5 Numerical Solutions of a System of Linear Equations

5.1 Introduction:

System of linear algebraic equations arise in a large number of problems in science and technology. The most common form of the system in n unknowns $x_1, x_2, ..., x_n$ is of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n &= b_2 \\ ... \quad\quad ... \quad\quad &... \\ ... \quad\quad ... \quad\quad &... \\ a_{n1}x_1 + a_{n2}x_2 + ... + a_{nn}x_n &= b_n \end{aligned}$$

$$\text{i.e., } \sum_{j=1}^{n} a_{ij}x_j = b_i, \ (i = 1, 2, ..., n) \quad ... \quad (1)$$

where $a_{ij}(i, j = 1, 2, ...n)$ and $b_i(i = 1, 2, ...n)$ are given numbers. We can also write the equation (1) in the matrix form as

$$AX = b \quad ... \quad (2)$$

where
$$A = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix} = [a_{ij}]_{n \times n},$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_n \end{bmatrix} \quad ... \quad (3)$$

in which it is supposed that the matrix A is non-singular, i.e., det $A \neq 0$ so that the system (2) has *a* unique solution. The system of equations (1) is said to be homogeneous if all $b_i(i = 1, 2, ..., n)$, are zero; otherwise, the system is called non-homogeneous.

To solve the above system of equations we apply, in general, two methods viz **(i)** direct method and **(ii)** indirect or iterative method. In direct method, the solution is obtained after a finite

Scanned by CamScanner

number of steps of elementary arithmetical operations. On the other hand, an indirect or iterative method we start with an arbitrary initial approximation to x and then improve this estimate in an infinite but convergent sequence of steps.

We discuss in this chapter both the above methods in various ways.

Direct methods.

(i) Gauss elimination method

(ii)Matrix inversion method

(iii) LU Factorization method

Indirect or iterative methods

(i) Gauss-Seidel method

5.2. Gauss elimination method.

In this method, the given system of equations is reduced to an equivalent upper triangular system by a systematic elimination procedure from which the unknowns are found by back substitution.

To illustrate the mehtod, we consider the system (1) given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \ldots + a_{3n}x_n &= b_3 \\ \ldots \qquad \ldots & \qquad \ldots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nn}x_n &= b_n \end{aligned} \qquad \ldots \quad (4)$$

First suppose that $a_{11} \neq 0$

Multiply the first equation of (4) by $\dfrac{a_{i1}}{a_{11}} (i = 2, 3, \ldots, n)$ and subtract the results from the i-th equation, $(i = 2, 3, \ldots, n)$ and obtain

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)} \\ a_{32}^{(1)}x_2 + a_{33}^{(1)}x_3 + \ldots + a_{3n}^{(1)}x_n &= b_3^{(1)} \\ \ldots \qquad \ldots & \qquad \ldots \\ a_{n2}^{(1)}x_2 + a_{n3}^{(1)}x_3 + \ldots + a_{nn}^{(1)}x_n &= b_n^{(1)} \end{aligned} \qquad \ldots \quad (5)$$

where

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{1j}a_{i1}}{a_{11}} \qquad \ldots \quad (6)$$

$$b_i^{(1)} = b_i - \frac{b_1 a_{i1}}{a_{11}} \quad (i, j = 2, 3, \ldots, n)$$

The numbers $\dfrac{a_{i1}}{a_{11}}$, $(i = 2, 3, \ldots, n)$ are called row multipliers.

The first equation of the syestem (5) contains x_1 while the remaining $(n-1)$ equations are independent of x_1.

Next assume $a_{22}^{(1)} \neq 0$

Multiplying the second equation of (5) by $\dfrac{a_{i2}^{(1)}}{a_{22}^{(1)}}$, $(i = 3, 4, \ldots, n)$ and subtracting the results from the i-th equation, $(i = 3, 4, \ldots, n)$ we get

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)} \\ a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= b_3^{(2)} \\ \ldots \qquad \ldots & \qquad \ldots \\ a_{n3}^{(2)}x_3 + \ldots + a_{nn}^{(2)}x_n &= b_n^{(2)} \end{aligned} \qquad \ldots \quad (7)$$

where

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{2j}^{(1)} a_{i2}^{(1)}}{a_{22}^{(1)}},$$

$$b_i^{(2)} = b_i^{(1)} - \frac{b_2^{(1)} a_{i2}^{(1)}}{a_{22}^{(1)}}, \quad (i, j = 3, 4, \ldots, n) \qquad \ldots \quad (8)$$

Here also the numbers $\dfrac{a_{i2}^{(1)}}{a_{22}^{(1)}}$ are row multipliers.

In the system (7), the last $(n-2)$ equations are independent of x_1 and x_2.

10

Scanned by CamScanner

Repeating the procedure, we obtain a system of n equations equivalent to an upper triangular system in the form

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1n}x_n &= b_1 \\
a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \ldots + a_{2n}^{(1)}x_n &= b_2^{(1)} \\
a_{33}^{(2)}x_3 + \ldots + a_{3n}^{(2)}x_n &= b_3^{(2)} \\
\ldots \quad \ldots \quad & \ldots \\
a_{nn}^{(n-1)}x_n &= b_n^{(n-1)}
\end{aligned} \qquad \ldots (9)$$

The coefficients of the leading terms in (9), i.e., the elements $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, a_{nn}^{(n-1)}$ are called *pivotal elements* and the corresponding equations are known as *pivotal equations*. The solutions of the (4) are then obtained from (9) by back substitutions as

$$x_n = \frac{b_n^{(n-1)}}{a_{nn}^{(n-1)}} \qquad \ldots (10)$$

$$x_{n-1} = \frac{1}{a_{n-1,n-1}^{(n-2)}}\left[b_{n-1}^{(n-2)} - \frac{a_{n-1,n}^{(n-2)}b_n^{(n-1)}}{a_{nn}^{(n-1)}}\right]$$

etc., provided none of the pivotal element is zero.

The above procedure can also be explained in a more compact form by matrix notation as following :

The augmented matrix is

$$\begin{array}{l} m\,(multiplier) \\ a_{21}/a_{11} \\ a_{31}/a_{11} \\ \vdots \\ a_{n1}/a_{11} \end{array}
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \vdots & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & \vdots & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & \vdots & b_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & \vdots & b_n \end{bmatrix}, \text{ provided } a_{11} \neq 0 \quad (11)$$

After the first elimination, we have

$$\begin{array}{l} m\,(multiplier) \\ \\ a_{32}^{(1)}/a_{22}^{(1)} \\ \vdots \\ a_{n2}^{(1)}/a_{22}^{(1)} \end{array}
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \vdots & b_1 \\ & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\ & a_{32}^{(1)} & \cdots & a_{3n}^{(1)} & \vdots & b_3^{(1)} \\ & \cdots & \cdots & \cdots & \cdots & \cdots \\ & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & \vdots & b_n^{(1)} \end{bmatrix}, \text{ provided } a_{22}^{(1)} \neq 0 \qquad \ldots (12)$$

in which $a_{ij}^{(1)}$ and $b_i^{(1)}, (i = 2, 3, \ldots, n)$ are given by (6).

The second elimination gives

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & \vdots & b_1 \\ & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & \vdots & b_3^{(2)} \\ & & \cdots & \cdots & \cdots & \vdots & \\ & & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & \vdots & b_n^{(2)} \end{bmatrix} \qquad \ldots (13)$$

Repeating the process for $(n-1)$ times, we obtain the following upper triangular matrix :

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & \vdots & b_1 \\ & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & \vdots & b_2^{(1)} \\ & & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & \vdots & b_3^{(2)} \\ & & \cdots & \cdots & \cdots & \vdots & \\ & & \cdots & \cdots & a_{nn}^{(n-1)} & \vdots & b_n^{(n-1)} \end{bmatrix} \qquad \ldots (14)$$

which is equivalent to the system (9) and hence by back substitution we get the required solutions of the system.

Note. **(i)** In Gauss elimination method, the total number of multiplications and divisions is $\dfrac{n^3}{3} + n^2 - \dfrac{n}{3}$ and those of additions and subtractions is $\dfrac{n^3}{3} + \dfrac{n^2}{2} - \dfrac{5}{6}n$

Scanned by CamScanner

(ii) The method fails if any of the pivotal elements $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, a_{nn}^{(n-1)}$ is zero. In such cases, we rearrange the equations in such a way that the pivotal elements do not vanish. If it is not at all possible, then the solution of the given system does not exist.

Example.1. Solve the following system of linear equations by Gauss-elimination method.

$$x-2y+9z=8$$
$$3x+y-z=3$$
$$2x-8y+z=-5$$

[W.B.U.T., CS-312 2007, 2008
M(CS)-401, 2016]

Solution. In order to eliminate x from the last two equations, we multiply the first equation successively by 3, 2 and subtract the results from the second and third equations respectively. Thus we have

$$7y-28z=-21 \qquad \cdots \quad (1)$$
$$-4y-17z=-21 \qquad \cdots \quad (2)$$

In the next step, we eliminate y from (2) by multiplying the equation (1) by $\frac{4}{7}$ and add the result from (2) to get

$$-33z=-33.$$

Thus the given system of equations reduces to the following upper triangular form as

$$x-2y+9z=8$$
$$7y-28z=-21$$
$$-33z=-33$$

from which the back substitution leads to the required solution as

$$x=1, y=1, z=1$$

Example.2. Solve the following system of equations by Gauss elimination method :

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

[M.A.K.A.U.T., M(CS)-301, 2014,
M(CS)-401, 2015]

Solution. The augmented matrix of the given system of equations is

$$\begin{array}{c} m\,(multiplier) \\ 2 \\ -1 \end{array} \begin{bmatrix} 1 & 2 & 1 & : & 0 \\ 2 & 2 & 3 & : & 3 \\ -1 & -3 & 0 & : & 2 \end{bmatrix}$$

Using the row operations $R_2 - 2R_1$ and $R_3 + R_1$ we get

$$\begin{array}{c} m\,(multiplier) \\ \\ \frac{1}{2} \end{array} \begin{bmatrix} 1 & 2 & 1 & : & 0 \\ 0 & -2 & 1 & : & 3 \\ 0 & -1 & 1 & : & 2 \end{bmatrix}$$

Again using the row operation $R_3 - \frac{1}{2}R_2$ we have

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

Hence the given system of equations is reduced to the upper triangular form given by

$$x+2y+z=0$$
$$-2y+z=3$$
$$\frac{1}{2}z=\frac{1}{2}$$

$\therefore$ By back substitution, the resulting solutions are

$$x=1, y=-1, z=1.$$

5.3. Matrix inversion method.

For the system (2) viz.

$$AX=b, \qquad \cdots \quad (15)$$

we suppose that det $A \neq 0$ and so A^{-1} exists. Multiplying both side of (15) by A^{-1}, we get

$$X=A^{-1}b \qquad \cdots \quad (16)$$

which gives the solution of the given system

Scanned by CamScanner

Noting that

$$A^{-1} = \frac{adj\ A}{\det A}$$

$$= \frac{(A_{ji})_{n\times n}}{|a_{ij}|}, (i,\ j = 1, 2, \ldots, n)$$

we have

$$X = \frac{(A_{ji})_{n\times n}\, b}{|a_{ij}|} \quad \ldots (17)$$

where *adj A* is the transpose of the matrix obtained from A by replacing each element a_{ij} of A by its corresponding co-factor A_{ij} $(i, j = 1, 2, \ldots, n)$.

Note. (i) The method fails if the matrix A is singular i.e., $\det A = 0$

(ii) The method is not suitable for $n > 4$, since it involves laborious numerical computation.

Example.3. Solve the following system of equations :

$$x + y + z = 4$$
$$2x - y + 3z = 1$$
$$3x + 2y - z = 1$$

by matrix inversion method.

Solution. The given system of equations can be written as

$$AX = b \quad \ldots (1)$$

where $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 3 & 2 & -1 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

Now $\det A = \begin{vmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 3 & 2 & -1 \end{vmatrix} = 13 \neq 0$

Hence A is non-singular.

$\therefore A^{-1}$ exist.

Since $adj\ A = \begin{bmatrix} -5 & 3 & 4 \\ 11 & -4 & -1 \\ 7 & 1 & -3 \end{bmatrix}$, so

$$A^{-1} = \frac{adjA}{\det A} = \frac{1}{13}\begin{bmatrix} -5 & 3 & 4 \\ 11 & -4 & -1 \\ 7 & 1 & -3 \end{bmatrix}$$

$\therefore$ From (1), we have

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

which gives

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{13}\begin{bmatrix} -5 & 3 & 4 \\ 11 & -4 & -1 \\ 7 & 1 & -3 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}$$

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

Thus the required solutions are

$$x = -1,\ y = 3,\ z = 2$$

5.4. LU-factorization method.

This method is also termed as *triangular decomposition* method. The method based on the fact that every square matrix can be expressed as the product of a lower and an upper triangular matrix provided all the principal minors of the given square matrix $A = (a_{ij})_{n\times n}$ are non-singular, i.e.,

$$a_{11} \neq 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0, \ldots, \det A \neq 0 \quad \ldots (18)$$

Further, if the matrix A can be factorized, then it is unique.

Assume that it is possible to decompose the coefficient matrix A of the given system of equation (2) and is expressible as the product of a lower triangular matrix L and an upper triangular matrix U so that

$$A = LU \quad \ldots (19)$$

Scanned by CamScanner

where

$$L = \begin{bmatrix} l_{11} & 0 & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ l_{n1} & l_{n2} & l_{n3} & l_{n4} & \cdots & l_{nn} \end{bmatrix}, \quad \cdots (20)$$

$$U = \begin{bmatrix} 1 & u_{12} & u_{13} & u_{14} & \cdots & u_{1n} \\ 0 & 1 & u_{23} & u_{24} & \cdots & u_{2n} \\ 0 & 0 & 1 & u_{34} & \cdots & u_{3n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad \cdots (21)$$

Hence the system of equations

$$AX = b$$

become

$$LUX = b \quad \cdots (22)$$

Putting $UX = Y$ in (22) we get ... (23)

$$LY = b \quad \cdots (24)$$

where $Y = (y_1, y_2, \ldots, y_n)^T$

Thus by forward substitution, the unknowns $y_1, y_2, \ldots, y_n$ are determined from (24) and thereafter the unknowns $x_1, x_2, \ldots, x_n$ are obtained from (23) by backward substitution.

For the sake of clarity and simplicity we now consider a system of three equations with three unknowns viz

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned} \quad \cdots (25)$$

Here the coefficient matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

can be written as

$$A = LU$$

where $L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$, $U = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}$

$\therefore$ Thus we have

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

leading to

$$l_{11} = a_{11},\ l_{21} = a_{21},\ l_{31} = a_{31}$$

$$l_{11}u_{12} = a_{12},\ l_{11}u_{13} = a_{13} \Rightarrow u_{12} = \frac{a_{12}}{l_{11}}, u_{13} = \frac{a_{13}}{l_{11}},$$

$$l_{21}u_{12} + l_{22} = a_{22}, \Rightarrow l_{22} = a_{22} - l_{21}u_{12} = a_{22} - \frac{a_{21}a_{12}}{a_{11}}$$

$$l_{31}u_{12} + l_{32} = a_{32} \Rightarrow l_{32} = a_{32} - l_{31}u_{12} = a_{32} - \frac{a_{31}a_{12}}{a_{11}}$$

$$l_{21}u_{13} + l_{22}u_{23} = a_{23} \Rightarrow u_{23} = \frac{a_{23} - l_{21}u_{13}}{l_{22}}$$

and $l_{31}u_{13} + l_{32}u_{23} + l_{33} = a_{33}$

$$\Rightarrow l_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23}.$$

Thus obtained values of $l_{11}, l_{21}, \ldots$ and $u_{12}, u_{13}, \ldots$ gives th matrices L and U.

Scanned by CamScanner

Example.4. Solve the following system of equations by LU-factorization method.

$$8x_1 - 3x_2 + 2x_3 = 20$$
$$4x_1 + 11x_2 - x_3 = 33$$
$$6x_1 + 3x_2 + 12x_3 = 36$$

[W.B.U.T.,CS-312, 2004, 2016]

Solution. The given system of equations can be written as

$$AX = b$$

where $A = \begin{bmatrix} 8 & -3 & 2 \\ 4 & 11 & -1 \\ 6 & 3 & 12 \end{bmatrix}$, $b = \begin{bmatrix} 20 \\ 33 \\ 36 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

Let $A = LU$

where

$$L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \text{ and } U = \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & u_{13} \\ 0 & 1 & u_{23} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 8 & -3 & 2 \\ 4 & 11 & -1 \\ 6 & 3 & 12 \end{bmatrix}$$

leading to

$$l_{11} = 8, l_{11}u_{12} = -3 \Rightarrow u_{12} = -\frac{3}{8}$$

$$l_{11}u_{13} = 2 \Rightarrow u_{13} = \frac{2}{8} = \frac{1}{4}$$

$$l_{21} = 4, l_{21}u_{12} + l_{22} = 11 \Rightarrow l_{22} = 11 - l_{21}u_{12}$$

$$\Rightarrow l_{22} = 11 - 4\left(-\frac{3}{8}\right)$$

$$\Rightarrow l_{22} = \frac{25}{2}$$

$$l_{21}u_{13} + l_{22}u_{23} = -1$$

$$\Rightarrow 4.\frac{1}{4} + \frac{25}{2}.u_{23} = -1 \Rightarrow u_{23} = -\frac{4}{25}$$

$$l_{31} = 6, l_{31}u_{12} + l_{32} = 3 \Rightarrow 6\left(-\frac{3}{8}\right) + l_{32} = 3$$

$$\Rightarrow l_{32} = \frac{21}{4}$$

$$l_{31}u_{13} + l_{32}u_{23} + l_{33} = 12 \Rightarrow 6.\frac{1}{4} + \frac{21}{4}\left(\frac{-4}{25}\right) + l_{33} = 12$$

$$\Rightarrow l_{33} = \frac{567}{50}$$

Hence $L = \begin{bmatrix} 8 & 0 & 0 \\ 4 & \frac{25}{2} & 0 \\ 6 & \frac{21}{4} & \frac{567}{50} \end{bmatrix}$, $U = \begin{bmatrix} 1 & -\frac{3}{8} & \frac{1}{4} \\ 0 & 1 & \frac{-4}{25} \\ 0 & 0 & 1 \end{bmatrix}$

Thus the equation $AX = b$ i.e., $LUX = b$ gives

$$LY = b \qquad \dots (1)$$

where $UX = Y$

i.e., $$\begin{bmatrix} 1 & -\frac{3}{8} & \frac{1}{4} \\ 0 & 1 & \frac{-4}{25} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \qquad \dots (2)$$

From (1), we get

$$\begin{bmatrix} 8 & 0 & 0 \\ 4 & \frac{25}{2} & 0 \\ 6 & \frac{21}{4} & \frac{567}{50} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 33 \\ 36 \end{bmatrix}$$

so that

$$8y_1 = 20 \Rightarrow y_1 = \frac{5}{2}$$

$$4y_1 + \frac{25}{2}.y_2 = 33$$

$$\therefore y_2 = \left(33 - 4.\frac{5}{2}\right)\frac{2}{25} = \frac{46}{25}$$

$$6y_1 + \frac{21}{4}y_2 + \frac{567}{50}y_3 = 36$$

$$\Rightarrow y_3 = \frac{50}{567}\left[36 - 15 - \frac{21 \times 23}{50}\right] = 1.$$

Scanned by CamScanner

Then from (2), we have

$$\begin{bmatrix} 1 & -\frac{3}{8} & \frac{1}{4} \\ 0 & 1 & \frac{-4}{25} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ \frac{46}{25} \\ 1 \end{bmatrix}$$

which gives

$x_3 = 1$.

$x_2 - \frac{4}{25}x_3 = \frac{46}{25}$

$\Rightarrow x_2 = \frac{46}{25} + \frac{4}{25} = 2$

$x_1 - \frac{3}{8}x_2 + \frac{x_3}{4} = \frac{5}{2}$

$\Rightarrow x_1 = \frac{5}{2} + \frac{3}{8}.2 - \frac{1}{4} = 3$

Hence, the required solution is

$x_1 = 3,\ x_2 = 2,\ x_3 = 1$.

5.5. Gauss-Seidel iteration method.

[*W.B.U.T.,CS-312 2002*]

This method is an improvement of the Gauss-Jacobi method in the sense that the improved values of are used here in each iteration instead of the values of the previous iteration and hence the method is also known as the *method of successive displacements.*

To illustrate the method, we rewrite the system of equations (1) in the following form

$$x_1 = (b_1 - a_{12}x_2 - a_{13}x_3 - \ldots - a_{1n}x_n)/a_{11}$$

$$x_2 = (b_2 - a_{21}x_1 - a_{23}x_3 - \ldots - a_{2n}x_n)/a_{22} \qquad \ldots \quad (26)$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$x_n = (b_n - a_{n1}x_1 - a_{n2}x_2 - \ldots - a_{nn-1}x_{n-1})/a_{nn},$$

provided $a_{ii} \neq 0,\ i = 1, 2, \ldots, n$

To solve the equations (26), suppose $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$ be the initial approximations (usually $x_i^{(0)}, i = 1$ to n are taken to be zero) of the solutions of (1). We substitute these initial values on the right hand side of the first equation of (26) and get the first approximation of x_1 as

$$x_1^{(1)} = \left(b_1 - a_{12}x_2^{(0)} - a_{13}x_3^{(0)} - \ldots - a_{1n}x_n^{(0)}\right)/a_{11}$$

In the second equation of (26), we substitute the improved value $x_1^{(1)}$ and initial values $x_3^{(0)}, x_4^{(0)}, \ldots, x_n^{(0)}$ and obtain the first approximation of x_2 as

$$x_2^{(1)} = \left(b_2 - a_{21}x_1^{(1)} - a_{23}x_3^{(0)} - \ldots - a_{2n}x_n^{(0)}\right)/a_{22}$$

We then substitute in the third equation of (26) the improved values $x_1^{(1)}, x_2^{(1)}$ and the initial values $x_4^{(0)}, x_5^{(0)}, \ldots, x_n^{(0)}$ to obtain the first approximation of x_3 as

$$x_3^{(1)} = \left(b_3 - a_{31}x_1^{(1)} - a_{32}x_2^{(1)} - a_{34}x_4^{(0)} - \ldots - a_{3n}x_n^{(0)}\right)/a_{33}$$

Proceeding in this way, the first approximation of x_n is given by

$$x_n^{(1)} = \left(b_n - a_{n1}x_1^{(1)} - a_{n2}x_2^{(1)} - \ldots - a_{nn-1}x_{n-1}^{(1)}\right)/a_{nn}$$

Thus at the end of the first stage of iteration, we get the first approximation $x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)}$ to the solutions $x_1, x_2, \ldots, x_n$.

Now if $x_i^{(k)} (k = 0, 1, 2, \ldots)$ be the k^{th} approximation to the solutions $x_i (i = 1, 2, \ldots, n)$, then the $(k+1)^{th}$ the approximation $x_i^{(k+1)}$ of $x_i (i = 1, 2, \ldots, n)$, are given by

$$x_1^{(k+1)} = \left(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \ldots - a_{1n}^{(k)}x_n^{(k)}\right)/a_{11}$$

$$x_2^{(k+1)} = \left(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \ldots - a_{2n}^{(k)}x_n^{(k)}\right)/a_{22} \qquad \ldots \quad (27)$$

$$\ldots \qquad \ldots \qquad \ldots$$

$$x_n^{(k+1)} = \left(b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \ldots - a_{nn-1}x_{n-1}^{(k+1)}\right)/a_{nn}$$

Scanned by CamScanner

The process is continued until we get the solutions $x_1, x_2, \ldots, x_n$ with sufficient degree of accuracy

The sequence $\left\{x_i^{(k)}\right\}$ generated from (27) can be shown to be convergent to the solution $\left\{x_i^*\right\}$ if

$$|a_{ii}| > \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|, \ (i = 1, 2, \ldots, n) \qquad \ldots \quad (28)$$

Hence the Gauss-Seidel iteration method is convergent if the system of equations (1) is *strictly diagonally dominant.*

Note. **(1)** It may be noted that the strictly diagonally dominant condition may not be necessary in some problems for the convergence of iteration.

(2) The order of convergence of iteration in Gauss-Seidel method is one.

(3) The rate of convergence is faster (roughly twice) than that of Gauss-Jacobi method.

Example.5. Using Gauss-Seidel method find the solution of the following system of linear equations correct upto 2 places of decimal:

$$3x + y + 5z = 13$$

$$5x - 2y + z = 4$$

$$x + 6y - 2z = -1$$

[*W.B.U.T., CS-312, 2004, M(CS)-301, 2015, M(CS)-401, 2013*]

Solution. **First we rearrange the given system of equations so that they are diagonally dominant as given below :**

$$5x - 2y + z = 4$$

$$x + 6y - 2z = -1$$

$$3x + y + 5z = 13$$

We rewrite the system in the form

$$x = (4 + 2y - z)/5 \qquad \ldots \quad (1)$$

$$y = (-1 - x + 2z)/6 \qquad \ldots \quad (2)$$

$$z = (13 - 3x - y)/5 \qquad \ldots \quad (3)$$

The initial approximations are chosen to be

$$x_1^{(0)} = 0, \ x_2^{(0)} = 0, \ x_3^{(0)} = 0$$

First iteration :

Putting $y^{(0)} = 0, \ z^{(0)} = 0$ in (1), we get $x^{(1)} = 0.8$

Putting $x^{(1)} = 0.8, \ z^{(0)} = 0$ in (2), we have $y^{(1)} = -0.3$

Putting $x^{(1)} = 0.8, \ y^{(1)} = -0.3$ in (3) yields $z^{(1)} = 2.18$.

Second iteration :

$$x^{(2)} = \{4 + 2 \times (-0.3) - 2.18\}/5 = 0.2441$$

$$y^{(2)} = \{-1 - 0.244 + 2 \times 2.18\}/6 = 0.5192$$

$$z^{(2)} = \{13 - 3 \times 0.244 - 0.519\}/6 = 2.3497$$

Proceeding as above, the successive iterations are obtained and are shown in the following table :

k	$x^{(k)}$	$y^{(k)}$	$z^{(k)}$
0	0	0	0
1	0.8	−0.3	2.18
2	0.2441	0.5192	2.3497
3	0.5377	0.5271	2.1720
4	0.5763	0.4615	2.1628
5	0.552	0.462	2.176
6	0.550	0.467	2.177

Thus the required solutions are

$x = 0.55, \ y = 0.47, \ z = 2.18$, correct to two decimal places.

5.6. Computation of Inverse of matrix

Method I. To compute the inverse of a matrix $A = (a_{ij})_{n \times n}$, we determine a matrix $X = (x_{ij})_{n \times n}$ of the same order such that

$$AX = I \qquad \ldots \quad (29)$$

where I is the unit matrix of the same order.

Scanned by CamScanner

So for determination of each element of X, we solve a system of linear equations given by (29). This can be done by a systemetic procedure using Gauss elemination method. We illustrate the technique for a third order matrix

Let us consider the equation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is equivalent to the three system of linear equations given by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix}=\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix}=\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \ldots (30)$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then applying Gauss elimination method to each of these system, we get the corresponding column of X, i.e., the inverse of the matrix A^{-1}. But the coefficient matrix of each system of equations are same and so we can solve the three system of equations simultaneously considering the following augmented matrix :

$$\left[\begin{array}{ccc:ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array}\right]$$

Then employing the same procedure as in Gauss elemination we can easily solve the three set of the system of equations.

Example.6. Find the inverse of the matrix

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

Solution. Consider the augmented matrix

$$\left[\begin{array}{ccc:ccc} 2 & -2 & 4 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ -1 & 4 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\sim\left[\begin{array}{ccc:ccc} 2 & -2 & 4 & 1 & 0 & 0 \\ 0 & 5 & -2 & -1 & 1 & 0 \\ 0 & 3 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right], \text{(using } R_2-R_1 \text{ and } R_3+\tfrac{1}{2}R_1)$$

$$\sim\left[\begin{array}{ccc:ccc} 2 & -2 & 4 & 1 & 0 & 0 \\ 0 & 5 & -2 & -1 & 1 & 0 \\ 0 & 0 & \frac{11}{5} & \frac{11}{10} & \frac{-3}{5} & 1 \end{array}\right], \text{(using } R_3-\frac{3}{5}R_1)$$

Thus we have an equivalent system of three equations given by

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 5 & -2 \\ 0 & 0 & \frac{11}{5} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 1 \\ -1 \\ \frac{11}{10} \end{bmatrix} \quad \ldots (1)$$

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 5 & -2 \\ 0 & 0 & \frac{11}{5} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 0 \\ 1 \\ -\frac{3}{5} \end{bmatrix} \quad \ldots (2)$$

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 5 & -2 \\ 0 & 0 & \frac{11}{5} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \ldots (3)$$

Scanned by CamScanner

So for determination of each element of X, we solve a system of linear equations given by (29). This can be done by systemetic procedure using Gauss elemination method. We illustrate the technique for a third order matrix

Let us consider the equation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

which is equivalent to the three system of linear equations given by

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad (30)$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_{13} \\ x_{23} \\ x_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then applying Gauss elimination method to each of these system, we get the corresponding column of X, i.e., the inverse of the matrix A^{-1}. But the coefficient matrix of each system of equations are same and so we can solve the three system of equations simultaneously considering the following augmented matrix :

$$\left[\begin{array}{ccc:ccc} a_{11} & a_{12} & a_{13} & 1 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & 0 & 1 & 0 \\ a_{31} & a_{32} & a_{33} & 0 & 0 & 1 \end{array}\right]$$

Then employing the same procedure as in Gauss elemination, we can easily solve the three set of the system of equations.

Example.6. Find the inverse of the matrix

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

Solution. Consider the augmented matrix

$$\left[\begin{array}{ccc:ccc} 2 & -2 & 4 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ -1 & 4 & -1 & 0 & 0 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc:ccc} 2 & -2 & 4 & 1 & 0 & 0 \\ 0 & 5 & -2 & -1 & 1 & 0 \\ 0 & 3 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right], \text{(using } R_2 - R_1 \text{ and } R_3 + \tfrac{1}{2}R_1\text{)}$$

$$\sim \left[\begin{array}{ccc:ccc} 2 & -2 & 4 & 1 & 0 & 0 \\ 0 & 5 & -2 & -1 & 1 & 0 \\ 0 & 0 & \frac{11}{5} & \frac{11}{10} & \frac{-3}{5} & 1 \end{array}\right], \text{(using } R_3 - \tfrac{3}{5}R_1\text{)}$$

Thus we have an equivalent system of three equations given by

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 5 & -2 \\ 0 & 0 & \frac{11}{5} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ \frac{11}{10} \end{bmatrix} \qquad (1)$$

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 5 & -2 \\ 0 & 0 & \frac{11}{5} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -\frac{3}{5} \end{bmatrix} \qquad (2)$$

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 5 & -2 \\ 0 & 0 & \frac{11}{5} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad (3)$$

Equation (1) is equivalent to the following system of equations:

$$2x-2y+4z=1$$
$$5y-2z=-1$$
$$\frac{11}{5}z=\frac{11}{10}$$

Solving by back substitutions, we get

$$x=-\frac{1}{2},\ y=0,\ z=\frac{1}{2}$$

Similarly solving (2) and (3) we get

$$x=\frac{7}{11},\ y=\frac{1}{11},\ z=\frac{-3}{11}$$

and $x=\frac{-8}{11},\ y=\frac{2}{11},\ z=\frac{5}{11}$

$$\text{Thus } A^{-1}=\begin{bmatrix}-\frac{1}{2} & \frac{7}{11} & -\frac{8}{11}\\ 0 & \frac{1}{11} & \frac{2}{11}\\ \frac{1}{2} & -\frac{3}{11} & \frac{5}{11}\end{bmatrix}$$

Method II. This method is very similar to method I to compute the inverse matrix A^{-1} of the matrix A. Here also we consider the given matrix A with the same order identity matrix simultaneously and convert the matrix A into an identity matrix. As a result, the identity matrix is converted into a matrix which is the inverse of A.

Example.7. Find the inverse of the matrix

$$A=\begin{bmatrix}8 & -4 & 0\\ -4 & 8 & -4\\ 0 & -4 & 8\end{bmatrix}$$ [W.B.U.T., C.S-312, 2007, 2008]

Solution. Consider the augmented matrix given by

$$\left[\begin{array}{ccc|ccc}8 & -4 & 0 & 1 & 0 & 0\\ -4 & 8 & -4 & 0 & 1 & 0\\ 0 & -4 & 8 & 0 & 0 & 1\end{array}\right]$$

$$\sim\left[\begin{array}{ccc|ccc}1 & -\frac{1}{2} & 0 & \frac{1}{8} & 0 & 0\\ -4 & 8 & -4 & 0 & 1 & 0\\ 0 & -4 & 8 & 0 & 0 & 1\end{array}\right], \text{ using } R_1\to\tfrac{1}{8}R_1$$

$$\sim\left[\begin{array}{ccc|ccc}1 & -\frac{1}{2} & 0 & \frac{1}{8} & 0 & 0\\ 0 & 6 & -4 & \frac{1}{2} & 1 & 0\\ 0 & -4 & 8 & 0 & 0 & 1\end{array}\right], \text{ using } R_2\to R_2+4R_1$$

$$\sim\left[\begin{array}{ccc|ccc}1 & \frac{-1}{2} & 0 & \frac{1}{8} & 0 & 0\\ 0 & 1 & \frac{-2}{3} & \frac{1}{12} & \frac{1}{6} & 0\\ 0 & -4 & 8 & 0 & 0 & 1\end{array}\right], \text{ using } R_2\to\tfrac{1}{6}R_2$$

$$\sim\left[\begin{array}{ccc|ccc}1 & 0 & -\frac{1}{3} & \frac{1}{8} & \frac{1}{12} & 0\\ 0 & 1 & -\frac{2}{3} & \frac{1}{12} & \frac{1}{6} & 0\\ 0 & 0 & \frac{16}{3} & \frac{1}{3} & \frac{2}{3} & 1\end{array}\right], \text{ using } R_1\to R_1+\tfrac{1}{2}R_2$$
$$R_3\to R_3+4R_2$$

$$\sim\left[\begin{array}{ccc|ccc}1 & 0 & -\frac{1}{3} & \frac{1}{8} & \frac{1}{12} & 0\\ 0 & 1 & -\frac{2}{3} & \frac{1}{12} & \frac{1}{6} & 0\\ 0 & 0 & 1 & \frac{1}{16} & \frac{1}{8} & \frac{3}{16}\end{array}\right], \text{ using } R_3\to\tfrac{3}{16}R_3$$

$$\sim\left[\begin{array}{ccc|ccc}1 & 0 & 0 & \frac{3}{16} & \frac{1}{8} & \frac{1}{16}\\ 0 & 1 & 0 & \frac{1}{8} & \frac{1}{4} & \frac{1}{8}\\ 0 & 0 & 1 & \frac{1}{16} & \frac{1}{8} & \frac{3}{16}\end{array}\right], \text{ using } R_1\to R_1+\tfrac{1}{3}R_3$$
$$R_2\to R_2+\tfrac{2}{3}R_3$$

Hence the required inverse matrix is

$$A^{-1}=\begin{bmatrix}\frac{3}{16} & \frac{1}{8} & \frac{1}{16}\\ \frac{1}{8} & \frac{1}{4} & \frac{1}{8}\\ \frac{1}{16} & \frac{1}{8} & \frac{3}{16}\end{bmatrix}=\frac{1}{16}\begin{bmatrix}3 & 2 & 1\\ 2 & 4 & 2\\ 1 & 2 & 3\end{bmatrix}$$

Scanned by CamScanner

6.1. Introduction.

In applied mathematics and engineering we frequently face the problem of finding one or more roots of the equation

$$f(x) = 0 \quad \cdots \quad (1)$$

where $f(x)$ is, in general, a nonlinear function of the real variable x. But in most cases, it is very difficult to have explicit solutions of the equation (1) and, therefore, we proceed to look for a root of (1) numercially with any specified degree of accuracy. The numerical methods of finding these roots are called *iterative methods.*

The function $f(x)$ may have any one of the following forms:

(i) $f(x)$ is an algebraic or polynomial function of degree n, say, so that

$$f(x) = a_\circ + a_1x + a_2x^2 + \cdots + a_{n-1}x^{x-1} + a_nx^n$$

where $a_i (i = 0,1,2,\cdots,n)$ are constants, real or complex, and $a_n \neq 0$. For example, $x^3 - 7x + 1$, $x^{12} + x^5 - 4x + 3$ etc. are algebraic functions. In such cases, the equation $f(x) = 0$ is called *algebraic equation.*

(ii) $f(x)$ is a transcendental function, i.e. $f(x)$ is a function of the form

$$f(x) = a_\circ + a_1x + a_2x^2 + \cdots \text{ to } \infty,$$

where $a_i, (i = 0,1,2,\cdots)$, are constants, real or complex, and all $a_i \neq 0$. For example, $\sin x + 2x^2 - 1$, $e^x + \log x + 5$ etc. are transcendental functions, Here the equation $f(x) = 0$ is called *transcendental equation.*

Every value α of x for which the function $f(x)$ is zero, i.e., $f(\alpha) = 0$ is called a root or zero of the equation (1). In this chapter we shall discuss different numerical methods to compute the approximate real roots of an algebraic or transcendental equation $f(x) = 0$

Scanned by CamScanner

To develop the methods we assume that

(i) The function $f(x)$ is continuous and continuously differentiable for a sufficient number of times.

(ii) $f(x)$ has no multiple root, i.e., if α is a real root of $f(x)=0$ then

$$f(\alpha)=0,\ f'(\alpha)\neq 0.$$

Determination of approximate (real) root of (1) by numerical methods to be discussed here, consists, in general, of the following two steps.

(i) Isolating the roots, i.e., finding the smallest possible interval $[a, b]$ containing one and only one root of (1).

(ii) Improving the values of the approximate roots, i.e refinding them to the desired degree of accuracy.

To implement the first step, we use the following theorem of a continuous function :

Theorem 1. If real valued function $f(x)$ is continuous in $[a, b]$ and $f(a), f(b)$ are of opposite signs, then there is at least one real root of $f(x)=0$ in (a, b)

6.2. Iteration Processes.

Let the sequence $\{x_n\}$ of iterates of a root α of the equation

$$f(x)=0$$

is produced by a given method. Then the error ε_n involved at the n th iteration is given by

$$\varepsilon_n=\alpha-x_n. \qquad \dots \qquad (2)$$

If $\varepsilon_n\to 0$ as $n\to\infty$, we say that the iteration converges and the sequence $\{x_n\}$ converges to α. Otherwise, the iteration is divergent and the method of computation fails. Thus our primary task is to find the *condition of convergence of the iteration processes.* The error ε_{n+1} can be expressed in terms of ε_n, $\varepsilon_{n-1}, \varepsilon_{n-2}, \cdots$ which we call *error equation.*

If we define h_n by $h_n=x_{n+1}-x_n=\varepsilon_n-\varepsilon_{n+1}$, then h_n is an approximation of ε_n if x_{n+1} approximates α. If the iteration converges, then we can find an upper bound for $|\varepsilon_{n+1}|$ in terms of h_n. This is called *estimation of error.*

In case an iterative method converges, we can find two constants $p\geq 1$ and $q>0$ such that

$$\lim_{n\to\infty}\left|\frac{\varepsilon_{n+1}}{\varepsilon_n^p}\right|=q \qquad \dots \qquad (3)$$

Here p is called the *order of convergence* and q is known as *asymptotic error constant.* The iterative method with $p>1$ generally converges rapidly.

The convergence of the iterative method also depends on the initial approximation x_0 of the root α. If this initial approximation is not satisfactory, the iterative method does not converge and then we look for the new computation.

The iterative method is self correct, i.e. if there is an accidental error in the calculations of iteration, the erroneous iterate acts as a new initial approximation leading to a correct result, provided that the error is not large enough for which the method fails.

6.3. Bisection Method.

A. *Basic principle and formula*

The method of bisection is the most simplest iterative method. It is also known as *half-interval* or *Bolzano method.* This method is based on Theorem 1 on the change of sign.

In this method, we first find out a sufficiently small interval $[a_0,b_0]$ containing the required root α of the equation (1). Then $f(a_0)f(b_0)<0$ and $f'(x)$ has the same sign in $[a_0,b_0]$ and so $f(x)$ is strictly monotonic in $[a_0,b_0]$

To generate the sequence $\{x_n\}$ of iterates, we put $x_0=a_0$ or b_0 and $x_1=\frac{1}{2}(a_0+b_0)$ and find $f(x_1)$. If $f(a_0)$ and $f(x_1)$ are of opposite signs, then set $a_1=a_0, b_1=x_1$ so that $[a_1,b_1]=[a_0,x_1]$. On the other hand, if $f(x_1)$ and $f(b_0)$ are of opposite signs then put $a_1=x_1, b_1=b_0$, i.e. $[a_1,b_1]=[x_1,b_0]$. Thus we see that $[a_1,b_1]$ contains the root α in either case.

Scanned by CamScanner

Next set $x_2 = \frac{1}{2}(a_1 + b_1)$ and repeat the above process till we obtain

$$x_{n+1} = \frac{1}{2}(a_n + b_n). \qquad \cdots \quad (4)$$

with desired accuracy with $x_n \to \alpha$ as $n \to \infty$.

B. Convergence of bisection method

Suppose the interval $[a_n, b_n]$ contains the root α and $f(a_n)f(b_n) < 0$.

Let $x_{n+1} = \frac{1}{2}(a_n + b_n)$.

If $f(a_n)f(x_{n+1}) < 0$, then set $a_{n+1} = a_n$, $b_{n+1} = x_{n+1}$.

On the other hand, if $f(x_{n+1})f(b_n) < 0$, then

$$a_{n+1} = x_{n+1}, b_{n+1} = b_n.$$

Thus in any case

$$\alpha \in [a_{n+1}, b_{n+1}], \; f(a_{n+1})f(b_{n+1}) < 0$$

and

$$b_{n+1} - a_{n+1} < \frac{1}{2}(b_n - a_n) < \cdots < \frac{b_0 - a_0}{2^n}.$$

If ε_{n+1} be the error in approximating α by x_{n+1}, then

$$\varepsilon_{n+1} = |\alpha - x_{n+1}| \le b_n - a_n < \frac{b_0 - a_0}{2^n} \to 0 \text{ as } n \to \infty.$$

Thus the iteration converges.

Since $\frac{\varepsilon_{n+1}}{\varepsilon_n} = \frac{1}{2}$, so the convergence in bisection method is linear.

C. Geometrical Interpretation of Bisection method

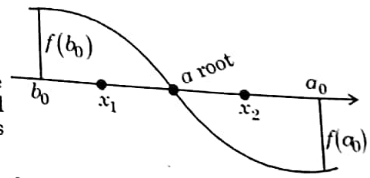
[M.A.K.A.U.T., MCS-401, 2014]

Let $f(x) = 0$ be the equation and $f(a_0)\, f(b_0) < 0$. Then a root lies between a_0 and b_0 as shown in the figure. Taking $x_1 = \frac{a_0 + b_0}{2}$, we divide the interval $[a_0, b_0]$ into two halves to locate a new sub-interval containing the root.

Let $f(x_1)\, f(a_0) < 0$.

$\therefore x_2 = \frac{x_1 + a_0}{2}$

This process can be repeated until the interval containing the root in as small as we desire.



D. Advantage and disadvantage of bisection method

Advantage. This method is very simple, as at any stage of iteration the approximate value of the desired root of the equation $f(x) = 0$ does not depend on the values $f(x_n)$ but on their signs only. Also the method is unconditionally and surely convergent.

Disadvantage. The method is very slow and requires large number of iteration to obtain moderately accurate results and hence it is laborious.

Example.1. Find the root of the equation $x \tan x = 1.28$, that lies in the interval $(0, 1)$, correct to four places of decimal, using bisection method. [W.B.U.T., M(CS)-312, 2005, M(CS)-301, 2015, M(CS)-401, 2015]

Solution. Let $f(x) = x \tan x - 1{\cdot}28$.

$\therefore f(0) = -1{\cdot}28 < 0$, $f(1) = 0{\cdot}277408 > 0$.

So a root lies between 0 and 1.

Take $a_0 = 0$, $b_0 = 1$ so that $x_1 = \frac{1}{2}(0 + 1) = 0.5$. Since $f(0.5) = -1.006849 < 0$ and $f(1) > 0$, the root lies between 0.5 and 1. Thus we have $x_2 = \frac{1}{2}(0.5 + 1) = 0.75$.

Proceeding in this way, we obtain the following table :

No. of iteration (n)	a_n $f(a_n) < 0$	b_n $f(b_n) > 0$	$x_{n+1} = \frac{a_n + b_n}{2}$	$f(x_{n+1})$
0	0	1	0.5	−1.006849
1	0.5	1	0.75	−0.581303
2	0.75	1	0.875	−0.232256
3	0.875	1	0.9375	−0.003058
4	0.9375	1	0.968750	0.129819

Scanned by CamScanner

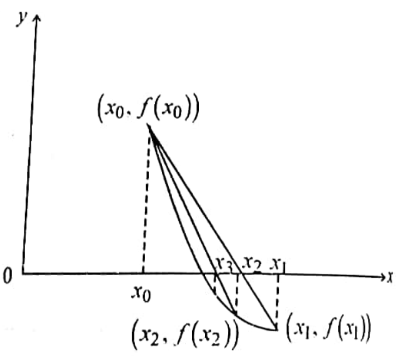
5	0.9375	0.96875	0.953125	0.061675
6	0.9375	0.961675	0.945312	0.028898
7	0.9375	0.945312	0.941406	0.012819
8	0.9375	0.941406	0.939453	0.004856
9	0.9375	0.939453	0.938477	0.000893
10	0.9375	0.938477	0.937988	−0.001084
11	0.937988	0.938477	0.938232	−0.000096
12	0.938232	0.938477	0.938354	0.000398
13	0.938232	0.938354	0.938293	0.000151
14	0.938232	0.938293	0.938263	0.000028

Thus a real root of the given equation is 0.9383 correct to four decimal places.

6.4 Regula-Falsi Method.

A. Basic principle and formula [W.B.U.T., CS-312, 2003]

The regula-falsi method or false position method is also sometimes referred to as the method of *linear interpolation* and it is the oldest method for computing real roots of an equation $f(x)=0$.



To find a real root α of $f(x)=0$, we first choose a sufficiently small interval $[x_0, x_1]$ in which the root α lies. Then $f(x_0)$ and $f(x_1)$ must be of opposite signs so that $f(x_0)f(x_1)<0$ and the graph of $f(x)$ must cross the x-axis between $x=x_0$ and $x=x_1$. Since the interval $[x_0, x_1]$ is sufficiently small, the portion of the curve between $A[x_0, f(x_0)]$ and $B[x_1, f(x_1)]$ can be approximated by a secant line (straight line) and so the intersection of the secant AB with the x-axis gives an approximate value x_2, say, of the root.

The equation of the secant line AB is

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

Putting $y=0, x=x_2$, we derive

$$x_2 = x_1 - \frac{f(x_1)}{f(x_1)-f(x_0)}(x_1-x_0) \qquad \dots (5)$$

If $f(x_2)=0$, then x_2 is a root of $f(x)=0$; otherwise, $f(x_2)<0$ or $f(x_2)>0$ It $f(x_\circ)$ and $f(x_2)$ are of opposite signs the root lies between $x_\circ$ and x_2 and in this case we set $x_1 = x_\circ$ and $x_2 = x_1$. On the other hand if $f(x_1)$ and $f(x_2)$ are of opposite signs, the root lies between x_1 and x_2 and thus, in either case,

$$f(x_1)f(x_2)<0$$

Hence the next approximation of the root, say x_3 lies between x_1 and x_2 and get

$$x_3 = x_2 - \frac{f(x_2)}{f(x_2)-f(x_1)}(x_2-x_1) \qquad \dots (6)$$

The general formula based on the above process is

$$x_{n+1} = x_n - \frac{f(x_n)}{f(x_n)-f(x_{n-1})}(x_n-x_{n-1}), n=1,2,\cdots \qquad \dots (7)$$

This is *regula falsi iteration formula.* The process is repeated until the root is obtained to required degree of accuracy.

B. Convergence of regula falsi method.

Let α be a simple root of the e uation $f(x)=0$. Then putting $x_n = \alpha + \varepsilon_n$ in (7), we get

$$\varepsilon_{n+1} = \varepsilon_n - \frac{(\varepsilon_n - \varepsilon_{n-1})f(\alpha+\varepsilon_n)}{f(\alpha+\varepsilon_n)-f(\alpha+\varepsilon_{n-1})}$$

Scanned by CamScanner

$$= \varepsilon_n - \frac{(\varepsilon_n - \varepsilon_{n-1})\left[\varepsilon_n f'(\alpha) + \frac{1}{2}\varepsilon_n^2 f''(\alpha) + \cdots\right]}{\left[\varepsilon_n f'(\alpha) + \frac{1}{2}\varepsilon_n^2 f''(\alpha) + \cdots\right] - \left[\varepsilon_{n-1} f'(\alpha) + \frac{1}{2}\varepsilon_{n-1}^2 f''(\alpha) + \cdots\right]}$$

(Expanding f in Taylor's series and noting $f(\alpha) = 0$)

$$= \varepsilon_n - \frac{\varepsilon_n f'(\alpha) + \frac{1}{2}\varepsilon_n^2 f''(\alpha) + \cdots}{f'(\alpha) + \frac{1}{2}(\varepsilon_n + \varepsilon_{n-1}) f''(\alpha) + \cdots}$$

$$= \varepsilon_n - \varepsilon_n\left[1 + \frac{1}{2}\varepsilon_n \frac{f''(\alpha)}{f'(\alpha)} + \cdots\right] \times \left[1 + \frac{1}{2}(\varepsilon_n + \varepsilon_{n-1})\frac{f''(\alpha)}{f'(\alpha)} + \cdots\right]^{-1}$$

$$= \frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)}\varepsilon_{n-1}\varepsilon_n + O\left(\varepsilon_n^2\varepsilon_{n-1} + \varepsilon_n\varepsilon_{n-1}^2\right), \quad \cdots (8)$$

so that

$$\varepsilon_{n+1} = C\,\varepsilon_{n-1}\varepsilon_n \quad \cdots (9)$$

where $C = \frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)}$ and we have neglect higher power of ε_n.

The relation (9) is called the error equation.

To find the order of convergence, we set $\varepsilon_{n+1} = A\varepsilon_n^m$ and $\varepsilon_n = A\varepsilon_{n-1}^m$ where the constants A and m are to be determined. Then the equation (9) gives

$$A\varepsilon_n^m = C \cdot A^{-\frac{1}{m}}\varepsilon_n^{\frac{1}{m}}\varepsilon_n = CA^{-\frac{1}{m}}\varepsilon_n^{1+\frac{1}{m}} \quad \cdots (10)$$

$$\Rightarrow m = 1 + \frac{1}{m}$$

$$\Rightarrow m = \frac{1}{2}\left(1 \pm \sqrt{5}\right)$$

Neglecting the minus sign, we find that the order of the convergence of $\{x_n\}$ is $m = 1 \cdot 618$. Also from (10) we get

$$A = C^{m/m+1}$$

Q. Advantage and disadvantage of regula falsi method. [W.B.U.T., CS-312,2003]

Advantage. The method is very simple and does no required to calculate the derivative of $f(x)$ which is difficult for some problems. Moreover, the method is evidently convergent.

Disadvantage. Sometimes the method is very slow and not suitable for practical computation. Also the initial interval in which the root lies is to be chosen very small.

Example.2. Find a root of the equation $x^3 - 2x - 5 = 0$ by Regula-Falsi method correct upto 4 places of decimal [W.B.U.T. CS-312,2004]

Solultion. Let $f(x) = x^3 - 2x - 5$.

$\therefore f(0) = -5,\ f(1) = -6,\ f(2) = -1,\ f(3) = 16$. So a real root lies between 2 and 3. We choose $x_0 = 2, x_1 = 3$ giving $f(x_0) = -1$ and $f(x_1) = 16$. Then the iteration (7) gives

$$x_2 = 3 - \frac{16}{16-(-1)}(3-2) = 2 \cdot 05882 \text{ and } f(x_2) = -0 \cdot 39082$$

Proceeding in this way, the iteration (7) gives the following table :

No, of iteration (n)	x_{n-1} $(f(x_{n-1})<0)$	x_n $(f(x_n)<0)$	$f(x_{n-1})$	$f(x_n)$	x_{n+1}	$f(x_{n+1})$
1	2	3	–1	16	2.05882	–0.39084
2	2.05882	3	–0.39084	16	2.08126	–0.147244
3	2.08126	3	–0.147244	16	2.08964	–0.054667
4	2.08964	3	–0.054667	16	2.09274	–0.020198
5	2.09274	3	–0.020198	16	2.09388	–0.007491
6	2.09388	3	–0.007491	16	2.09430	–0.002806
7	2.09430	3	–0.002806	16	2.09445	–0.001133
8	2.09445	3	–0.001133	16	2.094451	

Hence 2.0944 is a root of the given equation correct upto four decimal places.

Scanned by CamScanner

6.5. Newton-Raphson method.

A. Basic principle and formula [W.B.U.T., CS-312,2007]

Let x_0 be an initial approximation of the desired root α of the equation $f(x)=0$ and $x_1 = x_0 + h$ is the correct root

$$\therefore f(x_1) = 0$$

$$\text{i.e., } f(x_0+h) = 0, \left(\min(x_1, x_0) < h < \max(x_1, x_0)\right)$$

$$\text{i.e., } f(x_0) + hf'(x_0) + \frac{h^2}{2!}f''(x_0) + \cdots = 0,$$

by Taylor's series, expansion

Neglecting the second and higher order terms, we obtain

$$f(x_0) + hf'(x_0) = 0,$$

$$\text{i.e., } h = -\frac{f(x_0)}{f'(x_0)}$$

Thus a better approximation of the root α is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \quad \cdots (11)$$

Repeating the above process and replacing x_0 by x_1 we obtain the second approximation of the root as

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

Proceeding in this way, we get the successive approximations $x_3, x_4, \cdots, x_{n+1}$ where

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad \cdots (12)$$

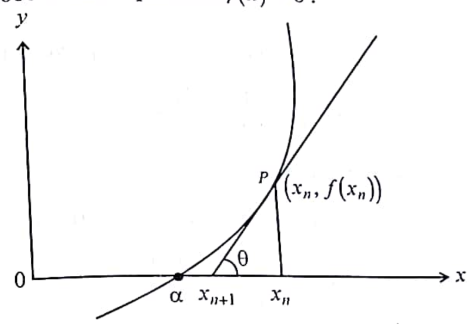
provided $f'(x_n) \neq 0, n = 1,2,\cdots$

The result (11) is known as *Newton-Raphson iteration formula.*

If $[a,b]$ be the initial interval in which the root α of the given equation $f(x)=0$ lies and $f'(x) \neq 0$, then the initial approximation may be started with $x_0 = a$ or b.

B. Geometrical meaning of Newton-Raphson formula

Let the curve $y = f(x)$ cuts the x–axis at the point $x = \alpha$ so that α is a root of the equation $f(x) = 0$.



If the tangent to the curve at the point $P(x_n, f(x_n))$ cuts the x–axis at the point $x = x_{n+1}$ and is inclined at an angle θ with the positive direction of the x–axis, then

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

so that

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

Accordingly, the Newton-Raphson method may also be called the *method of tangents.*

C. *Convergence of Newton–Raphson method.*

[W.B.U.T., CS-312,2003,2007, M(CS)-401,2013,2014,2015]

Let α be a root of the equation $f(x) = 0$

$$\therefore f(\alpha) = 0 \qquad \text{i.e., } f(x_n + \alpha - x_n) = 0$$

$$\text{i.e., } f(x_n) + (\alpha - x_n)f'(x_n) + \frac{(\alpha - x_n)^2}{2}f''(\xi_n) = 0$$

$\left[\min|\alpha, x_n| < \xi_n < \max|\alpha, x_n|\right]$, by Taylor's theorem

$$\therefore -\frac{f(x_n)}{f'(x_n)} = (\alpha - x_n) + \frac{1}{2}(\alpha - x_n)^2 \frac{f''(\xi_n)}{f'(x_n)} \quad \cdots (13)$$

Scanned by CamScanner

We have, therefore, from (12) and (13),

$$x_{n+1} - \alpha = \frac{1}{2}(\alpha - x_n)^2 \frac{f''(\xi_n)}{f'(x_n)}$$

Putting $\alpha - x_n = \varepsilon_n$ and $\alpha - x_{n+1} = \varepsilon_{n+1}$, we get

$$\varepsilon_{n+1} = -\frac{1}{2}\varepsilon_n^2 \frac{f''(\xi_n)}{f'(x_n)} \quad \cdots \quad (14)$$

which is the error equation.

If the iteration converges then $x_n, \xi_n \to \alpha$ as $n \to \infty$ so that

$$\lim_{n\to\infty}\left|\frac{\varepsilon_{n+1}}{\varepsilon_n^2}\right| = \frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)} \quad \cdots \quad (15)$$

Hence Newton–Raphson method is a second order iteration process. So the convergence is quadratic and the constant asymptotic, error is equal to $\frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)}$

D. Advantage and disadvantage of Newton–Raphson method. [M.A.K.A.U.T., MCS-401,2014]

Advantage. The rate of convergence of this method is quadratic. So the method converges more rapidly than other numerical method.

Disadvantage. In this method, the initial approximation must be chosen very close to the root; otherwise, the method will fail. Since the method depends on the derivative $f'(x)$, it may not be suitable for a function $f(x)$ whose derivative is difficult to compute. Also the method fails if $f'(x) = 0$ or small in the neighbourhood of the root.

Note. When $f'(x)$ is large in the neighbourhood of the real root, i.e., when the graph of function $y = f(x)$ is nearly vertical, then it crosses the x-axis. In this case, the method is very useful and the correct value of the root can be obtained more rapidly.

Example.3. Find the smallest positive root of the equation $3x^3 - 9x^2 + 8 = 0$ correct to four places decimal, using Newton–Raphson method. [W.B.U.T., CS-312,2009]

Solution. Let $f(x) = 3x^3 - 9x^2 + 8$

$\therefore f'(x) = 9x^2 - 18x$.

Then the iteration formula (12) gives

$$x_{n+1} = x_n - \frac{3x_n^3 - 9x_n^2 + 8}{9x_n^2 - 18x_n}$$

$$= \frac{6x_n^3 - 9x_n^2 - 8}{9x_n^2 - 18x_n}, n = 0, 1, 2, \cdots \quad \cdots \quad (1)$$

Now $f(0) = 8$, $f(1) = 2$, $f(2) = -4$. So a positive root lies between 1 and 2. Choose $x_\circ = 1$

$\therefore$ From (1)

$$x_1 = \frac{6x_\circ^3 - 9x_\circ^2 - 8}{9x_\circ^2 - 18x_\circ} = \frac{6\times 1^3 - 9\times 1^2 - 8}{9\times 1^2 - 18\times 1} = 1\cdot 22222$$

$$x_2 = \frac{6(1\cdot 22222)^3 - 9(1\cdot 22222)^2 - 8}{9(1\cdot 22222)^2 - 18\times 1\cdot 222222} = 1\cdot 22607$$

$$x_3 = \frac{6(1\cdot 22607)^3 - 9(1\cdot 22607)^2 - 8}{9(1\cdot 22607)^2 - 18\times 1\cdot 22607} = 1\cdot 22607$$

Hence positive real root of the given equation correct to four decimal places is 1.2261.

E. *Newton-Raphson method for finding an assigned root of a positive real number.* [W.B.U.T., CS-312,2009]

Suppose we are to find the m^{th} root of a real number R.

So let $x = \sqrt[m]{R}$

$\therefore x^m = R$

i.e., $x^m - R = 0$

Let $f(x) = x^m - R$

$\therefore f'(x) = mx^{m-1}$.

Scanned by CamScanner

The Newton–Raphson iteration formula (13) gives

$$x_{n+1} = x_n - \frac{x_n^m - R}{m x_n^{m-1}}$$

$$= \frac{(m-1)x_n^m + R}{m x_n^{m-1}}, n = 0, 1, 2, \ldots \qquad \ldots (16)$$

with $|x_{n+1} - x_n| < \varepsilon$, ε being the desired degree of accuracy.

It is obvious from (17) that for finding the square root of any positive number R (where $m = 2$), we have

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{R}{x_n}\right), \ n = 0, 1, 2, \ldots$$

Example.4. Find the cube root of 10 upto 5 significant figures by Newton–Rapshon method.

Solution. Let $x = \sqrt[3]{10}$

$\therefore x^3 - 10 = 0$

Let $f(x) = x^3 - 10$.

$\therefore f'(x) = 3x^2$

Now $f(0) = -10, f(1) = -9, f(2) = -2, f(3) = 17$ so that a real root of $f(x) = 0$ lies between 2 and 3. Hence using the iteration formula (16) with $m = 3$, we have

$$x_{n+1} = \frac{2x_n^3 + 10}{3x_n^2}, \ n = 0, 1, 2, \ldots$$

which gives with $x_0 = 2$

$$x_1 = \frac{2x_0^3 + 10}{3x_0^2} = 2.16666$$

Similarly, $x_2 = \dfrac{2(2.16666)^3 + 10}{3(2.16666)^2} = 2.15450$.

$x_3 = 2.15443$,

$x_4 = 2.15443$

Hence we have $x = \sqrt[3]{10} \simeq 2.1544$ correct to five significant figures.

Example.5. Use $N - R$ method find the cube root of 17.

[W.B.U.T., MCS-301,2009]

Solution. Let $x = \sqrt[3]{17}$

$\therefore x^3 - 17 = 0$

Let $f(x) = x^3 - 17$

$\therefore f'(x) = 3x^2$

$f(0) = -17, f(1) = -16, f(2) = -9, \ f(3) = 10$

So a root lies between 2 and 3.

We choose $x_0 = 2$

From, $N - R$ iterative formula,

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

we get

$$x_{n+1} = x_n - \frac{x_n^3 - 17}{3x_n^2} = \frac{2x_n^3 + 17}{3x_n^2}$$

$$\therefore \ x_1 = \frac{2x_0^3 + 17}{3x_0^2} = \frac{2 \times 2^3 + 17}{3 \times 2^2} = 2.75$$

$$\therefore \ x_2 = \frac{2x_1^3 + 17}{3x_1^2} = \frac{2(2.75)^3 + 17}{3(2.75)^2} = 2.5826$$

$$x_3 = \frac{2(2.5826)^3 + 17}{3(2.5826)^2} = 2.5713$$

$$x_4 = \frac{2(2.5713)^3 + 17}{3(2.5713)^2} = 2.5713$$

So the root of $x^3 - 17 = 0$ is 2.5713

Hence the cube root of 17 is 2.5713

Scanned by CamScanner

ILLUSTRATIVE EXAMPLES

Ex.1. Find a real root of the transcendental equation $x^x + 2x - 2 = 0$, correct upto two decimal places using bisection method.

Solution. Let $f(x) = x^x + 2x - 2$.

Then $f(0\cdot5) = -0\cdot293, f(1) = 1$.

So a real root lies between 0.5 and 1.

Take $a_0 = 0\cdot5, b_0 = 1$ so that $x_1 = \frac{0\cdot5+1}{2} = 0\cdot75$.

Again, since $f(0\cdot75) = 0\cdot3059$, so the root lies between 0.75 and 1. Proceeding in this way, we construct the following table:

No. of iteration(n)	a_n $(f(a_n)<0)$	b_n $(f(b_n)>0)$	$x_{n+1} = \frac{1}{2}(a_n + b_n)$	$f(x_{n+1})$
0	0.5	1	0.75	0.3059
1	0.5	0.75	0.625	−0.0045
2	0.625	0.75	0.687	0.1466
3	0.625	0.687	0.656	0.0704
4	0.625	0.656	0.640	0.0325
5	0.625	0.640	0.632	0.0041
6	0.625	0.632	0.628	0.0036
7	0.625	0.628	0.626	

Then the root correct upto two decimal places is 0.63.

Ex.2. Find the smallest positive root of the equation $e^x = 4\sin x$ correct to four decimal places by bisection method.

Solution. The given equation is

$$4\sin x - e^x = 0.$$

Let $f(x) = 4\sin x - e^x$.

Since $f(0) = -1$, $f(1) = 0.64760$, so the smallest positive root lies between 0 and 1. Take $a_0 = 0$, $b_0 = 1$ so that

$$x_1 = \frac{1}{2}(a_0 + b_0) = 0.5.$$

Since $f(0.5) = 0.26898$, so the root lies between 0 and 0.5. Proceeding in this way, we construct the following table: (shown in the next page)

No. of iteration (n)	a_n $(f(a_n)<0)$	b_n $(f(b_n)>0)$	$x_{n+1} = \frac{a_n + b_n}{2}$	$f(x_{n+1})$
0	0	1	0.5	0.26898
1	0	0.5	0.25	−0.29440
2	0.25	0.5	0.375	0.01010
3	0.25	0.375	0.33333	−0.08690
4	0.33333	0.375	0.35416	−0.03770
5	0.35416	0.375	0.36458	−0.01360
6	0.36458	0.375	0.36979	−0.00168
7	0.36979	0.375	0.37240	0.00420
8	0.36979	0.37240	0.37109	0.00126
9	0.36979	0.37109	0.37044	−0.00240
10	0.37044	0.37109	0.37077	0.00047
11	0.37044	0.37077	0.37060	0.00011
12	0.37044	0.37060	0.37052	−0.00007
13	0.37052	0.37060	0.37056	

Thus the required smallest positive root is 0.3706 correct to four decimal places.

Ex.3. Find the root of the equation $3x - \cos x - 1 = 0$ that lies between 0 and 1 correct to four places of decimal using bisection method.

[*W.B.U.T.,MCS-301, 2009*]

Solution. Let $f(x) = 3x - \cos x - 1$

$\therefore f(0) = -2, f(2) = 1.4597$

Scanned by CamScanner

So a root lies between 0 and 1

Take $a_0 = 0, b_0 = 1$

$\therefore x_1 = \frac{1}{2}(0+1) = 0.5$

Now $f(x_1) = -0.3776$

So the root lies between 0.5 and 1

Thus we have $x_2 = \frac{0.5+1}{2} = 0.75$.

Proceeding in this way, we obtain the following tables :

No of iteration	a_n $(f(a_n)<0)$	b_n $(f(b_n)>0)$	$x_{n+1} = \frac{a_n + b_n}{2}$	$f(x_{n+1})$
0	0	1	0.5	−0.3776
1	0.5	1	0.75	0.5183
2	0.5	0.75	0.625	0.0640
3	0.5	0.025	0.5625	−0.1584
4	0.5625	0.625	0.59375	−0.0476
5	0.59375	0.625	0.60938	0.0081
6	0.59375	0.60938	0.60157	−0.0198
7	0.60157	0.60938	0.60548	−0.0058
8	0.60548	0.60938	0.60743	−0.0012
9	0.60548	0.60743	0.60646	−0.0023
10	0.60646	0.60743	0.60695	−0.0005
11	0.60695	0.60743	0.60719	0.0003
12	0.60695	0.60719	0.60707	−0.0001
13	0.60707	0.60719	0.60713	0.0001

Thus the required root is 0.6071, correct upto four decimal places.

Ex.4. Using Bisection method obtain a root between 1 and 2 of the equation $e^x - 3x = 0$. *[M.A.K.A.U.T.,M(CS)-401, 2016]*

Solution. Let $f(x) = e^x - 3x$

$f(0) = 1 > 0$

$f(1) = -0.281718 < 0.$

So, a root lies between 0 & 1.

No of iteration	a_n $f(a_n)<0$	b_n $f(b_n)>0$	$x_{n+1} = \frac{a_n + b_n}{2}$	$f(x_{n+1})$
0	0	1	0.5	0.1487
1	0.5	1	0.75	−0.133
2	0.5	0.75	0.625	−0.0067
3	0.5	0.625	0.5625	0.068
4	0.5625	0.625	0.59375	0.03
5	0.59375	0.625	0.609375	0.011
6	0.609375	0.625	0.6171875	0.0021
7	0.6171875	0.625	0.621094	−0.00232
8	0.6171875	0.621094	0.619141	−0.00009
9	0.6171875	0.619141	0.618164	0.0010262
10	0.618164	0.619141	0.6186525	

$\therefore$ The root is 0.62 correctup to 2 decimal places.

Ex.5. Find a positive root of $x + \ln x - 2 = 0$ by Newton-Raphson method correct upto six significant figure. *[M.A.K.A.U.T.,M(CS)-401,2015,2016]*

Solution. Let $f(x) = x + \ln x - 2$

$f(1) = -1 < 0$

$f(2) = 0.693147 > 0$

So, a positive lies between 1 & 2.

We choose $x_0 = 1.5$

$f'(x) = 1 + 1/x$

Scanned by CamScanner

By Newton Raphson method

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

$$= x_n - \frac{(x_n + \ln x_n - 2)x_n}{1 + x_n}$$

$$= \frac{x_n^2 + x_n - x_n^2 - x_n \ln x_n + 2x_n}{1 + x_n}$$

$$= \frac{3x_n - x_n \ln x_n}{1 + x_n}$$

$$x_1 = \frac{3x_0 - x_0 \ln x_0}{1 + x_0} = 1.556721$$

$$x_2 = \frac{3x_1 - x_1 \ln x_1}{1 + x_1} = 1.556714$$

Similarly,

$x_3 = 1.557146$

$x_4 = 1.557146$

$\because$ The required root is 1.55715 correct up to 6 singnificant figure.

Ex.6. Find a root of the equation $x \sin x + \cos x = 0$ using Newton-Raphson method correct upto 5 places of decimal.

[*W.B.U.T., CS-312, 2004*]

Solution. Let $f(x) = x \sin x + \cos x$

$\therefore f'(x) = x \cos x + \sin x - \sin x = x \cos x$

So the Newton-Raphson iteration formula

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

gives

$$x_{n+1} = x_n - \frac{x_n \sin x_n + \cos x_n}{x_n \cos x_n}$$

$$= x_n - \frac{1}{x_n} - \tan x_n \quad \ldots (1)$$

Now $f(0) = 1, f(1) = 1.38, f(2) = 1.4, f(3) = -0.56$

So a root lies between 2 and 3.

Choose $x_0 = 2.5$, for quick convergence.

Then we have, from (1),

$$x_1 = 2.5 - \tfrac{1}{2.5} - \tan 2.5$$

$$= 2.847022$$

$$x_2 = 2.847022 - \frac{1}{2.847022} - \tan 2.847022$$

$$= 2.799175$$

Similarly $x_3 = 2.798386$

$x_4 = 2.798386$

Hence a positive real root of the given equation correct to five decimal places is 2.79839.

Ex.7. Find out the root of the following equation using Regula falsi method $x^3 - 5x - 7 = 0$ that lies between 2 and 3, correct to 4 decimal places. [*W.B.U.T.CS-312, 2006, 2009*]

Solution. Let $f(x) = x^3 - 5x - 7$

We choose $x_0 = 2, x_1 = 3$

so that $f(x_0) = -9, f(x_1) = 5$.

Then the regula-falsi iteration formula

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

gives

$$x_2 = 3 - \frac{5}{5 - (-9)}(3 - 2) = 2.642857$$

Scanned by CamScanner

Proceeding in this way, the iteration formula (1) gives the following table :

No of iteration (n)	x_{n-1} $(f(x_{n-1})<0)$	x_n $(f(x_n)<0)$	$f(x_n - 1)$	$f(x_n)$	x_{n+1}	$f(x_{n+1})$
1	2	3	–9	5	2.642857	–1.754740
2	2.642857	3	–1.754740	5	2.735635	–0.205506
3	2.735635	3	–0.205506	5	2.746072	–0.022474
4	2.746072	3	–0.022474	5	2.747208	–0.002444
5	2.747208	3	–0.002444	5	2.747332	–0.000257
6	2.747332	3	–0.000257	5	2.747345	–0.000027

Hence a real root of the given equation correct to four decimal places is 2.7473.

Ex.8. Find the root of the equation $xe^x = \cos x$ using the Regular-False method.

[*W.B.U.T. MCS-401,2006*]

Solution. The given equation can be written as

$$\cos x - xe^x = 0$$

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

$\therefore f(0) = 1, f(1) = -2.17798$

So a real root lies between 0 and 1

We choose $x_0 = 0, x_1 = 1$

$\therefore f(x_0) = 1, f(x_1) = -2.17798$

Then Regula-falsi iterative method

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

gives

$$x_2 = 1 - \frac{-2.17798}{-2.17798 - 1}(1 - 0)$$

$$= 0.31467$$

Now $f(x_2) = 0.51987$

So the root lies between 0.31467 and 1

$\therefore$ From (1),

$$x_3 = 1 - \frac{-2.17798}{-2.17718 - 0.51987}(1 - 0.31467)$$

$$= 0.44673$$

Again $f(x_3) = 0.20356$

So the root lies between 0.44673 and 1.

Then applying (1), we get

$$x_4 = 0.49402$$

Repeating the process, we can obtain

$$x_5 = 0.50995, x_6 = 0.51520, x_7 = 0.51692$$

$$x_8 = 0.51748, x_9 = 0.51767, x_{10} = 0.51775$$

Thus the required root is 0.5177, correct upto four decimal places.

Ex.9. Find the root of the equation $3x - \cos x - 1 = 0$ using Regula-Falsi method.

[*W.B.U.T. MCS-301, 2007*]

Solution. Let $f(x) = 3x - \cos x - 1$

$\therefore f(0) = -2, f(1) = 1.4597$

So a real root lies between 0 and 1

we choose $x_0 = 0, x_1 = 1$. Then the iteration formula

$$x_{n+1} = x_n - \frac{f(x_n)}{f(x_0) - f(x_{n-1})}(x_n - x_{n-1}) \quad n = 1, 2, \ldots\ldots \quad (1)$$

gives

$$x_2 = 1 - \frac{1.4597}{1.4597 - (-2)}(1 - 0)$$

$$= 0.578085$$

$\therefore f(x_2) = -0.103256$

So the root lies between 0.578085 and 1.

Scanned by CamScanner