.2)^2 - 1]}{6}\left(\frac{0.00058 - 0.00045}{2}\right) + \frac{(0.2)^2[(0.2)^2 - 1]}{24}(-0.00013)$$

$$= 0.31788 + 0.00715 - 0.00006 - 0.000002 + 0.0000002 = 0.32497.$$

**Example 7.12.** *Given*

| $\theta°$ : | 0 | 5 | 10 | 15 | 20 | 25 | 30 |
|---|---|---|---|---|---|---|---|
| $\tan\theta$ : | 0 | 0.0875 | 0.1763 | 0.2679 | 0.3640 | 0.4663 | 0.5774 |

*Using Stirling's formula, estimate the value of tan 16°.* (Anna. B. Tech., 2005)

**Sol.** Taking the origin at $\theta° = 15°$, $h = 5°$ and $p = \frac{\theta - 15}{5}$, we have the following central difference table :

| $p$ | $y = \tan\theta$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ | $\Delta^5 y$ |
|---|---|---|---|---|---|---|
| −3 | 0.0000 | | | | | |
| | | 0.0875 | | | | |
| −2 | 0.0875 | | 0.0013 | | | |
| | | 0.0888 | | 0.0015 | | |
| −1 | 0.1763 | | 0.0028 | | 0.0002 | |
| | | 0.0916 | | 0.0017 | | −0.0002 |
| 0 | 0.2679 | | 0.0045 | | 0.0000 | |
| | | 0.0961 | | 0.0017 | | 0.0009 |
| 1 | 0.3640 | | 0.0062 | | 0.0009 | |
| | | 0.1023 | | 0.0026 | | |
| 2 | 0.4663 | | 0.0088 | | | |
| | | 0.1111 | | | | |
| 3 | 0.5774 | | | | | |

At $\theta = 16°$, $\quad p = \frac{16 - 15}{5} = 0.2$

Stirling's formula is

$$y_p = y_0 + \frac{p}{1} \cdot \frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{p^2}{2!} \Delta^2 y_{-1} + \frac{p(p^2 - 1)}{3!} \cdot \frac{\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \frac{p^2(p^2 - 1)}{4!} \Delta^4 y_{-2} + \ldots$$

$$\therefore \quad y_{0.2} = 0.2679 + (0.2)\left(\frac{0.0916 + 0.0961}{2}\right) + \frac{(0.2)^2}{2!}(0.0045) + \ldots$$

$$= 0.2679 + 0.01877 + 0.00009 + \ldots = 0.28676$$

Hence $\quad \tan 16° = 0.28676.$

**Example 7.13.** *Apply Bessel's formula to obtain $y_{25}$, given $y_{20} = 2854$, $y_{24} = 3162$, $y_{28} = 3544$, $y_{32} = 3992$.* (S.V.T.U., B. Tech., 2007)

**Sol.** Taking the origin at $x_0 = 24$, $h = 4$, we have $p = \frac{1}{4}(x - 24)$.



$\therefore$ The central difference table is

| $p$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ |
|---|---|---|---|---|
| −1 | 2854 | | | |
| | | 308 | | |
| 0 | 3162 | | 74 | |
| | | 382 | | −8 |
| 1 | 3544 | | 66 | |
| | | 448 | | |
| 2 | 3992 | | | |

At $x = 25$, $p = (25 - 24)/4 = \frac{1}{4}$. (As $p$ lies between $\frac{1}{4}$ and $\frac{3}{4}$, the use of Bessel's formula will yield accurate result.)

Bessel's formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!} \frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2} + \frac{(p - \frac{1}{2})p(p-1)}{3!} \Delta^3 y_{-1} + \ldots \quad \ldots(1)$$

When $\quad p = 0.25$, we have

$$y_p = 3162 + 0.25 \times 382 + \frac{0.25(-0.75)}{2}\left(\frac{74 + 66}{2}\right) + \frac{(-0.25)\,0.25\,(-0.75)}{2}(-8)$$

$$= 3162 + 95.5 - 6.5625 - 0.0625$$

$$= 3250.875 \text{ approx.}$$

**Example 7.14.** *Apply Bessel's formula to find the value of $f(27.5)$ from the table :*

| $x$ : | 25 | 26 | 27 | 28 | 29 | 30 |
|---|---|---|---|---|---|---|
| $f(x)$ : | 4.000 | 3.846 | 3.704 | 3.571 | 3.448 | 3.333 |

(U.P.T.U., M.C.A., 2009)

**Sol.** Taking the origin at $x_0 = 27$, $h = 1$, we have $p = x - 27$

The central difference table is

| $x$ | $p$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|---|---|---|---|---|---|---|
| 25 | −2 | 4.000 | | | | |
| | | | −0.154 | | | |
| 26 | −1 | 3.846 | | 0.012 | | |
| | | | −0.142 | | −0.003 | |
| 27 | 0 | 3.704 | | 0.009 | | 0.004 |
| | | | −0.133 | | −0.001 | |
| 28 | 1 | 3.571 | | 0.010 | | −0.001 |
| | | | −0.123 | | −0.002 | |
| 29 | 2 | 3.448 | | 0.008 | | |
| | | | −0.115 | | | |
| 30 | 3 | 3.333 | | | | |

At $x = 27.5$, $p = 0.5$ (As $p$ lies between 1/4 and 3/4, the use of Bessel's formula will yield accurate result)

Bessel's formula is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\left(\frac{\Delta^2 y_{-1} + \Delta^2 y_0}{2}\right) + \frac{\left(p-\frac{1}{2}\right)p(p-1)}{3!}\Delta^3 y_{-1}$$

$$+ \frac{(p+1)p(p-1)(p-2)}{4!}\left(\frac{\Delta^4 y_{-2} + \Delta^4 y_{-1}}{2}\right) + \ldots\ldots$$

When $p = 0.5$, we have

$$y_p = 3.704 - \frac{(0.5)(0.5-1)}{2}\left(\frac{0.009 + 0.010}{2}\right) + 0$$

$$+ \frac{(0.5+1)(0.5)(0.5-1)(0.5-2)}{24}\left(\frac{-0.001 - 0.004}{2}\right)$$

$$= 3.704 - 0.11875 - 0.00006 = 3.585$$

Hence $f(27.5) = 3.585$.

**Example 7.15.** *Using Everett's formula, evaluate $f(30)$ if $f(20) = 2854$, $f(28) = 3162$, $f(36) = 7088$, $f(44) = 7984$.* *(U.P.T.U., B. Tech., 2006)*

**Sol.** Taking the origin at $x_0 = 28$, $h = 8$, we have $p = \frac{x-28}{8}$. The central table is

| $x$ | $p$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ |
|---|---|---|---|---|---|
| 20 | −1 | 2854 | | | |
| | | | 308 | | |
| 28 | 0 | 3162 | | 3618 | |
| | | | 3926 | | −6648 |
| 36 | 1 | 7088 | | −3030 | |
| | | | 896 | | |
| 44 | 2 | 7984 | | | |

At $x = 30$, $p = \frac{30-28}{8} = 0.25$ and $q = 1 - p = 0.75$

Everett's formula is

$$y_p = qy_0 + \frac{q(q^2-1^2)}{3!}\Delta^2 y_{-1} + \frac{q(q^2-1^2)(q^2-2^2)}{5!}\Delta^4 y_{-2} + \ldots\ldots$$

$$+ py_1 + \frac{p(p^2-1^2)}{3!}\Delta^2 y_0 + \frac{p(p^2-1^2)(p^2-2^2)}{5!}\Delta^4 y_{-1} + \ldots\ldots$$

$$= (0.75)(3162) + \frac{0.75(0.75^2-1)}{6}(3618) + \ldots\ldots$$

$$+ 0.25(7080) + \frac{0.25(0.25^2-1)}{6}(-3030) + \ldots\ldots$$

$$= 2371.5 - 351.75 + 1770 + 94.69 = 3884.4$$

Hence $f(30) = 3884.4$



**Example 7.16.** *Given the table*

| $x$: | 310 | 320 | 330 | 340 | 350 | 360 |
|---|---|---|---|---|---|---|
| $\log x$: | 2.49136 | 2.50515 | 2.51851 | 2.53148 | 2.54407 | 2.55630 |

*find the value of log 337.5 by Everett's formula.*

**Sol.** Taking the origin at $x_0 = 330$ and $h = 10$, we have $p = \frac{x-330}{10}$.

∴ The central difference table is

| $p$ | $y$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ | $\Delta^5 y$ |
|---|---|---|---|---|---|---|
| −2 | 2.49136 | | | | | |
| | | 0.01379 | | | | |
| −1 | 2.50515 | | − 0.00043 | | | |
| | | 0.01336 | | 0.00004 | | |
| 0 | 2.51881 | | − 0.00039 | | − 0.00003 | |
| | | 0.01297 | | 0.00001 | | 0.00004 |
| 1 | 2.53148 | | − 0.00038 | | 0.00001 | |
| | | 0.01259 | | 0.00002 | | |
| 2 | 2.54407 | | − 0.00036 | | | |
| | | 0.01223 | | | | |
| 3 | 2.55630 | | | | | |

To evaluate log 337.5 *i.e.* for $x = 337.5$, $p = \frac{337.5-330}{10} = 0.75$

(As $p > 0.5$ and $= 0.75$, Everett's formula will be quite suitable)

Everett's formula is

$$y_p = qy_0 + \frac{q(q^2-1^2)}{3!}\Delta^2 y_{-1} + \frac{q(q^2-1^2)(q^2-2^2)}{5!}\Delta^4 y_{-2} + \ldots\ldots$$

$$+ py_1 + \frac{p(p^2-1^2)}{3!}\Delta^2 y_0 + \frac{p(p^2-1^2)(p^2-2^2)}{5!}\Delta^4 y_{-1} + \ldots\ldots$$

$$= 0.25 \times 2.51851 + \frac{0.25(0.0625-1)}{6} \times (-0.00039)$$

$$+ \frac{0.25(0.0625-1)(0.0625-4)}{120} \times (-0.00003)$$

$$+ 0.75 \times 2.53148 + \frac{0.75(0.5625-1)}{6} \times (-0.00038)$$

$$+ \frac{0.75(0.5625-1)(0.5625-4)}{120} \times (0.00001)$$

$$= 0.62963 + 0.00002 - 0.0000002 + 1.89861 + 0.00002 + 0.0000001$$

$$= 2.52828 \text{ nearly.}$$

## PROBLEMS 7.2

1. Find the $y$ (25), given that $y_{20} = 24$, $y_{24} = 32$, $y_{28} = 35$, $y_{32} = 40$, using Gauss forward difference formula. (J.N.T.U., B. Tech., 2006)
2. Using Gauss forward formula, find a polynomial of degree four which takes the following values of the function $f(x)$ :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 |
| $f(x)$ : | 1 | $-1$ | 1 | $-1$ | 1 |

3. Using Gauss's forward formula, evaluate $f(3.75)$ from the table :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 2.5 | 3.0 | 3.5 | 4.0 | 4.5 | 5.0 |
| $y$ : | 24.145 | 22.043 | 20.225 | 18.644 | 17.262 | 16.047 |

(Bhopal, B. Tech., 2002)

4. From the following table :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 1.00 | 1.05 | 1.10 | 1.15 | 1.20 | 1.25 | 1.30 |
| $e^x$ : | 2.7183 | 2.8577 | 3.0042 | 3.1582 | 3.3201 | 3.4903 | 3.6693 |

Find $e^{1.17}$, using Gauss forward formula. (U.P.T.U., B. Tech., 2006)

5. Using Gauss's backward formula, estimate the number of persons earning wages between Rs. 60 and Rs. 70 from the following data :

| | | | | | |
|---|---|---|---|---|---|
| Wages (Rs.) : | Below 40 | 40—60 | 60—80 | 80—100 | 100—120 |
| *No. of persons* (in thousands) : | 250 | 120 | 100 | 70 | 50 |

(Tirchirapalli, B.E., 2001)

6. Apply Gauss's backward formula to find sin 45° from the following table :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\theta°$ : | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| $\sin\theta$ : | 0.34202 | 0.502 | 0.64279 | 0.76604 | 0.86603 | 0.93969 | 0.98481 |

7. Using Stirling's formula find $y_{35}$, given $y_{20} = 512$, $y_{30} = 439$, $y_{40} = 346$, $y_{50} = 243$, where $y_x$ represents the number of persons at age $x$ years in a life table. (Nagarjuna, B.E., 2003 S)
8. The pressure $p$ of wind corresponding to velocity $v$ is given by the following data. Estimate $p$ when $v = 25$.

| | | | | |
|---|---|---|---|---|
| $v$ : | 10 | 20 | 30 | 40 |
| $p$ : | 1.1 | 2 | 4.4 | 7.9 |

9. Use Stirling's formula to evaluate $f(1.22)$, given

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1.0 | 1.1 | 1.2 | 1.3 | 1.4 |
| $f(x)$ : | 0.841 | 0.891 | 0.932 | 0.963 | 0.985 |

(Tirchirapalli, B.E., 2001)

10. Calculate the value of $f(1.5)$ using Bessels' interpolation formula, from the table :

| | | | | |
|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 |
| $f(x)$ : | 3 | 6 | 12 | 15 |

(U.P.T.U., B. Tech., 2008)

11. Use Bessel's formula to obtain $y_{25}$, given $y_{20} = 24$, $y_{24} = 32$, $y_{28} = 35$, $y_{32} = 40$. (Gurukul, M.Sc., 2000)
12. Employ Bessel's formula to find the value of $F$ at $x = 1.95$, given that

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $x$ : | 1.7 | 1.8 | 1.9 | 2.0 | 2.1 | 2.2 | 2.3 |
| $F$ : | 2.979 | 3.144 | 3.283 | 3.391 | 3.463 | 3.997 | 4.491 |

Which other interpolation formula can be used here ? Which is more appropriate ? Give reasons.



13. From the following table :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 20 | 25 | 30 | 35 | 40 |
| $f(x)$ : | 11.4699 | 12.7834 | 13.7648 | 14.4982 | 15.0463 |

Find $f(34)$ using Everett's formula. (Madras, B.E., 2000 S)

14. Apply Everett's formula to obtain $u_{25}$, given $u_{20} = 2854$, $u_{24} = 3162$, $u_{28} = 3544$, $u_{32} = 3992$.
15. Given the table :

| | | | | | | |
|---|---|---|---|---|---|---|
| $x$ : | 310 | 320 | 330 | 340 | 350 | 360 |
| $\log x$ : | 2.4914 | 2.5052 | 2.5185 | 2.5315 | 2.5441 | 2.5563, |

find the value of log 337.5 by Gauss, Stirling, Bessel and Everette formulae.

## 7.11. INTERPOLATION WITH UNEQUAL INTERVALS

The various interpolation formulae derived so far possess the disadvantage of being applicable only to equally spaced values of the argument. It is, therefore, desirable to develop interpolation formulae for unequally spaced values of $x$. Now we shall study two such formulae :

(i) Lagrange's interpolation formula

(ii) Newton's general interpolation formula with divided differences.

## 7.12. LAGRANGE'S INTERPOLATION FORMULA

*If $y = f(x)$ takes the value $y_0, y_1, \ldots\ldots, y_n$ corresponding to $x = x_0, x_1, \ldots\ldots, x_n$, then*

$$f(x) = \frac{(x-x_1)(x-x_2)\ldots\ldots(x-x_n)}{(x_0-x_1)(x_0-x_2)\ldots\ldots(x_0-x_n)}y_0 + \frac{(x-x_0)(x-x_2)\ldots\ldots(x-x_n)}{(x_1-x_0)(x_1-x_2)\ldots\ldots(x_1-x_n)}y_1 + \ldots\ldots + \frac{(x-x_0)(x-x_1)\ldots\ldots(x-x_{n-1})}{(x_n-x_0)(x_n-x_1)\ldots\ldots(x_n-x_{n-1})}y_n \quad \ldots(1)$$

This is known as *Lagrange's interpolation formula for unequal intervals.*

*Proof.* Let $y = f(x)$ be a function which takes the values $(x_0, y_0)$, $(x_1, y_1)$, ......, $(x_n, y_n)$. Since there are $n + 1$ pairs of values of $x$ and $y$, we can represent $f(x)$ by a polynomial in $x$ of degree $n$. Let this polynomial be of the form

$$y = f(x) = a_0(x-x_1)(x-x_2)\ldots\ldots(x-x_n) + a_1(x-x_0)(x-x_2)\ldots\ldots(x-x_n) + a_2(x-x_0)(x-x_1)(x-x_3)\ldots\ldots(x-x_n) + \ldots + a_n(x-x_0)(x-x_1)\ldots\ldots(x-x_{n-1}) \quad \ldots(2)$$

Putting $x = x_0$, $y = y_0$, in (2), we get

$$y_0 = a_0(x_0-x_1)(x_0-x_2)\ldots\ldots(x_0-x_n)$$
$$a_0 = y_0/[(x_0-x_1)(x_0-x_2)\ldots\ldots(x_0-x_n)]$$

Similarly putting $x = x_1$, $y = y_1$ in (2), we have $a_1 = y_1/[(x_1-x_0)(x_1-x_2)\ldots\ldots(x_1-x_n)]$

Proceeding the same way, we find $a_2, a_3 \ldots\ldots a_n$.

Substituting the values of $a_0, a_1, \ldots\ldots, a_n$ in (2), we get (1).

**Obs.** Lagranges interpolation formula (1) for $n$ points is a polynomial of degree $(n-1)$ which is known as *Lagrangian polynomial* and is very simple to implement on a computer.

*This formula can also be used to split the given function into partial fractions.*

For on dividing both sides of (1) by $(x-x_0)(x-x_1)\ldots\ldots(x-x_n)$, we get

$$\frac{f(x)}{(x-x_0)(x-x_1)\ldots\ldots(x-x_n)} = \frac{y_0}{(x_0-x_1)(x_0-x_2)\ldots\ldots(x_0-x_n)}\cdot\frac{1}{x-x_0} + \frac{y_1}{(x_1-x_0)(x_1-x_2)\ldots\ldots(x_1-x_n)}\cdot\frac{1}{x-x_1} + \ldots\ldots + \frac{y_n}{(x_n-x_0)(x_n-x_1)\ldots\ldots(x_n-x_{n-1})}\cdot\frac{1}{x-x_n}.$$

**Example 7.17.** *Given the values*

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 5 | 7 | 11 | 13 | 17 |
| $f(x)$ : | 150 | 392 | 1452 | 2366 | 5202, |

*evaluate $f(9)$, using Lagrange's formula* (V.T.U., B. Tech., 2006)

**Sol.** (*i*) Here $x_0 = 5, \quad x_1 = 7, \quad x_2 = 11, \quad x_3 = 13, \quad x_4 = 17$

and $y_0 = 150, y_1 = 392, y_2 = 1452, y_3 = 2366, y_4 = 5202.$

Putting $x = 9$ and substituting the above values in Lagrange's formula, we get

$$f(9) = \frac{(9-7)(9-11)(9-13)(9-17)}{(5-7)(5-11)(5-13)(5-17)} \times 150 + \frac{(9-5)(9-11)(9-13)(9-17)}{(7-5)(7-11)(7-13)(7-17)} \times 392$$

$$+ \frac{(9-5)(9-7)(9-13)(9-17)}{(11-5)(11-7)(11-13)(11-17)} \times 1452$$

$$+ \frac{(9-5)(9-7)(9-11)(9-17)}{(13-5)(13-7)(13-11)(13-17)} \times 2366$$

$$+ \frac{(9-5)(9-7)(9-11)(9-13)}{(17-5)(17-7)(17-11)(17-13)} \times 5202$$

$$= -\frac{50}{3} + \frac{3136}{15} + \frac{3872}{3} - \frac{2366}{3} + \frac{578}{5} = 810$$

**Example 7.18.** *Find the polynomial $f(x)$ by using Lagrange's formula and hence find $f(3)$ for*

| | | | | |
|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 5 |
| $f(x)$ : | 2 | 3 | 12 | 147 |

(Anna, B. Tech., 2005)

**Sol.** Here $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 5$

and $y_0 = 2, y_1 = 3, y_2 = 12, y_3 = 147.$

Lagrange's formula is

$$y = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} y_1$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} y_3$$

$$= \frac{(x-1)(x-2)(x-5)}{(0-1)(0-2)(0-5)}(2) + \frac{(x-0)(x-2)(x-5)}{(1-0)(1-2)(1-5)}(3)$$

$$+ \frac{(x-0)(x-1)(x-5)}{(2-0)(2-1)(2-5)}(12) + \frac{(x-0)(x-1)(x-2)}{(5-0)(5-1)(5-2)}(147)$$

Hence $f(x) = x^3 + x^2 - x + 2$

$\therefore \quad f(3) = 27 + 9 - 3 + 2 = 35.$

**Example 7.19.** *A curve passes through the points (0, 18), (1, 10), (3, – 18) and (6, 90). Find the slope of the curve at $x = 2$.* (J.N.T.U., B.Tech., 2009)

**Sol.** Here $x_0 = 0, x_1 = 1, x_2 = 3, x_3 = 6$ and $y_0 = 18, y_1 = 10, y_2 = -18, y_3 = 90$



Since the values of $x$ are unequally spaced, we use the Lagrange's formula :

$$y = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} y_1$$

$$+ \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} y_3$$

$$= \frac{(x-1)(x-3)(x-6)}{(0-1)(0-3)(0-6)}(18) + \frac{(x-0)(x-3)(x-6)}{(1-0)(1-3)(1-6)}(10)$$

$$+ \frac{(x-0)(x-1)(x-6)}{(3-0)(3-1)(3-6)}(-18) + \frac{(x-0)(x-1)(x-3)}{(6-0)(6-1)(6-3)}(90)$$

$$= (-x^3 + 10x^2 - 27x + 18) + (x^3 - 9x^2 + 18x) + (x^3 - 7x^2 + 6x) + (x^3 - 4x^2 + 3x)$$

*i.e.,* $y = 2x^3 - 10x^2 + 18$

Thus the slope of the curve at $x = 2 = \left(\frac{dy}{dx}\right)_{x=2}$

$$= (6x^2 - 20x)_{x=2} = -16.$$

**Example 7.20.** *Using Lagrange's formula, express the function $\frac{3x^2+x+1}{(x-1)(x-2)(x-3)}$ as a sum of partial fractions.*

**Sol.** Let us evaluate $y = 3x^2 + x + 1$ for $x = 1$, $x = 2$ and $x = 3$

These values are

| | | | |
|---|---|---|---|
| $x$ : | $x_0 = 1$ | $x_1 = 2$ | $x_2 = 3$ |
| $y$ : | $y_0 = 5$ | $y_1 = 15$ | $y_2 = 31$ |

Lagrange's formula is

$$y = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} y_0 + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} y_1 + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} y_2$$

Substituting the above values, we get

$$y = \frac{(x-2)(x-3)}{(1-2)(1-3)}(5) + \frac{(x-1)(x-3)}{(2-1)(2-3)}(15) + \frac{(x-1)(x-2)}{(3-1)(3-2)}(31)$$

$$= 2.5\,(x-2)(x-3) - 15\,(x-1)(x-3) + 15.5\,(x-1)(x-2)$$

Thus $$\frac{3x^2+x+1}{(x-1)(x-2)(x-3)} = \frac{2.5(x-2)(x-3) - 15(x-1)(x-3) + 15.5(x-1)(x-2)}{(x-1)(x-2)(x-3)}$$

$$= \frac{2.5}{x-1} - \frac{15}{x-2} + \frac{15.5}{x-3}$$

**Example 7.21.** *Find the missing term in the following table using interpolation :*

| $x$ : | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $y$ : | 1 | 3 | 9 | ... | 81 |

**Sol.** Since the given data is unevenly spaced, therefore we use Lagrange's interpolation formula :

$$y = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)} y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)} y_1 + \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)} y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)} y_3$$

Here we have $x_0 = 0 \quad x_1 = 1 \quad x_2 = 2 \quad x_3 = 4$

$y_0 = 1 \quad y_1 = 3 \quad y_2 = 9 \quad y_3 = 81$

$$\therefore \quad y = \frac{(x-1)(x-2)(x-4)}{(0-1)(0-2)(0-4)}(1) + \frac{(x-0)(x-2)(x-4)}{(1-0)(1-2)(1-4)}(3) + \frac{(x-0)(x-1)(x-4)}{(2-0)(2-1)(2-4)}(9) + \frac{(x-0)(x-1)(x-2)}{(4-0)(4-1)(4-2)}(81)$$

When $x = 3$, then

$$\therefore \quad y = \frac{(3-1)(3-2)(3-4)}{-8} + 3(3-2)(3-4) + \frac{3(3-1)(3-4)}{-4}(9) + \frac{3(3-1)(3-2)}{24}(81) = \frac{1}{4} - 3 + \frac{27}{2} + \frac{81}{24} = 31.$$

Hence the missing term for $x = 3$ is $y = 31$.

**Example 7.22.** *Find the distance moved by a particle and its acceleration at the end of 4 seconds, if the time verses velocity data is as follows :*

| $t$ : | 0 | 1 | 3 | 4 |
|---|---|---|---|---|
| $v$ : | 21 | 15 | 12 | 10 |

**Sol.** Since the values of $t$ are not equispaced, we use Lagrange's formula :

$$v = \frac{(t-t_1)(t-t_2)(t-t_3)}{(t_0-t_1)(t_0-t_2)(t_0-t_3)} v_0 + \frac{(t-t_0)(t-t_2)(t-t_3)}{(t_1-t_0)(t_1-t_2)(t_1-t_3)} v_1 + \frac{(t-t_0)(t-t_1)(t-t_3)}{(t_2-t_0)(t_2-t_1)(t_2-t_3)} v_2 + \frac{(t-t_0)(t-t_1)(t-t_2)}{(t_3-t_0)(t_3-t_1)(t_3-t_2)} v_3$$

*i.e.,*
$$v = \frac{(t-1)(t-3)(t-4)}{(-1)(-2)(-4)}(21) + \frac{t(t-3)(t-4)}{(1)(-2)(-3)}(15) + \frac{t(t-1)(t-4)}{(3)(2)(-1)}(12) + \frac{t(t-1)(t-3)}{(4)(3)(1)}(10)$$

*i.e.,*
$$v = \frac{1}{12}(-5t^3 + 38t^2 - 105t + 252)$$

$$\therefore \quad \text{Distance moved } s = \int_0^4 v\,dt = \frac{1}{12}\int_0^4 (-5t^3 + 38t^2 - 105t + 252)\,dt \qquad \left[\because\ v = \frac{ds}{dt}\right]$$



$$= \frac{1}{12}\left(\frac{-5t^4}{4} + \frac{38t^3}{3} - \frac{105t^2}{2} + 252t\right)_0^4$$

$$= \frac{1}{12}\left(-320 + \frac{2432}{3} - 840 + 1008\right) = 54.9$$

Also acceleration $= \dfrac{dv}{dt} = \dfrac{1}{2}(-15t^2 + 76t - 105 + 0)$

Hence acceleration at $(t = 4) = \dfrac{1}{12}(-15(16) + 76(4) - 105) = -3, 4.$

## PROBLEMS 7.3

1. Use Lagrange's interpolation formula to find the value of $y$ when $x = 10$, if the following values of $x$ and $y$ are given :

| $x$ : | 5 | 6 | 9 | 11 |
|---|---|---|---|---|
| $y$ : | 12 | 13 | 14 | 16. |

(*J.N.T.U., B. Tech., 2008*)

2. The following table gives the viscosity of an oil as a function of temperature. Use Lagrange's formula to find viscosity of oil at a temperature of 140°.

| Temp.° : | 110 | 130 | 160 | 190 |
|---|---|---|---|---|
| Viscosity : | 10.8 | 8.1 | 5.5 | 4.8 |

3. Given $\log_{10} 654 = 2.8156$, $\log_{10} 658 = 2.8182$, $\log_{10} 659 = 2.8189$, $\log_{10} 661 = 2.8202$, find by using Lagrange's formula, the value of $\log_{10} 656$. (*Hazaribagh, B.E., 2009*)

4. The following are the measurements $T$ made on a curve recorded by oscilograph representing a change of current $I$ due to a change in the conditions of an electric current.

| $T$ : | 1.2 | 2.0 | 2.5 | 3.0 |
|---|---|---|---|---|
| $I$ : | 1.36 | 0.58 | 0.34 | 0.20 |

Using Lagrange's formula, find $I$ and $T = 1.6$. (*J.N.T.U., B. Tech., 2009*)

5. Using Lagrange's interpolation, calculate the profit in the year 2000 from the following data :

| Year : | 1997 | 1999 | 2001 | 2002 |
|---|---|---|---|---|
| Profit in Lakhs of Rs : | 43 | 65 | 159 | 248 |

(*Anna, B. Tech., 2004*)

6. Use Lagrange's formula to find the form of $f(x)$, given

| $x$ : | 0 | 2 | 3 | 6 |
|---|---|---|---|---|
| $f(x)$ : | 648 | 704 | 729 | 792. |

(*Madras, B.E., 2003 S*)

7. If $y(1) = -3$, $y(3) = 9$, $y(4) = 30$, $y(6) = 132$, find the Lagrange's interpolation polynomial that takes the same values as $y$ at the given points. (*V.T.U., B. Tech., 2006*)

8. Given $f(0) = -18$, $f(1) = 0$, $f(3) = 0$, $f(5) = -248$, $f(6) = 0$, $f(9) = 13104$, find $f(x)$.

(*Nagarjuna, B.E., 2003 S*)

**9.** Find the missing term in the following table using interpolation

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 4 | 5 | 6 |
| $y$ : | 14 | 15 | 5 | ... | 9. |

**10.** Using Lagrange's formula, express the function $\dfrac{x^2+x-3}{x^3-2x^2-x+2}$ as sum of partial fractions.

**11.** Using Lagrange's formula, express the function $\dfrac{x^2+6x-1}{(x^2-1)(x-4)(x-6)}$ as a sum of partial fractions.

[**Hint.** Tabulate the values of $f(x) = x^2 + 6x - 1$ for $x = -1, 1, 4, 6$ and apply Lagrange's formula.]

## 7.13 DIVIDED DIFFERENCES

The Lagrange's formula has the drawback that if another interpolation value were inserted, then the interpolation coefficients are required to be recalculated. This labour of recomputing the interpolation coefficients is saved by using Newton's general interpolation formula which employs what are called 'divided differences'. Before deriving this formula, we shall first define these differences.

If $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots\ldots$ be given points, then the *first divided difference* for the arguments $x_0, x_1$ is defined by the relation $[x_0, x_1] = \dfrac{y_1 - y_0}{x_1 - x_0}$.

Similarly $[x_1, x_2] = \dfrac{y_2 - y_1}{x_2 - x_1}$ and $[x_2, x_3] = \dfrac{y_3 - y_2}{x_3 - x_2}$ etc.

The *second divided difference* for $x_0, x_1, x_2$ is defined as $[x_0, x_1, x_2] = \dfrac{[x_1, x_2] - [x_0, x_1]}{x_2 - x_0}$.

The *third divided difference* for $x_0, x_1, x_2, x_3$ is defined as

$[x_0, x_1, x_2, x_3] = \dfrac{[x_1, x_2, x_3] - [x_0, x_1, x_2]}{x_3 - x_0}$ and so on.

**Obs. 1.** *The divided differences are symmetrical in their arguments i.e. independent of the order of the arguments.* For it is easy to write $[x_0, x_1] = \dfrac{y_0}{x_0 - x_1} + \dfrac{y_1}{x_1 - x_0} = [x_1, x_0]$, $[x_0, x_1, x_2]$

$$= \frac{y_0}{(x_0 - x_1)(x_0 - x_2)} + \frac{y_1}{(x_1 - x_0)(x_1 - x_2)} + \frac{y_2}{(x_2 - x_0)(x_2 - x_1)}$$

$= [x_1, x_2, x_0]$ or $[x_2, x_0, x_1]$ and so on.

**Obs. 2.** *The nth divided differences of a polynomial of the nth degree are constant.*

Let the arguments be equally spaced so that $x_1 - x_0 = x_2 - x_1 = \ldots = x_n - x_{n-1} = h$. Then

$$[x_0, x_1] = \frac{y_1 - y_0}{x_1 - x_0} = \frac{\Delta y_0}{h}$$

$$[x_0, x_1, x_2] = \frac{[x_1, x_2] - [x_0 - x_1]}{x_2 - x_0} = \frac{1}{2h}\left\{\frac{\Delta y_1}{h} - \frac{\Delta y_0}{h}\right\}$$



DRAFT TUBR

$$= \frac{1}{2!\,h^2}\Delta^2 y_0 \text{ and in general, } [x_0, x_1, x_2, \ldots\ldots, x_n] = \frac{1}{n!\,h^n}\Delta^n y_0.$$

If the tabulated function is a $n$th degree polynomial, then $\Delta^n y_0$ will be constant. Hence the $n$th divided differences will also be constant.

## 7.14. NEWTON'S DIVIDED DIFFERENCE FORMULA

Let $y_0, y_1, \ldots\ldots, y_n$ be the values of $y = f(x)$ corresponding to the arguments $x_0, x_1, \ldots\ldots, x_n$. Then from the definition of divided differences, we have

$$[x, x_0] = \frac{y - y_0}{x - x_0}$$

so that $\quad y = y_0 + (x - x_0)[x, x_0] \qquad \ldots(1)$

Again $\quad [x, x_0, x_1] = \dfrac{[x, x_0] - [x_0, x_1]}{x - x_1}$

which gives $\quad [x, x_0] = [x_0, x_1] + (x - x_1)[x, x_0, x_1]$

Substituting this value of $[x, x_0]$ in (1), we get

$$y = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x, x_0, x_1] \qquad \ldots(2)$$

Also $\quad [x, x_0, x_1, x_2] = \dfrac{[x, x_0, x_1] - [x_0\, x_1, x_2]}{x - x_2}$

which gives $\quad [x, x_0, x_1] = [x_0, x_1, x_2] + (x - x_2)[x, x_0, x_1, x_2]$

Substituting this value of $[x, x_0, x_1]$ in (2), we obtain

$$y = y_0 + (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] + (x - x_0)(x - x_1)(x - x_2)[x, x_0, x_1, x_2]$$

Proceeding in this manner, we get

$$\begin{aligned} y = f(x) = y_0 &+ (x - x_0)[x_0, x_1] + (x - x_0)(x - x_1)[x_0, x_1, x_2] \\ &+ (x - x_0)(x - x_1)(x - x_2)[x_0, x_1, x_2, x_3] + \ldots\ldots \\ &+ (x - x_0)(x - x_1) \ldots\ldots (x - x_n)[x, x_0, x_1, \ldots\ldots x_n] \end{aligned} \qquad \ldots(3)$$

$\dfrac{y_2 - y_1}{x_1 - x_1}$

which is called *Newton's general interpolation formula with divided differences.*

## 7.15 RELATION BETWEEN DIVIDED AND FORWARD DIFFERENCES

If $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots\ldots$ be the given points, then

$$[x_0, x_1] = \frac{y_1 - y_0}{x_1 - x_0}$$

Also $\quad \Delta y_0 = y_1 - y_0$

If $x_0, x_1, x_2, \ldots\ldots$ are equispaced, then $x_1 - x_0 = h$, so that

$$[x_0, x_1] = \frac{\Delta y_0}{h}$$

Similarly $\quad [x_1, x_2] = \dfrac{\Delta y_1}{h}$

Now $\quad [x_0, x_1, x_2] = \dfrac{[x_1, x_2] - [x_0, x_1]}{x_2 - x_0}$

$$= \frac{\Delta y_1/h - \Delta y_0/h}{2h} \qquad [\because \; x_2 - x_0 = 2h]$$

$$= \frac{\Delta y_1 - \Delta y_0}{2h^2} = \frac{\Delta^2 y_0}{2h^2}$$

Thus $\quad [x_0, x_1, x_2] = \dfrac{\Delta^2 y_0}{2!\,h^2}$

Similarly $\quad [x_1, x_2, x_3] = \dfrac{\Delta^2 y_1}{2!\,h^2}$

$$\therefore \quad [x_0, x_1, x_2, x_3] = \frac{\Delta^2 y_1/2h^2 - \Delta^2 y_0/2h^2}{x_3 - x_0} = \frac{\Delta^2 y_1 - \Delta^2 y_0}{2h^2(3h)} \qquad [\because \; x_3 - x_0 = 3h]$$

Thus $\quad [x_0, x_1, x_2, x_3] = \dfrac{\Delta^3 y_0}{3!\,h^3}$

In general, $[x_0, x_1, \ldots\ldots, x_n] = \dfrac{\Delta^n y_0}{n!\,h^n}$.

This is the relation between divided and forward differences.

**Example 7.23.** *Given the values*

| | | | | | | |
|---|---|---|---|---|---|---|
| *x* : | *5* | *7* | *11* | *13* | *17* | |
| *f(x)* : | *150* | *392* | *1452* | *2366* | *5202,* | |

*evaluate f(9), using Newton's divided difference formula.*

**Sol.** The divided differences table is

| $x$ | $y$ | 1st divided differences | 2nd divided differences | 3rd divided differences |
|---|---|---|---|---|
| 5 | 150 | | | |
| | | $\frac{392-150}{7-5} = 121$ | | |
| 7 | 392 | | $\frac{265-121}{11-5} = 24$ | |
| | | $\frac{1452-392}{11-7} = 265$ | | $\frac{32-24}{13-5} = 1$ |
| 11 | 1452 | | $\frac{457-265}{13-7} = 32$ | |
| | | $\frac{2366-1452}{13-11} = 457$ | | $\frac{42-32}{17-7} = 1$ |
| 13 | 2366 | | $\frac{709-457}{17-11} = 42$ | |
| | | $\frac{5202-2366}{17-13} = 709$ | | |
| 17 | 5202 | | | |



Taking $x = 9$ in the Newton's divided difference formula, we obtain

$$f(9) = 150 + (9-5) \times 121 + (9-5)(9-7) \times 24 + (9-5)(9-7)(9-11) \times 1$$
$$= 150 + 484 + 192 - 16 = 810.$$

**Example 7.24.** *Using Newton's divided differences formula, evaluate f(8) and f(15) given :*

| | | | | | | |
|---|---|---|---|---|---|---|
| *x* : | *4* | *5* | *7* | *10* | *11* | *13* |
| *f(x)* : | *48* | *100* | *294* | *900* | *1210* | *2028* |

(*V.T.U., B. Tech., 2008*)

**Sol.** The divided differences table is

| $x$ | $f(x)$ | 1st divided differences | 2nd divided differences | 3rd divided differences | 4th divided differences |
|---|---|---|---|---|---|
| 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 | | | | |

Taking $x = 8$ in the Newton's divided difference formula, we obtain

$$f(8) = 48 + (8-4)\,52 + (8-4)(8-5)\,15 + (8-4)(8-5)(8-7)\,1$$
$$= \mathbf{448.}$$

Similarly $f(15) = 3150$.

**Example 7.25.** *Determine f(x) as a polynomial in x for the following data :*

| | | | | | |
|---|---|---|---|---|---|
| *x* : | *– 4* | *– 1* | *0* | *2* | *5* |
| *f(x)* : | *1245* | *33* | *5* | *9* | *1335.* |

(*V.T.U., B. Tech., 2007*)

**Sol.** The divided differences table is

| $x$ | $f(x)$ | 1st divided differences | 2nd divided differences | 3rd divided differences | 4th divided differences |
|---|---|---|---|---|---|
| – 4 | 1245 | | | | |
| | | – 404 | | | |
| – 1 | 33 | | 94 | | |
| | | – 28 | | – 14 | |
| 0 | 5 | | 10 | | 3 |
| | | 2 | | 13 | |
| 2 | 9 | | 88 | | |
| | | 442 | | | |
| 5 | 1335 | | | | |

Applying Newton's divided difference formula

$$f(x) = f(x_0) + (x - x_0)\,[x_0, x_1] + (x - x_0)(x - x_1)\,[x_0, x_1, x_2] + \ldots$$
$$= 1245 + (x + 4)(-404) + (x + 4)(x + 1)(94)$$
$$+ (x + 4)(x + 1)(x - 0)(-14) + (x + 4)(x + 1)x(x - 2)(3)$$
$$= 3x^4 - 5x^3 + 6x^2 - 14x + 5.$$

**Example 7.26.** *Using Newton's divided difference formula, find the missing value from the table :*

| $x$ : | 1 | 2 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| $y$ : | 14 | 15 | 5 | ... | 9 |

**Sol.** The divided difference table is

| $x$ | $y$ | *1st divided differences* | *2nd divided differences* | *3rd divided differences* |
|---|---|---|---|---|
| 1 | 14 | | | |
| | | $\frac{15-14}{2-1} = 1$ | | |
| 2 | 15 | | $\frac{-5-1}{4-1} = -2$ | |
| | | $\frac{5-15}{4-2} = -5$ | | $\frac{7/4+2}{6-1} = \frac{3}{4}$ |
| 4 | 5 | | $\frac{2+5}{6-2} = \frac{7}{4}$ | |
| | | $\frac{9-5}{6-4} = 2$ | | |
| 6 | 9 | | | |

Newton's divided difference formula is

$$y = y_0 + (x - x_0)\,[x_0, x_1] + (x - x_0)(x - x_1)\,[x_0, x_1, x_2]$$
$$+ (x - x_0)(x - x_1)(x - x_2)\,[x_0, x_1, x_2, x_3] + \ldots$$

$$= 14 + (x - 1)(1) + (x - 1)(x - 2)(-2) + (x - 1)(x - 2)(x - 4)\left(\frac{3}{4}\right)$$

Putting $x = 5$, we get

$$y(5) = 14 + 4 + (4)(3)(-2) + (4)(3)(1)\left(\frac{3}{4}\right) = 3.$$

Hence missing value is 3.



## PROBLEMS 7.4

1. Find the third divided difference with arguments 2, 4, 9, 10 of the function $f(x) = x^3 - 2x$.
   (U.P.T.U., B. Tech., 2005)
2. Obtain the Newton's divided difference interpolating polynomial and hence find $f(6)$ :

| $x$ : | 3 | 7 | 9 | 10 |
|---|---|---|---|---|
| $f(x)$ : | 168 | 120 | 72 | 63 |

(U.P.T.U., B. Tech., 2007)

3. Using Newton's divided differences interpolation, find $u(3)$, given that $u(1) = -26$, $u(2) = 12$, $u(4) = 256$, $u(6) = 844$.
   (Anna, B.E., 2004)
4. A thermocouple gives the following output for rise in temperature

| Temp (°C) | 0 | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|---|
| Output (mV) | 0.0 | 0.4 | 0.8 | 1.2 | 1.6 | 2.0 |

Find the output of thermocouple for 37°C temperature using Newton's divided difference formula.

5. Using Newton's divided difference interpolation, find the polynomial of the given data :

| $x$ : | −1 | 0 | 1 | 3 |
|---|---|---|---|---|
| $f(x)$ : | 2 | 1 | 0 | −1 |

(Anna, B.E., 2007)

6. For the following table, find $f(x)$ as a polynomial in $x$ using Newton's divided difference formula :

| $x$ : | 5 | 6 | 9 | 11 |
|---|---|---|---|---|
| $f(x)$ : | 12 | 13 | 14 | 16 |

7. 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 |

8. The observed values of a function are respectively 168, 120, 72 and 63 at the four positions 3, 7, 9 and 10 of the independent variable. What is the best estimate for the value of the function at the position 6.
9. Find the equation of the cubic curve which passes through the points (4, − 43), (7, 83), (9, 327) and (12, 1053).
10. Find the missing term in the following table using Newton's divided difference formula

| $x$ : | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $y$ : | 1 | 3 | 9 | ... | 81 |

11. Certain corresponding values of $x$ and $\log_{10} x$ are given below :

| $x$ : | 300 | 304 | 305 | 307 |
|---|---|---|---|---|
| $\log_{10} x$ : | 2.4771 | 2.4829 | 2.4843 | 2.4871 |

Find $\log_{10} 310$ by (*i*) Lagrange's formula.
(*ii*) Newton's divided difference formula.

## 7.16. HERMITE'S INTERPOLATION FORMULA

This formula is similar to the Lagrange's interpolation formula. In Lagrange's method, the interpolating polynomial $P(x)$ agrees with $y(x)$ at the points $x_0, x_1, \ldots\ldots, x_n$, whereas in Hermite's method $P(x)$ and $y(x)$ as well as $P'(x)$ and $y'(x)$ coincide at the $(n + 1)$ points *i.e.*

$$P(x_i) = y(x_i) \text{ and } P'(x_i) = y'(x_i)\,;\ i = 0, 1, \ldots\ldots, n \qquad \ldots(1)$$

As there are $2(n + 1)$ conditions in (1), $(2n + 2)$ coefficients are to be determined. Therefore $P(x)$ is a polynomial of degree $(2n + 1)$.

We assume that $P(x)$ is expressible in the form

$$P(x) = \sum_{i=0}^{n} U_i(x)\, y(x_i) + \sum_{i=0}^{n} V_i(x)\, y'(x_i) \qquad \ldots(2)$$

where $U_i(x)$ and $V_i(x)$ are polynomials in $x$ of degree $(2n + 1)$. These are to be determined. Using the conditions (1), we get

$$\left.\begin{aligned} &U_i(x_j) = \begin{cases} 0 \text{ when } i \neq j \\ 1 \text{ when } i = j \end{cases}; \quad V_i(x_j) = 0 \text{ for all } i \\ &U_i'(x_j) = 0, \text{ for all } i\,; \quad V_i'(x_j) = \begin{cases} 0 \text{ when } i \neq j \\ 1 \text{ when } i = j \end{cases} \end{aligned}\right\} \quad \ldots(3)$$

We now write

$$U_i(x) = A_i(x)\,[L_i(x)]^2\,;\; V_i(x) = B_i(x)\,[L_i(x)]^2$$

where
$$L_i(x) = \frac{(x - x_0)(x - x_1)\ldots\ldots(x - x_{i-1})(x - x_{i+1})\ldots\ldots(x - x_n)}{(x_i - x_0)(x_i - x_1)\ldots\ldots(x_i - x_{i-1})(x_i - x_{i+1})\ldots\ldots(x_i - x_n)}$$

Since $[L_i(x)]^2$ is of degree $2n$ and $U_i(x)$, $V_i(x)$ are of degree $(2n + 1)$, therefore $A_i(x)$ and $B_i(x)$ are both linear functions.

$\therefore$ We can write
$$\left.\begin{aligned} U_i(x) &= (a_i + b_i x)\,[L_i(x)]^2 \\ \text{and} \quad V_i(x) &= (c_i + d_i x)[L_i(x)]^2 \end{aligned}\right\} \quad \ldots(4)$$

Using conditions (3) in (4), we get
$$\left.\begin{aligned} &a_i + b_i x = 1,\; c_i + d_i x = 0 \\ \text{and} \quad &b_i + 2L_i'(x_i) = 0,\; d_i = 1 \end{aligned}\right\} \quad \ldots(5)$$

Solving these equations, we obtain
$$\left.\begin{aligned} &b_i = -\,2L_i'(x_i),\; a_i = 1 + 2x_i L_i'(x_i) \\ &d_i = 1 \quad \text{and} \quad c_i = -\,x_i \end{aligned}\right\} \quad \ldots(6)$$

Now putting the above values in (4), we get

$$U_i(x) = [1 + 2x_i L_i'(x_i) - 2x L_i'(x_i)][L_i(x)]^2 = [1 - 2(x - x_i)L_i'(x_i)][L_i(x)]^2$$

and
$$V_i(x) = (x - x_i)\,[L_i(x)]^2$$

Finally substituting $U_i(x)$ and $V_i(x)$ in (2), we obtain

$$P(x) = \sum_{i=0}^{n} [1 - 2(x - x_i)\,L_i'(x_i)][L_i(x)]^2\,y(x_i) + \sum_{i=0}^{n} (x - x_i)[L_i(x)]^2\,y'(x_i) \quad \ldots(7)$$

This is the required *Hermite's interpolation formula* which is sometimes known as *osculating interpolation formula.*

**Obs.** In comparison to Lagrange's interpolation formula, the Hermite interpolation formula is computationally uneconomical.

**Example 7.27.** *For the following data :*

| $x$ : | $f(x)$ | $f'(x)$ |
|---|---|---|
| 0.5 | 4 | − 16 |
| 1 | 1 | − 2 |

*Find the Hermite interpolating polynomial.* (U.P.T.U., B. Tech., 2008)

**Sol.** We have $x_0 = 0.5$, $x_1 = 1$, $y(x_0) = 4$, $y(x_1) = 1$; $y'(x_0) = -\,16$, $y'(x_1) = -\,2$.

Also
$$L(x_0) = \frac{x - x_1}{x_0 - x_1} = \frac{x - 1}{-\,0.5} = -\,2\,(x - 1)\,;\; L'(x_0) = -\,2$$

$$L(x_1) = \frac{x - x_0}{x_1 - x_0} = \frac{x - 0.5}{1 - 0.5} = 2x - 1\,;\; L'(x_1) = 2$$



Hermite's interpolation formula in this case, is

$$\begin{aligned} P(x) &= [1 - 2\,(x - x_0)\,L'(x_0)]\,[L(x_0)]^2\,y(x_0) + (x - x_0)\,[L\,(x_0)]^2\,y'\,(x_0) \\ &\quad + [1 - 2\,(x - x_1)\,L'(x_1)]\,[L(x_1)]^2\,y(x_1) + (x - x_1)\,[L(x_1)]^2\,y'(x_1) \\ &= [1 - 2\,(x - 0.5)\,(-\,2)]\,[-\,2\,(x - 1)]^2\,(4) + (x - 0.5)\,[-\,2\,(x - 1)]^2\,(-\,16) \\ &\quad + [1 - 2\,(x - 1)\,(2)]\,(2x - 1)^2\,(1) + (x - 1)\,(2x - 1)^2\,(-\,2) \\ &= 16\,[1 + 4\,(x - 0.5]\,(x^2 - 2x + 1) - 164\,(x - 0.5)\,(x^2 - 2x + 1) \\ &\quad + [1 - 4\,(x - 1)]\,(4x^2 - 4x + 1) - 2(x - 1)\,(4x^2 - 4x + 1) \end{aligned}$$

Hence $P(x) = -\,24x^3 + 324x^2 - 130x + 23$

**Example 7.28.** *Apply Hermite's formula to interpolate for sin (1.05) from the following data :*

| *x* | *sin x* | *cos x* |
|---|---|---|
| *1.00* | *0.84147* | *0.54030* |
| *1.10* | *0.89121* | *0.45360* |

**Sol.** Here $y(x) = \sin x$ and $y'(x) = \cos x$

so that $y(x_0) = 0.84147$, $y'(x_0) = 0.54030$, $y(x_1) = 0.89121$, $y'(x_1) = 0.45360$

Also
$$L(x_0) = \frac{x - x_1}{x_0 - x_1} = 11 - 10x,\; L'(x_0) = -\,10$$

$$L(x_1) = \frac{x - x_0}{x_1 - x_0} = -\,10 + 10x,\; L'(x_1) = 10$$

Hence the Hermite's interpolation formula in this case is

$$\begin{aligned} P(x) &= [1 - 2(x - x_0)\,L'(x_0)]\,[L(x_0)]^2\,y(x_0) + (x - x_0)\,[L(x_0)]^2\,y'(x_0) \\ &\quad + [1 - 2\,(x - x_1)\,L'(x_1)]\,[L(x_1)]^2\,y(x_1) + (x - x_1)\,[L(x_1)]^2\,y'(x_1) \\ &= [1 - 2(x - 1)(-\,10)]\,(11 - 10x)^2\,(0.84147) + (x - 1)\,(-\,11 + 10x)^2\,(0.54030) \\ &\quad + [1 - 2\,(x - 1.1)\,(10)]\,(-\,10 + 10x)^2\,(0.89121) \\ &\quad + (x - 1.1)\,(-\,10 + 10x)^2\,(0.4536) \end{aligned}$$

Putting $x = 1.05$ in $P(x)$, we get

$$\begin{aligned} \sin\,(1.05) &= 1 - 2(0.05)(-\,10)\,[-\,10(1.05) + 11]^2\,(0.84147) + (0.05)(-\,0.5)^2\,(0.54030) \\ &\quad + [1 - 2(0.05)(10)]\,(0.5)^2 \times (0.89121) + (-\,0.05)(0.5)^2(0.4536) = 0.86742. \end{aligned}$$

**Example 7.29.** *Determine the Hermite polynomial of degree 4 which fits the following data :*

| | | | |
|---|---|---|---|
| *x* : | *0* | *1* | *2* |
| *y(x)* : | *0* | *1* | *0* |
| *y′(x)* : | *0* | *0* | *0* |

**Sol.** Here $x_0 = 0$, $x_1 = 1$, $x_2 = 2$, $y(x_0) = 0$, $y(x_1) = 1$, $y(x_2) = 0$ and $y'(x_0) = 0$, $y'(x_1) = 0$, $y'(x_2) = 0$.

Hermite's formula in this case is

$$\begin{aligned} P(x) &= [1 - 2L_0'\,(x_0)(x - x_0)]\,[L_0(x)]^2\,y(x_0) + (x - x_0)[L_0(x)]^2\,y'(x_0) \\ &\quad + [1 - 2L_1'\,(x_1)\,(x - x_1)][L_1(x)]^2 \times y(x_1) + (x - x_1)\,[L_1(x)]^2\,y'(x_1) \\ &\quad + [1 - 2L_2'\,(x_2)(x - x_2)] \times [L_2(x)]^2\,y(x_2) + (x - x_2)[L_2(x)]^2\,y'(x_2) \end{aligned}$$

Substituting the above values in $P(x)$, we get

$$P(x) = [1 - 2L_1'(x_1)(x - 1)][L_1(x)]^2$$

where $\quad L_1(x) = \dfrac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = 2x - x^2$ and $L_1'(x_1) = (2 - 2x)_{x=1} = 0$

Hence $P(x) = [L_1(x)]^2 = (2x - x^2)^2$.

**Example 7.30.** *Using Hermita's intropolation, find the value of $f(-0.5)$ from the following data.*

| | | | |
|---|---|---|---|
| $x$ : | $-1$ | $0$ | $1$ |
| $f(x)$ : | $1$ | $1$ | $3$ |
| $f'(x)$ : | $-5$ | $1$ | $7$ |

(*Anna, B. Tech.,* 2005)

**Sol.** Here $x_0 = -1, x_1 = 0, x_2 = 1$ ; $f(x_0) = 1, f(x_1) = 1, f(x_2) = 3$ and $f'(x_0) = -5, f'(x_1) = 1, f'(x_2) = 7$.

Hermite's formula in this case is

$$P(x) = U_0 f(x_0) + V_0 f'(x_0) + U_1 f(x_1) + V_1 f'(x_1) + U_2 f(x_2) + V_2 f'(x_2) \quad \text{...(i)}$$

where $U_0 = [1 - 2L_0'(x_0)(x - x_0)]\,[L_0(x)]^2,\ V_0 = (x - x_0)\,[L_0(x)]^2$

$U_1 = [1 - 2L_1'(x_1)(x - x_1)]\,[L_1(x)]^2,\ V_1 = (x - x_1)\,[L_1(x)]^2$

$U_2 = [1 - 2L_2'(x_2)(x - x_2)]\,[L_2(x_2)]^2,\ V_2 = (x - x_2)\,[L_2(x)]^2$

and $L_0(x) = \dfrac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \dfrac{x(x-1)}{2},\ L_0'(x) = x - \dfrac{1}{2}$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = 1 - x^2,\ L_1'(x) = -2x$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{x(x+1)}{2},\ L_2'(x) = x + \frac{1}{2}.$$

Substituting the values of $L_0, L_0'$ ; $L_1, L_1'$ and $L_2, L_2'$, we get

$$U_0 = [1 + 3(x + 1)]\,\frac{x^2(x-1)^2}{4} = \frac{1}{4}(3x^5 - 2x^4 - 5x^3 + 4x^2)$$

$$V_0 = (x + 1)\,\frac{x^2(x-1)^2}{4} = \frac{1}{4}(x^5 - x^4 - x^3 + x^2)$$

$$U_1 = x^4 - 2x^2 + 1, V_1 = x^5 - 2x^3 + x$$

$$U_2 = -\frac{1}{4}(3x^5 + 2x^4 - 5x^3 - 4x^2),\ V_2 = \frac{1}{4}(x^5 + x^4 - x^3 - x^2)$$

Substituting the values of $U_0, V_0, U_1, V_1$ ; $U_2, V_2$ in (*i*), we get

$$\begin{aligned} P(x) &= \frac{1}{4}(3x^5 - 2x^4 - 5x^3 + 4x^2)(1) + \frac{1}{4}(x^5 - x^4 - x^3 + x^2)(-5) \\ &\quad + (x^4 - 2x^2 + 1)(1) + (x^5 - 2x^3 + x)(1) \\ &\quad - \frac{1}{4}(3x^5 + 2x^4 - 5x^3 - 4x^2)(3) + \frac{1}{4}(x^5 + x^4 - x^3 - x^2)(7) \\ &= 2x^4 - x^2 + x + 1 \end{aligned}$$

Hence $\ f(-0.5) = 2(-0.5)^4 - (-0.5)^2 + (-0.5) + 1 = 0.375.$



## PROBLEMS 7.5

1. Find the Hermite's polynomial which fits the following data :

| | | | |
|---|---|---|---|
| $x$ : | 0 | 1 | 2 |
| $f(x)$ : | 1 | 3 | 21 |
| $f'(x)$ : | 0 | 3 | 36 |

2. A switching path between parallel rail road tracks is to be a cubic polynomial joining positions (0, 0) and (4, 2) and tangents to the lines $y = 0$ and $y = 2$. Using Hermite's method, find the polynomial, given :

| $x$ | $y$ | $y'$ |
|---|---|---|
| 0 | 0 | 0 |
| 4 | 2 | 0 |

3. Apply Hermite's formula to estimate the value of log 3.2 from the following data :

| $x$ | $y = \log x$ | $y' = 1/x$ |
|---|---|---|
| 3.0 | 1.0986 | 0.3333 |
| 3.5 | 1.2528 | 0.2857 |
| 4.0 | 1.3863 | 0.2500 |

## 7.17. (1) SPLINE INTERPOLATION

In the interpolation methods so far explained, a single polynomial has been fitted to the tabulated points. If the given set of points belong to the polynomial, then this method works well, otherwise the results are rough approximations only. If we draw lines through every two closest points, the resulting graph will not be smooth. Similarly we may draw a quadratic curve through points $A_i, A_{i+1}$ and another quadratic curve through $A_{i+1}, A_{i+2}$, such that the slopes of the two quadratic curves match at $A_{i+1}$ (Fig. 7.1). The resulting curve looks better but is not quite smooth. We can ensure this by drawing a cubic curve through $A_i, A_{i+1}$ and another cubic through $A_{i+1}, A_{i+2}$ such that the slopes and curvatures of the two curves match at $A_{i+1}$. Such a curve is called a **cubic spline**. We may use polynomials of higher order but the resulting graph is not better. As such, cubic splines are commonly used. This technique of 'spline-fitting' is of recent origin and has important applications.

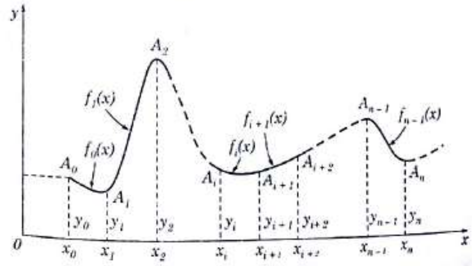


Fig. 7.1

**(2) Cubic spline**

Consider the problem of interpolating between the data points $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ by means of spline fitting.

Then the cubic spline $f(x)$ is such that

(*i*) $f(x)$ is a linear polynomial outside the interval $(x_0, x_n)$,

(*ii*) $f(x)$ is a cubic polynomial in each of the subintervals,

(*iii*) $f'(x)$ and $f''(x)$ are continuous at each point.

Since $f(x)$ is cubic in each of the subintervals $f''(x)$ shall be linear.

$\therefore$ Taking equally-spaced values of $x$ so that $x_{i+1} - x_i = h$, we can write

$$f''(x) = \frac{1}{h}\left[(x_{i+1} - x) f''(x_i) + (x - x_i) f''(x_{i+1})\right]$$

Integrating twice, we have

$$f(x) = \frac{1}{h}\left[\frac{(x_{i+1} - x)^3}{3!} f''(x_i) + \frac{(x - x_i)^3}{3!} f''(x_{i+1})\right] + a_i(x_{i+1} - x) + b_i(x - x_i) \qquad \ldots(1)$$

The constants of integration $a_i$, $b_i$ are determined by substituting the values of $y = f(x)$ at $x_i$ and $x_{i+1}$. Thus

$$a_i = \frac{1}{h}\left[y_i - \frac{h^2}{3!} f''(x_i)\right] \quad \text{and} \quad b_i = \frac{1}{h}\left[y_{i+1} - \frac{h^2}{3!} f''(x_{i+1})\right]$$

Substituting the values of $a_i$, $b_i$ and writing $f''(x_i) = M_i$, (1) takes the form

$$f(x) = \frac{(x_{i+1} - x)^3}{6h} M_i + \frac{(x - x_i)^3}{6h} M_{i+1}$$

$$+ \frac{x_{i+1} - x}{h}\left(y_i - \frac{h^2}{6} M_i\right) + \frac{x - x_i}{h}\left(y_{i+1} - \frac{h^2}{6} M_{i+1}\right) \qquad \ldots(2)$$

$$\therefore \quad f'(x) = -\frac{(x_{i+1} - x)^2}{2h} M_i + \frac{(x - x_i)^2}{h} M_{i+1} - \frac{h}{6}(M_{i+1} - M_i) + \frac{1}{h}(y_{i+1} - y_i)$$

To impose the condition of continuity of $f'(x)$, we get

$$f'(x - \epsilon) = f'(x + \epsilon) \text{ as } \epsilon \to 0.$$

$$\therefore \quad \frac{h}{6}(2M_i + M_{i-1}) + \frac{1}{h}(y_i - y_{i-1}) = -\frac{h}{6}(2M_i + M_{i+1}) + \frac{1}{h}(y_{i+1} - y_i)$$

or
$$M_{i-1} + 4M_i + M_{i+1} = \frac{6}{h^2}(y_{i-1} - 2y_i + y_{i+1}), \; i = 1 \text{ to } n - 1 \qquad \ldots(3)$$

Now since the graph is linear for $x < x_0$ and $x > x_n$, we have

$$M_0 = 0, \; M_n = 0 \qquad \ldots(4)$$

(3) and (4) give $(n + 1)$ equations in $(n + 1)$ unknowns $M_i$ ($i = 0, 1, \ldots, n$) which can be solved. Substituting the value of $M_i$ in (2) gives the concerned cubic spline.



**Example 7.31.** *Obtain the cubic spline for the following data :*

| | | | | |
|---|---|---|---|---|
| $x$ : | 0 | 1 | 2 | 3 |
| $y$ : | 2 | − 6 | − 8 | 2 |

(U.P.T.U., M.C.A., 2009)

**Sol.** Since the points are equispaced with $h = 1$ and $n = 3$, the cubic spline can be determined from $M_{i-1} + 4M_i + M_{i+1} = 6(y_{i-1} - 2y_i + y_{i+1})$, $i = 1, 2$.

$$\therefore \quad M_0 + 4M_1 + M_2 = 6(y_0 - 2y_1 + y_2)$$
$$M_1 + 4M_2 + M_3 = 6(y_1 - 2y_2 + y_3)$$

*i.e.*,
$$4M_1 + M_2 = 36 \; ; \; M_1 + 4M_2 = 72 \qquad [\because \; M_0 = 0, M_3 = 0]$$

Solving these, we get $M_1 = 4.8$, $M_2 = 16.8$.

Now the cubic spline in $(x_i \le x \le x_{i+1})$ is

$$f(x) = \frac{1}{6}(x_{i+1} - x)^3 M_i + \frac{1}{6}(x - x_i)^3 M_{i+1} + (x_{i+1} - x)\left(y_i - \frac{1}{6} M_i\right)$$
$$+ (x - x_i)\left(y_{i+1} - \frac{1}{6} M_{i+1}\right) \qquad \ldots(A)$$

Taking $i = 0$ in $(A)$ the cubic spline in $(0 \le x \le 1)$ is

$$f(x) = \frac{1}{6}(1 - x)^3(0) + \frac{1}{6}(x - 0)^3(4.8) + (1 - x)(x - 0) + x\left[-6 - \frac{1}{6}(4.8)\right]$$
$$= 0.8x^3 - 8.8x + 2 \quad (0 \le x \le 1)$$

Taking $i = 1$ in $(A)$, the cubic spline in $(1 \le x \le 2)$ is

$$f(x) = \frac{1}{6}(2 - x)^3(4.8) + \frac{1}{6}(x - 1)^3(16.8) + (2 - x)\left[-6 - \frac{1}{6}(4.8)\right]$$
$$+ (x - 1)\left[-8 - \frac{1}{6}(16.8)\right]$$
$$= 2x^3 - 5.84x^2 - 1.68x + 0.8$$

Taking $i = 2$ in $(A)$, the cubic spline in $(2 \le x \le 3)$ is

$$f(x) = \frac{1}{6}(3 - x)^3(4.8) + \frac{1}{6}(x - 2)^3(0) + (3 - x)\left[-8 - \frac{1}{6}(16.8)\right]$$
$$+ (x - 2)\left[2 - \frac{1}{6}(2)\right]$$
$$= -0.8x^3 + 2.64x^2 + 9.68x - 14.8$$

**Example 7.32.** *The following values of x and y are given :*

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 1 | 2 | 5 | 11 |

*Find the cubic splines and evaluate y(1.5) and y′(3).*

**Sol.** Since the points are equispaced with $h = 1$ and $n = 3$, the cubic splines can be obtained from

$$\therefore \quad M_{i-1} + 4M_i + M_{i+1} = 6(y_{i-1} - 2y_i + y_{i+1}), \; i = 1, 2.$$
$$M_0 + 4M_1 + M_2 = 6(y_0 - 2y_1 + y_2)$$
$$M_1 + 4M_2 + M_3 = 6(y_1 - 2y_2 + y_3)$$

*i.e.,* $4M_1 + M_2 = 12, M_1 + 4M_2 = 18$ [$\because$ $M_0 = 0, M_3 = 0$]

which give $M_1 = 2, M_2 = 4$.

Now the cubic spline in $(x_i \le x \le x_{i+1})$ is

$$f(x) = \frac{1}{6}[(x_{i+1} - x)^3 M_i + (x - x_i)^3 M_{i+1}] + (x_{i+1} - x)\left(y_i - \frac{1}{6}M_i\right) + (x - x_i)\left(y_{i+1} - \frac{1}{6}M_{i+1}\right) \quad ...(A)$$

Thus, taking $i = 0, i = 1, i = 2$ in $(A)$, the cubic splines are

$$f(x) = \begin{cases} \frac{1}{3}(x^3 - 3x^2 + 5x), & 1 \le x \le 2 \\ \frac{1}{3}(x^3 - 3x^2 + 5x), & 2 \le x \le 3 \\ \frac{1}{3}(-2x^3 + 24x^2 - 76x + 81), & 3 \le x \le 4 \end{cases}$$

$\therefore$ $y(1.5) = f(1.5) = 11/8.$

Also $y'(3) = 14/3$, from both the splines of intervals [2, 3] and [3, 4] as they should be.

**Example 7.33.** *Find the cubic spline interpolation for the data :*

| $x$ : | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $f(x)$ : | 1 | 0 | 1 | 0 | 1 |

*(Anna, B.E., 2007)*

**Sol.** Since the points are equispaced with $h = 1$, $n = 4$, the cubic spline can be found by means of

$$M_{i-1} + 4M_i + M_{i+1} = 6(y_{i-1} - 2y_i + y_{i+1}), \; i = 1, 2, 3$$

$\therefore$
$$M_0 + 4M_1 + M_2 = 6(y_0 - 2y_1 + y_2) = 12$$
$$M_1 + 4M_2 + M_3 = 6(y_1 - 2y_2 + y_3) = -12$$
$$M_2 + 4M_3 + M_4 = 6(y_2 - 2y_3 + y_4) = 12$$

Since $M_0 = y''_0 = 0$ and $M_4 = y''_4 = 0$

$\therefore$ $4M_1 + M_2 = 12 \; ; M_1 + 4M_2 + M_3 = -12 \; ; M_1 + 4M_3 = 12$

Solving these equations, we get $M_1 = 30/7, M_2 = -36/7, M_3 = 30/7$

Now the cubic spline in $(x_i \le x \le x_{i+1})$ is $f(x) = \frac{1}{6}(x_{i+1} - x) M_i + (x - x_i) M_{i+1} (x_{6+1} - x)$

$$\left(y_i - \frac{1}{6}M_i\right) + (x - x_i)\left(y_{i+1} - \frac{1}{6}M_{i+1}\right) \quad ...(A)$$

Taking $i = 0$, in $(A)$, the cubic spline in $(1 \le x \le 2)$ is

$$y = \frac{1}{6}[(x_1 - x)^3 M_0 + (x - x_0)^3 M_1] + (x_1 - x)\left(y_0 - \frac{1}{6}M_0\right) + (x - x_0)\left(y_1 - \frac{1}{6}M_1\right)$$

$$= \frac{1}{6}[(2 - x)^3 (0) + (x - 1)^3 (30/7)] + (2 - x)\left[1 - \frac{1}{6}(0)\right] + (x - 1)\left[0 - \frac{1}{6}\left(\frac{30}{7}\right)\right]$$

*i.e.,* $y = 0.71x^3 - 2.14x^2 + 0.42x + 2 \quad (1 < x \le 2)$

Taking $i = 1$ in $(A)$, the cubic spline in $(2 \le x \le 3)$ is

$$y = \frac{1}{6}\left[(3 - x)^3 \frac{30}{7} + (x - 2)^3\left(-\frac{36}{7}\right)\right] + (3 - x)\left[0 - \frac{1}{6}\left(\frac{30}{7}\right)\right] + (x - 2)\left[1 - \frac{1}{6}\left(-\frac{36}{7}\right)\right]$$

*i.e.,* $y = -1.57x^3 + 11.57x^2 - 27x + 20.28. \quad (2 \le x \le 3)$

Taking $i = 2$ in $(A)$, the cubic spline in $(3 \le x \le 4)$ is



$$y = \frac{1}{6}\left[(4 - x)^3\left(-\frac{36}{7}\right) + (x - 3)^3\left(\frac{30}{7}\right)\right] + (4 - x)\left[1 - \frac{1}{6}\left(-\frac{36}{7}\right)\right] + (x - 3)\left(0 - \frac{5}{7}\right)$$

*i.e.,* $y = 1.57x^3 - 16.71x^2 + 57.86x - 64.57 \quad (3 \le x \le 4)$

Taking $i = 3$ in $(A)$, the cubic spline in $(4 \le x \le 5)$ is

$$y = \frac{1}{6}\left[(1 - x)^3\left(\frac{30}{7}\right)\right] + (5 - x)\left(-\frac{5}{7}\right) + (x - 4)(1)$$

*i.e.,* $y = -0.71x^3 + 2.14x^2 - 0.43x - 6.86. \quad (4 \le x \le 5)$

## PROBLEMS 7.6

1. Find the cubic splines for the following table of values :

| | | | |
|---|---|---|---|
| $x$ : | 1 | 2 | 3 |
| $y$ : | −6 | −1 | 16 |

Hence evaluate $y(1.5)$ and $y'(2)$.

2. The following values of $x$ and $y$ are given :

| | | | | |
|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 |
| $y$ : | 1 | 5 | 11 | 8 |

Using cubic splines, show that

(*i*) $y(1.5) = 2.575$ (*ii*) $y'(3) = 2.067$.

3. Find the cubic spline corresponding to the interval [2, 3] from the following table :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1 | 2 | 3 | 4 | 5 |
| $y$ : | 30 | 15 | 32 | 18 | 25 |

Hence compute (*i*) $y(2.5)$ and (*ii*) $y'(3)$.

## 7.18. DOUBLE INTERPOLATION

So far, we have derived interpolation formulae to approximate a function of a single variable. In case of functions, of two variables, we interpolate with respect to the first variable keeping the other variable constant. Then interpolate with respect to the second variable.

Similarly, we can extend the said procedure for functions of three variables.

## 7.19. INVERSE INTERPOLATION

So far, given a set of values of $x$ and $y$, we have been finding the value of $y$ corresponding to a certain value of $x$. On the other hand, the process of estimating the value of $x$ for a value of $y$ (which is not in the table) is called *inverse interpolation*. When the values of $x$ are unequally spaced *Lagrange's method* is used and when the values of $x$ are equally spaced, the *Iterative method* should be employed.

## 7.20. LAGRANGE'S METHOD

This procedure is similar to Lagrange's interpolation formula (p. 207), the only difference being that $x$ is assumed to be expressible as a polynomial in $y$.

Lagrange's formula is merely a relation between two variables either of which may be taken as the independent variable. Therefore, on inter-changing $x$ and $y$ in the Lagrange's formula, we obtain

$$x = \frac{(y-y_1)(y-y_2)\ldots\ldots(y-y_n)}{(y_0-y_1)(y_0-y_2)\ldots\ldots(y_0-y_n)} x_0 + \frac{(y-y_0)(y-y_2)\ldots\ldots(y-y_n)}{(y_1-y_0)(y_1-y_2)\ldots\ldots(y_1-y_n)} x_1 + \ldots\ldots + \frac{(y-y_0)(y-y_1)\ldots\ldots(y-y_{n-1})}{(y_n-y_0)(y_n-y_1)\ldots\ldots(y_n-y_{n-1})} \quad \ldots(1)$$

which is used for inverse interpolation.

**Example 7.34.** *The following table gives the values of x and y :*

| | | | | | | |
|---|---|---|---|---|---|---|
| *x :* | *1.2* | *2.1* | *2.8* | *4.1* | *4.9* | *6.2* |
| *y :* | *4.2* | *6.8* | *9.8* | *13.4* | *15.5* | *19.6* |

*Find the value of x corresponding to y = 12, using Lagrange's technique.*

(V.T.U., B.E., 2009)

**Sol.** Here $x_0 = 1.2,\ x_1 = 2.1,\ x_2 = 2.8,\ x_3 = 4.1,\ x_4 = 4.9,\ x_5 = 6.2$

and $y_0 = 4.2,\ y_1 = 6.8,\ y_2 = 9.8,\ y_3 = 13.4,\ y_4 = 15.5,\ y_5 = 19.6.$

Taking $y = 12$, the above formula (1) gives

$$x = \frac{(12-6.8)(12-9.8)(12-13.4)(12-15.5)(12-19.6)}{(4.2-6.8)(4.2-9.8)(4.2-13.4)(4.2-15.5)(4.2-19.6)} \times 1.2$$
$$+ \frac{(12-4.2)(12-9.8)(12-13.4)(12-15.5)(12-19.6)}{(6.8-4.2)(6.8-9.8)(6.8-13.4)(6.8-15.5)(6.8-19.6)} \times 2.1$$
$$+ \frac{(12-4.2)(12-6.8)(12-13.4)(12-15.5)(12-19.6)}{(9.8-4.2)(9.8-6.8)(9.8-13.4)(9.8-15.5)(9.8-19.6)} \times 2.8$$
$$+ \frac{(12-4.2)(12-6.8)(12-9.8)(12-15.5)(12-19.6)}{(13.4-4.2)(13.4-6.8)(13.4-9.8)(13.4-15.5)(13.4-19.6)} \times 4.1$$
$$+ \frac{(12-4.2)(12-6.8)(12-9.8)(12-13.4)(12-19.6)}{(15.5-4.2)(15.5-6.8)(15.5-9.8)(15.5-13.4)(15.5-19.6)} \times 4.9$$
$$+ \frac{(12-4.2)(12-6.8)(12-9.8)(12-13.4)(12-15.5)}{(19.6-4.2)(19.6-6.8)(19.6-9.8)(19.6-13.4)(19.6-15.5)} \times 6.2$$

$$= 0.022 - 0.234 + 1.252 + 3.419 - 0.964 + 0.055 = 3.55.$$

**Example 7.35.** *Apply Lagrange's formula inversely to obtain a root of the equation $f(x) = 0$, given that $f(30) = -30$, $f(34) = -13$, $f(38) = 3$, and $f(42) = 18$.* (V.T.U., B. Tech., 2009)

**Sol.** Here $x_0 = 30, x_1 = 34, x_2 = 38, x_3 = 42$

and $y_0 = -30, y_1 = -13, y_2 = 3, y_3 = 18$

It is required to find $x$ corresponding to $y = f(x) = 0$.

Taking $y = 0$, the Lagrange's formula gives,

$$x = \frac{(y-y_1)(y-y_2)(y-y_3)}{(y_0-y_1)(y_0-y_2)(y_0-y_3)} x_0 + \frac{(y-y_0)(y-y_2)(y-y_3)}{(y_1-y_0)(y_1-y_2)(y_1-y_3)} x_1 + \frac{(y-y_0)(y-y_1)(y-y_3)}{(y_2-y_0)(y_2-y_1)(y_2-y_3)} x_2 + \frac{(y-y_0)(y-y_1)(y-y_2)}{(y_3-y_0)(y_3-y_1)(y_3-y_2)} x_3$$



$$= \frac{13(-3)(-18)}{(-17)(-33)(-48)} \times 30 + \frac{30(-3)(-18)}{17(-16)(-31)} \times 34 + \frac{30(13)(-18)}{33(16)(-15)} \times 38 + \frac{30(13)(-3)}{48(31)(15)} \times 42$$

$$= -0.782 + 6.532 + 33.682 - 2.202 = 37.23.$$

Hence the desired root of $f(x) = 0$ is 37.23.

## 7.21. ITERATIVE METHOD

Newton's forward interpolation formula (p. 163) is

$$y_p = y_0 + p\Delta y_0 + \frac{p(p-1)}{2!}\Delta^2 y_0 + \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 + \ldots\ldots$$

From this, we get

$$p = \frac{1}{\Delta y_0}\left[y_p - y_0 - \frac{p(p-1)}{2!}\Delta^2 y_0 - \frac{p(p-1)(p-2)}{3!}\Delta^3 y_0 - \ldots\ldots\right] \quad \ldots(1)$$

Neglecting the second and higher differences, we obtain the first approximation to $p$ as

$$p_1 = (y_p - y_0)/\Delta y_0 \quad \ldots(2)$$

To find the second approximation, retaining the term with second differences in (1) and replacing $p$ by $p_1$, we get

$$p_2 = \frac{1}{\Delta y_0}\left[y_p - y_0 - \frac{p_1(p_1-1)}{2!}\Delta^2 y_0\right] \quad \ldots(3)$$

To find the third approximation, retaining the term with third differences in (1) and replacing every $p$ by $p_2$, we have

$$p_3 = \frac{1}{\Delta y_0}\left[y_p - y_0 - \frac{p_2(p_2-1)}{2!}\Delta^2 y_0 - \frac{p_2(p_2-1)(p_2-2)}{3!}\Delta^3 y_0\right]$$

and so on. This process is continued till two successive approximations of $p$ agree with each other.

**Obs.** This technique can equally well be applied by starting with any other interpolation formula.

This method is a powerful *iterative procedure for finding the roots of an equation to a good degree of accuracy.*

**Example 7.36.** *The following values of $y = f(x)$ are given*

| | | | |
|---|---|---|---|
| *x :* | *10* | *15* | *20* |
| *y :* | *1754* | *2648* | *3564* |

*Find the value of x for y = 3000 by iterative method.*

**Sol.** Taking $x_0 = 10$ and $h = 5$, the difference table is

| $x$ | $y$ | $\Delta y$ | $\Delta^2 y$ |
|---|---|---|---|
| 10 | 1754 | | |
| | | 894 | |
| 15 | 2648 | | 22 |
| | | 916 | |
| 20 | 3564 | | |

Here $y_p = 3000, y_0 = 1754, \Delta y_0 = 894$ and $\Delta^2 y_0 = 22$.

$\therefore$ The successive approximations to $p$ are

$$p_1 = \frac{1}{894}(3000 - 1754) = 1.39$$

$$p_2 = \frac{1}{894}\left[3000 - 1754 - \frac{1.39\,(1.39 - 1)}{2} \times 22\right] = 1.387$$

$$p_3 = \frac{1}{894}\left[3000 - 1754 - \frac{1.387\,(1.387 - 1)}{2} \times 22\right] = 1.3871$$

We, therefore, take $p = 1.387$ correct to three decimal places. Hence the value of $x$ (corresponding to $y = 3000$) $= x_0 + ph = 10 + 1.387 \times 5 = 16.935$.

**Example 7.37.** *Using inverse interpolation, find the real root of the equation $x^3 + x - 3 = 0$, which is close to 1.2.*

**Sol.** The difference table is

| $x$ | $v$ | $y(= x^3 + x - 3)$ | $\Delta y$ | $\Delta^2 y$ | $\Delta^3 y$ | $\Delta^4 y$ |
|---|---|---|---|---|---|---|
| 1 | − 0.2 | − 1 | | | | |
| | | | 0.431 | | | |
| 1.1 | − 0.1 | − 0.569 | | 0.066 | | |
| | | | 0.497 | | 0.006 | |
| 1.2 | 0 | − 0.072 | | 0.072 | | 0 |
| | | | 0.569 | | 0.006 | |
| 1.3 | 0.1 | 0.497 | | 0.078 | | |
| | | | 0.647 | | | |
| 1.4 | 0.2 | 1.144 | | | | |

Clearly the root of the given equation lies between 1.2 and 1.3. Assuming the origin at $x = 1.2$ and using Stirling's formula.

$$y = y_0 + x\,\frac{\Delta y_0 + \Delta y_{-1}}{2} + \frac{x^2}{2}\,\Delta^2 y_{-1} + \frac{x(x^2 - 1)}{6}\,\frac{(\Delta^3 y_{-1} + \Delta^3 y_{-2})}{2}, \text{ we get}$$

$$0 = -\,0.072 + x \,.\, \frac{0.569 + 0.467}{2} + \frac{x^2}{2}(0.072) + \frac{x(x^2 - 1)}{6}\,\frac{0.006 + 0.006}{2} \qquad (\because\ y = 0)$$

or $$0 = -\,0.072 + 0.532x + 0.036x^2 + 0.001x^3$$

This equation can be written as

$$x = \frac{0.072}{0.532} - \frac{0.036}{0.532}\,x^2 - \frac{0.001}{0.532}\,x^3 \qquad ...(i)$$



$\therefore$ First approximation $x^{(1)} = \dfrac{0.072}{0.532} = 0.1353$

Putting $x = x^{(1)}$ on R.H.S. of $(i)$, we get

Second approximation $x^{(2)} = 0.1353 - 0.067(0.1353)^2 - 1.8797(0.1353)^3 = 0.134$

Hence the desired root $= 1.2 + 0.1 \times 0.134 = 1.2134$.

## PROBLEMS 7.7

1. Apply Lagrange's method to find the value of $x$ when $f(x) = 15$ from the given data :

| | | | | |
|---|---|---|---|---|
| $x$ : | 5 | 6 | 9 | 11 |
| $f(x)$ : | 12 | 13 | 14 | 16 |

(*Madras, B.E., 2000*)

2. Obtain the value of $t$ when $A = 85$ from the following table, using Lagrange's method :

| | | | | |
|---|---|---|---|---|
| $t$ : | 2 | 5 | 8 | 14 |
| $A$ : | 94.8 | 87.9 | 81.3 | 68.7 |

3. Apply Lagrange's formula inversely to obtain the root of the equation $f(x) = 0$, given that $f(30) = -30$, $f(34) = -13$, $f(38) = 3$ and $f(42) = 18$.

(*Kerala, B. Tech., 1995 S*)

4. From the following data :

| | | | | | |
|---|---|---|---|---|---|
| $x$ : | 1.8 | 2.0 | 2.2 | 2.4 | 2.6 |
| $y$ : | 2.9 | 3.6 | 4.4 | 5.5 | 6.7. |

find $x$ when $y = 5$ using the iterative method.

5. The equation $x^3 - 15x + 4 = 0$ has a root close to 0.3, obtain this root upto 4 decimal places using inverse interpolation.
6. Solve the equation $x = 10 \log x$, by iterative method, given that

| | | | | |
|---|---|---|---|---|
| $x$ : | 1.35 | 1.36 | 1.37 | 1.38 |
| $\log x$ : | 0.1303 | 0.1355 | 0.1367 | 0.1392. |

## 7.22. OBJECTIVE TYPE OF QUESTIONS

## PROBLEMS 7.8

*Select the correct answer or fill up the blanks in the following questions :*

1. Newton's backward interpolation formula is ........ .
2. Bessel's formula is most appropriate when $p$ lies between
   (*a*) − 0. 25 and 0.25 (*b*) 0.25 and 0.75 (*c*) 0.75 and 1.00
3. Form the divided difference table for the following data :

| | | | |
|---|---|---|---|
| $x$ : | 5 | 15 | 22 |
| $y$ : | 7 | 36 | 160 |

4. Interpolation is the technique of estimating the value of a function for any
5. Bessel's formula for interpolation is ..... .
6. The 4th divided differences for $x_0, x_1, x_2, x_3, x_4$ = ...... .
7. Stirling's formula is best suited for $p$ lying between ..... .
8. Newton's divided differences formula is ...... .

9. Given $(x_0, y_0), (x_1, y_1), (x_2, y_2)$, Lagrange's interpolation formula is ...... .
10. If $f(0) = 1, f(2) = 5, f(3) = 10$ and $f(x) = 14$, then $x =$ ..... .
11. Difference between Lagrange's interpolating polynomial and Hermite's interpolating polynomial is ...... .
12. If $y(1) = 4, y(3) = 12, y(4) = 19$ and $y(x) = 7$, find $x$ using Lagrange's formula.
13. Extrapolation is defined as ...... .
14. The second divided difference of $f(x) = 1/x$, with arguments $a, b, c$ is ......

(*Anna, B.E.*, 2007)

15. Gauss-forward interpolation formula is used to interpolate values of $y$ for
    (*a*) $0 < p < 1$ (*b*) $-1 < 1 < 0$
    (*c*) $0 < p < -\alpha$ (*d*) $-\alpha < p < 0$.
16. Given

| | | | | |
|---|---|---|---|---|
| $x$: | 0 | 1 | 3 | 4 |
| $y$: | $-12$ | 0 | 6 | 12 |

Using Lagrange's formula, a polynomial that can be fitted to the data is ......

17. The $n$th divided difference of a polynomial of degree $n$ is
    (*a*) zero (*b*) a constant
    (*c*) a variable (*d*) none of these.
18. Gauss forward interpolation formula involves
    (*a*) even differences above the central line and odd differences on the central line
    (*b*) even differences below the central line and odd differences on the central line
    (*c*) odd differences below the central line and even differences on the central line
    (*d*) odd differences above the central line and even differences on the central line.

# 8

# NUMERICAL DIFFERENTIATION & INTEGRATION



R

## ~~8.1.~~ NUMERICAL DIFFERENTIATION

It is the process of calculating the value of the derivative of a function at some assigned value of $x$ from the given set of values $(x_i, y_i)$. To compute $dy/dx$, we first replace the exact relation $y = f(x)$ by the best interpolating polynomial $y = \phi(x)$ and then differentiate the latter as many times as we desire. The choice of the interpolation formula to be used, will depend on the assigned value of $x$ at which $dy/dx$ is desired.

If the values of $x$ are equi-spaced and $dy/dx$ is required near the beginning of the table, we employ Newton's forward formula. If it is required near the end of the table, we use Newton's backward formula. For values near the middle of the table, $dy/dx$ is calculated by means of Stirling's or Bessel's formula.

If the values of $x$ are not equi-spaced, we use Newton's divided difference formula to represent the function.

Hence corresponding to each of the interpolation formulae, we can derive a formula for finding the derivative.

**Obs.** While using these formulae, it must be observed that the table of values defines the function at these points only and does not completely define the function and the function may not be differentiable at all. As such, *the process of numerical differentiation should be used only if the tabulated values are such that the differences of some order are constants.* Otherwise, errors are bound to creep in which go on increasing as derivatives of higher order are found. This is due to the fact that the difference between $f(x)$ and the approximating polynomial $\phi(x)$ may be small at the data points but $f'(x) - \phi'(x)$ may be large.

R

## ~~8.2~~ FORMULAE FOR DERIVATIVES

Consider the function $y = f(x)$ which is tabulated for the values $x_i (= x_0 + ih), i = 0, 1, 2, \ldots n$.