CHAPTER 20

NUMERICAL COMPUTATION

A Bit of History

First of all we arise following questions

- Why are computational methods in mathematics important?
- What can we do with these methods?
- What is the difference between computation by hand and by computer? What do we need to know to perform computations on computers?

These are natural questions for a student to ask before starting a course on computational methods. And therefore it is also appropriate to try and provide some short answers already in this introduction. By the time you reach the end of the chapter you will hopefully have more substantial answers to these as well as many other questions

Characteristics of Numerical Computation

1. **Accuracy:** Every method of numerical computing introduces errors. They may be either due to using an appropriate method in place of an exact mathematical procedure or due to inexact representation and manipulation of numbers in the computer. These errors affect the accuracy of the results.
2. **Efficiency:** A more efficient numerical method is one which gives the more desired accuracy with less amount of work.
3. **Rate of Convergence:** In numerical analysis, the speed at which a convergent sequence approaches its limit is called the rate of convergence. Although strictly speaking, a limit does not give information about any finite first part of the sequence, this concept is of practical importance if we deal with a sequence of successive approximations for an iterative method, as then typically fewer iterations are needed to yield a useful approximation if the rate of convergence is higher.

   Suppose that the sequence $\{x_k\}$ converges to the number L.

   We say that this sequence converges to L with order q, if there exists a number $\mu > 0$ such that

$$\lim_{k \to \infty} \frac{|x_{k+1} - L|}{|x_k - L|^q} = \mu.$$

   The number μ is called the *rate of convergence*.

   If the sequence converges, and

   - $\mu = 0$, then the sequence is said to converge superlinearly.
   - $\mu = 1$, then the sequence is said to converge sublinearly.
   - $q = 1$, then the sequence is said to converge linearly.
   - $q = 2$, then the sequence is said to converge quadratically and so on.

Approximation of Errors and their Computations

Since numerical solutions are approximations, and since the computer program that executes the numerical method might have errors, a numerical solution needs to be examined closely. There are three major sources of error in computation: human errors, round-off errors and truncation errors.

1. **Human errors:** Typical human errors are arithmetic errors, and/or programming errors: These errors can be very hard to detect unless they give obviously incorrect solution. In discussing errors, we shall assume that human errors are not present.
2. **Truncation and Round-off errors:** Error in computation is the difference between the exact answer X_{ex} and the computed answer X_{cp}. This is also known as true error $\therefore$ true error $= X_{cp} - X_{ex}$

   Since we are usually interested in the magnitude or absolute value of the error we define true absolute error $= | X_{cp} - X_{ex} |$

   There are two kinds of numbers, exact and approximate numbers. Examples of exact numbers are 1, 2, 3 ...., 1/2, 3/2, ... $\sqrt{2}$, π, e, etc. written in this manner. Approximate numbers are those that represent the numbers to a certain degree of accuracy. Thus, an approximate value of π is 3.1416, or if we desire a better approximation, it is 3.14159265. But we cannot write that exact value of π.

   The digits that are used to express a number are called significant digits or significant figures. Thus, the numbers 3.1416, 0.66667 and 4.0687 contain five significant digits each. The number 0.00023 has, however, only two significant digits, viz., 2 and 3, since the zeros serve only to fix the position of the decimal point. Similarly, the numbers 0.00145, 0.000145 and 0.0000145 all have three significant digits. In case of ambiguity, the scientific noation should be used. For example, in the number 25,600, the number of significant figures is uncertain whereas the number 2.56×10^4, 2.560×10^4 and 2.5600×10^4 have three, four and five significant digits, respectively.

   In numberical compuations, we come across numbers which have large number of digits and it will be necessary to cut them to a usuable number of figures. This process is called rounding off. It is usual to round off numbers according to the following rule:

Truncation Error

Truncation error is defined as the error caused by truncating a mathematical procedure. This series has an infinite number of terms but when we use only finite number of terms we get truncation error. Numerical methods use approximations for solving problems. The errors introduced by the approximations are the truncation errors.

For example, consider the series expansion

$$e^x = 1 + x + \frac{x^2}{1.2} + \ldots + \frac{x^n}{1.2.3\ldots n} + \ldots$$

If the formula is used to calculate $e^{0.3}$ we get $e^{0.3} = 1 + 0.3 + \frac{(0.3)^2}{1.2} + \ldots + \frac{(0.3)^n}{1.2.3\ldots n} + \ldots$

Theoretically the calculation will never stop. There are always more terms to add on. If we do stop after a finite number of terms, we will not get the exact answer. For example if we do take only a finite number of terms then we get truncation error.

For example, if one uses three terms to calculate e^x,

If $r = e^x = 1 + x + \frac{x^2}{1.2} + \ldots + \frac{x^n}{1.2.3\ldots n} + \ldots$

and we take

$r_1 = e^x = 1 + \frac{x}{1} + \frac{x^2}{1.2}$ then truncarion error is

$$r - r_1 = \ldots + \frac{x^3}{1.2.3} + \ldots + \frac{x^n}{1.2.3\ldots n} + \ldots$$

Types of Errors

The practice of numerical analysis is well known for focusing on algorithms as they are used to solve issues in continuous mathematics. The practice is familiar not only in engineering and physical sciences, but also in astrology, stock portfolio analysis, data analysis, medicine etc. Part of the application of numerical analysis involves the use of errors. Specific errors are sought out and applied to arrive at mathematical conclusions.

True Error, Absolute Error, Relative Error and Percent Error

The absolute error is the magnitude of the difference between the exact value and the approximation. The relative error is the absolute error divided by the magnitude of the exact value. The percent error is the relative error expressed in terms of per 100.

Given some exact value v and its approximation v_{approx},
the absolute error is

$$\epsilon = | v - v_{\text{approx}} |$$

If $v \neq 0$ the relative error is

$$\frac{| v - v_{\text{approx}} |}{| v |}$$

and the percent error is $\frac{| v - v_{\text{approx}} |}{| v |} \times 100\%$.

Methods of Solving Non-linear Equations

A problem of great importance in science and engineering is that of determinig the roots/zeros of an equation of the form

$$f(x) = 0 \qquad \text{......(i)}$$

If f(x) is a non-linear function, then (i) is called non-linear equation.

1. A polynomial equation of the form

   $f(x) = P_n(x) = a_0 x^n + a_1x^{n-1} + a_2x^{n-2} + \ldots + a_{n-1}x + a_n = 0$ with $n \geq 2$ is an example of non-linear algebraic equation. In particular, $3x^3 - 2x^2 - x - 5 = 0$, $x^4 - 3x^2 + 1 = 0$, $x^2 - 3x + 1 = 0$, are non-linear algebraic (polynomial) equations.

2. An equation which contains exponential functions, logarithmic functions, trigonometric functions etc., called a transcendental equation, is also non-linear equation .

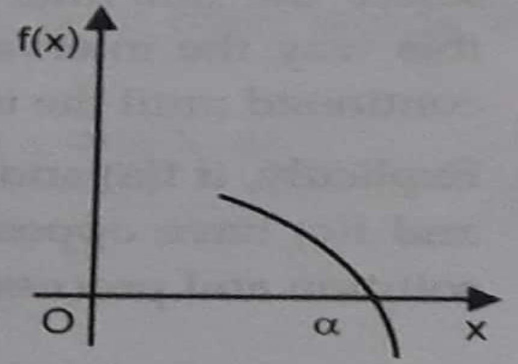
For example,

$xe^{2x} - 1 = 0$, $\cos x - xe^x = 0$, $\tan x = x$ are transcendental equations.

We assume that the function f(x) is continuous in the required interval.

We define the following.

Root/zero: A number α, for which f(α) = 0 is called a root of the equation f(x) = 0, or a zero of f(x). Geometrically, a root of an equation f(x) = 0 is the value of x at which the graph of the equation y = f(x) intersects the x-axis.



'Root of f(x) = 0'

Various methods of computing numerical solutions of non linear equations are known. Here we discuss only Bisection Method.

Bisection Method

One of the first numerical methods developed to find the root of a nonlinear equation $f(x) = 0$ was the bisection method (also called binary-search method or **The Dichotomy Method**). The bisection method in mathematics is a root finding method which repeatedly bisects an interval and then selects a subinterval in which a root must lie for further processing. It is very simple but relatively slow method. Because of this, it is often used to obtain a rough approximation to a solution which is then used as a starting point for more rapidly converging methods.

This method is applicable when we wish to solve the equation

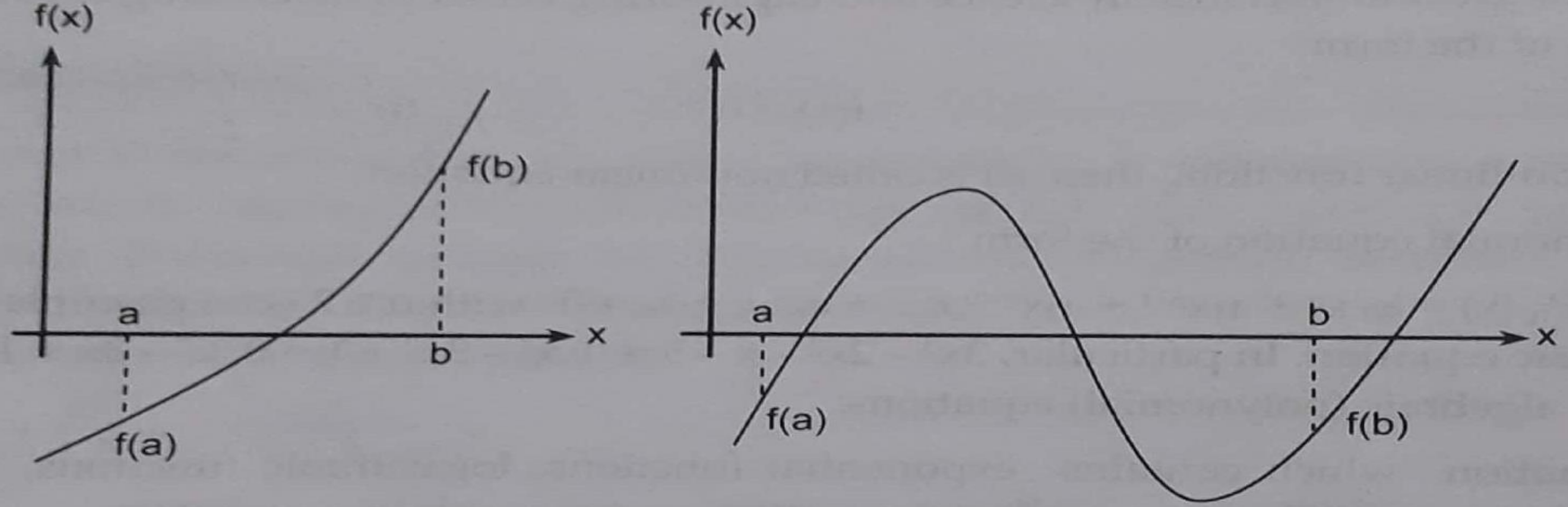
$f(x) = 0$ for the real variable x, where f is continuous function defined on an interval $[a, b]$ and $f(a)$ and $f(b)$ have opposite signs. This is based on the following intermediate value theorem.

Intermediate Value Theorem

Let f be continuous on $[a, b]$. If $f(a) \neq f(b)$, then for every k, between $f(a)$ and $f(b)$ there is a point $\alpha \in (a, b)$ such that $f(\alpha) = k$.

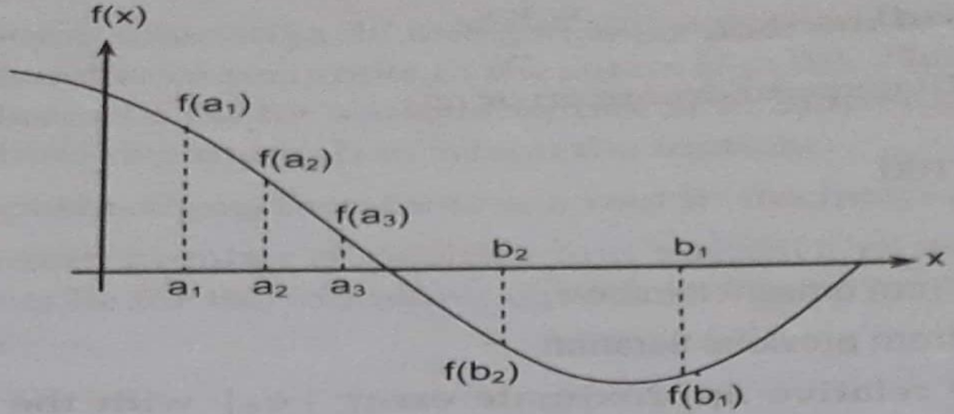
Bolzano's Theorem

Let f(x) be a real continuous function on [a, b]. If f (a). f (b) < 0 then an equation f(x) = 0 has at least root between a and b.



At each step the bisection method divides the interval in two by computing the midpoint $c = \frac{a + b}{2}$ of the interval and the value of the function f(c) at that point. Unless c is itself a root (which is very unlikely, but possible) there are now two possibilities: either f(a) and f(c) have opposite signs or f(c) and f(b) have opposite signs. If f(a) and f(c) have opposite sign then we select the sub interval [a, c] as a new interval to be used in next step, otherwise select [c, b]. In this way the interval that contains a zero of f is reduced by 50% at each step. The process is continued until the interval is sufficiently small.

Explicitly, if f(a) and f(c) have opposite signs then, the method sets c as new value of b and if f(b) and f(c) have opposite signs then the method sets c as new value of a. If f(c) = 0 then c is the solution and process stops.



The method is guaranteed to converge to a root of $f(x)=0$ if it is continuous function on the interval $[a, b]$ and $f(a)$ and $f(b)$ have opposite signs. The absolute error is halved at each step so the method converges linearly, which is comparatively slow.

Specifically, if $c_1 = \dfrac{a+b}{2}$ is the midpoint of the initial interval and c_n is the midpoint of the interval in n^{th} step, then the different between c_n and solution c is bounded by

$$|c_n - c| \leq \frac{b-a}{2^n}$$

This formula can be used to determine in advance the number of iterations that the bisection method would need to converge to a root to within a certain tolerance. The number of iterations needed, n, to achieve a given error (or tolerance) $\in$, is given by

$$2^n \leq \frac{b-a}{|c_n - c|} = \frac{\in_0}{\in}$$

where $\in_0 = b - a$ and $\in = \text{error} = |c_n - c|$

Taking log on both sides, we get

$$\log 2^n \leq \log \frac{\in_0}{\in}$$

$$n \log 2 \leq \log \in_0 - \log \in$$

$$n \leq \frac{\log \in_0 - \log \in}{\log 2}.$$

Algorithm for the Bisection Method

The steps to apply the bisection method to find the root equation $f(x) = 0$ are

1. Choose x_ℓ and x_u as two guesses for the root such that $f(x_\ell).f(x_u) < 0$, or in other words, $f(x)$ change signs between x_ℓ and x_u.

2. Estimate the root, x_m, of the equation $f(x) = 0$ as the mid-point between x_ℓ and x_u as

$$x_m = \frac{x_\ell + x_u}{2}$$

3. Now check the following

- If $f(x_\ell) .f(x_m) < 0$, then the root lies between x_ℓ and x_m; then $x_\ell = x_\ell$ and $x_u = x_m$.
- If $f(x_\ell). f(x_m) > 0$, then the root lies between x_m and x_u ; then $x_\ell = x_m$ and $x_u = x_u$.
- If $f(x_\ell). f(x_m) = 0$; then root is x_m. Stop the algorithm if this is true.

WORKED OUT EXAMPLES

EXAMPLE 1

Determine an interval of length one unit in which the negative real root, which is smallest in magnitude lies for the equation $9x^3 + 18x^2 - 37x - 70 = 0$.

Solution:

Let $f(x) = 9x^3 + 18x^2 - 37x - 70 = 0$. Since, the smallest negative real root in magnitude is required, we form a table of values for $x < 0$,

Table

x	-5	-4	-3	-2	-1	0
f(x)	- 560	-210	-40	4	-24	-70

Since, $f(-2).f(-1) < 0$, the negative root of smallest magnitude lies in the interval $(-2, -1)$.

EXAMPLE 2

Locate the smallest positive root of the equations

a. $xe^x = \cos x$.

b. $\tan x = 2x$.

Solution:

a. Let $f(x) = xe^x - \cos x = 0$.

We have

$f(0) = -1$, $f(1) = e - \cos 1 = 2.718 - 0.540 = 2.178$.

Since, $f(0) . f(1) < 0$, there is a root in the interval $(0, 1)$.

b. Let $f(x) = \tan x - 2x = 0$.

We have the following function values.

$f(0) = 0$, $f(0.1) = -0.0997$, $f(0.5) = -0.4537$, $f(1) = -0.4426$,

$f(1.1) = -0.2352$, $f(1.2) = 0.1722$.

Since, $f(1.1) . f(1.2) < 0$, the root lies in the interval $(1.1, 1.2)$.

EXAMPLE 3

interval (1.1, 1.2).

EXAMPLE 3

Determine a formula which relates the number of iterations, n, required by the bisection method to coverage to within an absolute error tolerance of ϵ, starting from the initial interval [a, b].

Solution:

The bisection method generates a sequence $\{c_n\}$ approximating a root c_0 of $f(x) = 0$

with $|c_n - c| \le \frac{b-a}{2^n}$

To converge to within an absolute error tolerance of ϵ means we need to have $|c_n - c| \le \epsilon$

So, $\frac{b-a}{2^n} \le \epsilon$

or, $2^n \ge \frac{b-a}{\epsilon}$

or, $n \log 2 \ge \log \frac{(b-a)}{\epsilon}$

$\therefore\ n \ge \frac{\log\left(\frac{b-a}{\epsilon}\right)}{\log 2}$ which gives the formula to find the number of iterations.

EXAMPLE 4

Write the three stopping criteria for bisection method.

Solution:

The three stopping criteria are

a. $|x_n - x_{n-1}| < \in$

b. $\dfrac{|x_n - x_{n-1}|}{|x_n|} < \in$

c. $|f(x_n)| < \in$

EXAMPLE 5

Use bisection method to find the solution $f(x) = x^3 - x - 2 = 0$ in the interval [1, 2] with accuracy 10^{-4}

Solution:

Here, $f(x) = x^3 - x - 2 = 0$

Then, $f(1) = (1)^3 - (1) - 2 = -2$ and

$f(2) = 2^3 - 2 - 2 = 4$

Since $f(1) . f(2) = -8 < 0$, then by Bolazo therom there is at least one root lying between $a = 1$ and $b = 2$.

No. of iterations	a_n	b_n	$c_n = \frac{a_n + b_n}{2}$	$f(c_n)$
1	1	2	1.5	- 0.125
2	1.5	2	1.75	1.609375
3	1.5	1.75	1.625	0.6660156
4	1.5	1.625	1.5625	0.2521973
5	1.5	1.5625	1.53125	0.0591125
6	1.5	1.53125	1.515625	- 0.0340538
7	1.515625	1.53125	1.5234375	0.0122504
8	1.515625	1.5234375	1.5195313	- 0.0109712
9	1.5195313	1.5234375	1.5214844	0.000622

Here $|f(c_9)| = 0.0006222 < 10^{-4}$

$\therefore$ the root of given equation is 1.521 with accuracy 10^{-4}.

EXAMPLE 6

Find the solution of $f(x) = x^2 - 3 = 0$ in the interval [1, 2] with error = 0.01.

Solution:

Here, $f(x) = x^2 - 3$

$f(1) = 1 - 3 = -2 < 0$

$f(2) = 4 - 3 = 1 > 0$

$\therefore \quad f(1).\, f(2) = -2 < 0$

So there is at least one root lying in the interval [1, 2]. So,

n	a	b	$c = \frac{a+b}{2}$	f(a)	f(b)	f(c)
1	1	2	1.5	-2	1	-0.75
2	1.5	2	1.75	-0.75	1	0.062
3	1.5	1.75	1.625	-0.75	0.0625	-0.3594
4	1.625	1.75	1.6875	-0.3594	0.0625	-0.1523
5	1.6875	1.75	1.7188	-0.1523	0.0625	-0.0457
6	1.7188	1.75	1.7344	-0.0457	0.0625	0.0081

Now, $f(1.7344) = 0.0081 < 0.01$.

$\therefore x = 1.7344$ is the solution of the equation $x^2 - 3 = 0$

EXAMPLE 7

Use Bisection method to find the root of the equation $f(x) = x^3 + 4x^2 - 10 = 0$ with error 10^{-3}.

Solution:

Here, $f(x) = x^3 + 4x^2 - 10$

$f(1) = 1 + 4 - 10 = -5 < 0$

$f(2) = 8 + 16 - 10 = 14 > 0$

$f(1).f(2) = -70 < 0$

So there is at least one root in the interval $[1, 2]$

n	a	b	$c = \frac{a+b}{2}$	f(a)	f(b)	f(c)
1	1	2	1.5	-5	14	2.375
2	1	1.5	1.25	-5	2.375	-1.797
3	1.25	1.5	1.375	-1.797	2.375	0.162
4	1.25	1.375	1.3125	-1.797	0.162	-0.848
5	1.3125	1.375	1.343750	-0.848	0.162	-0.351
6	1.34375	1.375	1.359375	-0.351	0.162	-0.096
7	1.359375	1.375	1.367188	-0.096	0.162	0.032
8	1.359375	1.367188	1.363281	-0.096	-0.032	0.032
9	1.363281	1.367188	1.365234	-0.032	-0.032	0.000

Now, $f(1.365234) = 0.000 < 10^{-3}$

$\therefore$ $1.365234 \approx 1.365$ is the solution.

EXAMPLE 8

Use bisection method to find the solution of the equation $x^2 - \sin x - 0.5 = 0$ in the interval [0, 2] with $\in = 10^{-2}$.

Solution:

Here,

$$x^2 - \sin x - 0.5 = 0$$

Let $f(x) = x^2 - \sin x - 0.5$

Then, $f(0) = -0.5 < 0$ and

$f(2) = 2.59 > 0$

$\therefore$ $f(0).\ f(2) < 0$

So there is at least one root lying between 0 and 2.

n	a	b	$c = \frac{a+b}{2}$	$\frac{b-a}{2}$
1	0	2	1	1
2	1	2	1.5	0.5
3	1	1.5	1.25	0.25
4	1	1.25	1.125	0.125
5	1.125	1.25	1.1825	0.0625
6	1.1875	1.25	1.2188	0.03125
7	1.1875	1.2188	1.2031	0.015625
8	1.1875	1.2031	1.1953	0.0078125

Also $f(1.1953)| = |-0.0015836|$

$= 0.001536 < 10^{-2}$

$\therefore$ $x = 1.1953$ is the solution

EXAMPLE 9

Find a bound for the number of iterations required by the bisection method to achieve an approximation with accuracy 10^{-4} lying in the interval [1, 2].

Solution:

Let n be a bound for the number of iterations.

Then we have

$$\frac{b-a}{2^n} \leq \epsilon \text{ where } a = 1, b = 2 \ \epsilon = 10^{-4}$$

$$\Rightarrow \frac{2-1}{2^n} \leq 10^{-4}$$

$$\Rightarrow 2^n \geq 10^4$$

$$\Rightarrow n \log 2 \geq 4 \log 10$$

$$\Rightarrow n \times 0.3010 \geq 4 \times 1$$

$$\Rightarrow n = 13.28 \approx 14$$

$\therefore$ Minimum number of iterations required n = 14.

EXAMPLE 10

Determine the number of positive and negative roots of the equation $x^5 - 4x^4 + 3x^3 - 2x^2 + x + 1 = 0$

Solution:

Here the equation is

$$f(x) = x^5 - 4x^4 + 3x^3 - 2x^2 + x + 1$$

$$+ \quad - \quad + \quad - \quad + \quad +$$

$$f(-x) = -x^5 - 4x^4 - 3x^3 - 2x^2 - x + 1$$

$$- \quad - \quad - \quad - \quad - \quad +$$

Since f(x) has 4 changes in sign so the function f(x) has 4 positive roots and there is only one change in sign of f(−x) so f(x) has 1 negative root.

Numerical Computing

- Numerical computing is an approach for solving complex mathematical problems using simple arithmetic operations.

- Most of the Numerical computations depend upon the method of computation.

- A computational methods deals with process for solving problems with a computer device or computer.

- The computational method is judged on the basis of accuracy, rate of convergence, numerical stability and efficiency.

Accuracy

Accuracy refers to how well a numerical method matches the expected result.

Approximation + Error = Accuracy

For e.g. Value of π is not fixed. When we estimate 3.1416 then better estimation is 3.141592965.

Rate of Convergence:

The rate of convergence is rate in which it converges to the true solution. It is important to test for convergence before a numerical technique is used.

Numerical Stability:

A numerical method is said to be stable if it decreases error step by step or iteration by iteration.

Efficiency:

Efficiency of the numerical method means the amount of effort required by both human and computing devices for implementation of method to obtain the true solution.

Approximation and Error in Computing:

- Approximation and errors are integral part of every walk of human life.
- Dealing with the issue of error, we need to identify where the error is coming from, followed by quantify the error and lastly minimize the error as per our needs.

True error:

True error is denoted by E_t is the difference between the true value (or exact value) and the approximate value. i.e. True error = True value – Approximate value.

Round-off-error:

Round off error occurs because of the computing device's inability to deal with inexact numbers. Such numbers need to be rounded off to some near approximation which is dependent on the word size used to represent numbers of the device.

Rounded –off a number can be done in two ways: one is known as chopping and the other is known as symmetric rounding.

a) **Chopping:**

In chopping the extra digits, after the required number of significant digits are dropped. Suppose we are using a computing device with fixed word length of four digits then a number like 25.46859 will be stored as 25.46, dropping the digits 859.

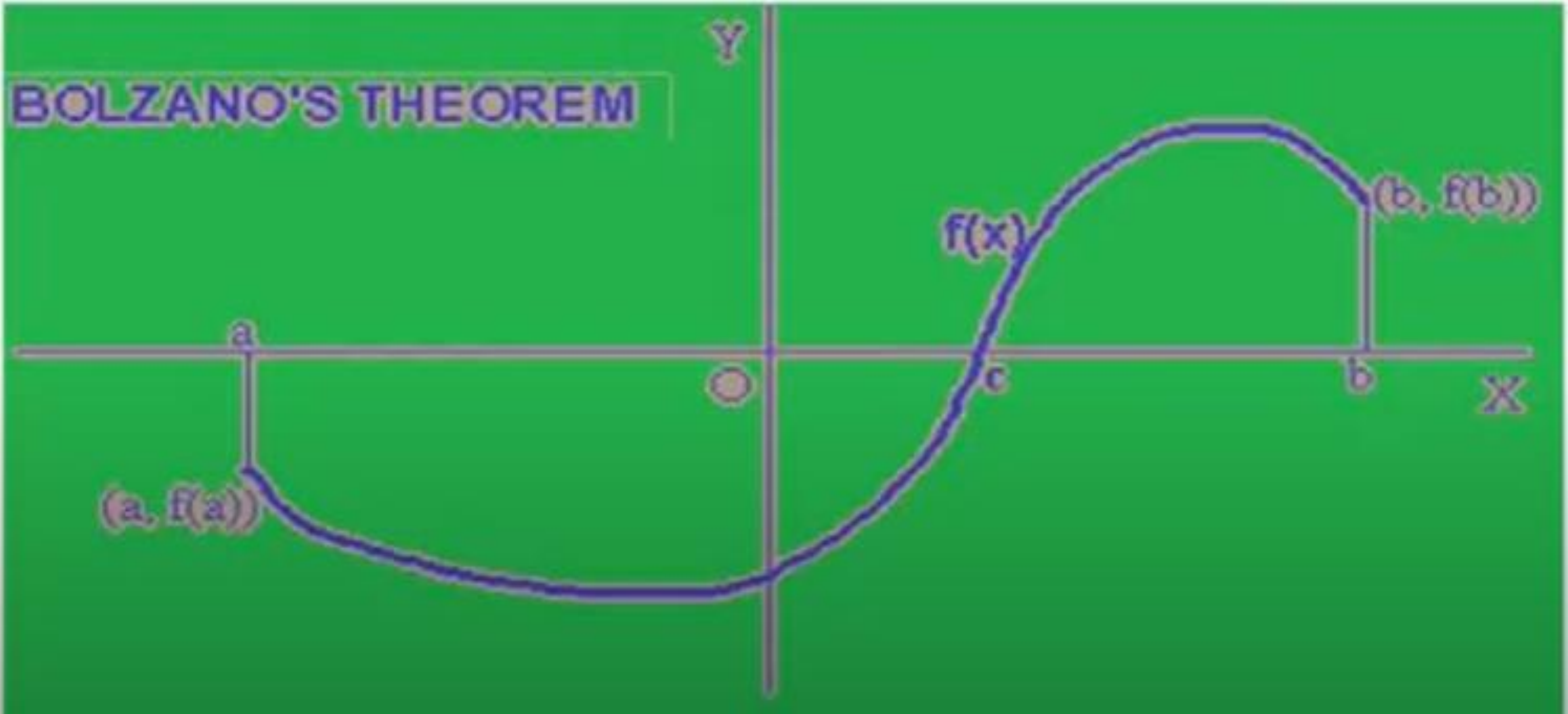
b) **Symmetric Round off**:

In symmetric round off, the last retained significant digit is rounded up by 1 (increased by 1) if the first discarded digit is 5 or greater than 5; otherwise the retained digit is unchanged. Example: 456 rounded up in 100 place is 500.

Method of finding roots of non-linear equations : Iteration Method:

Bolzano's Theorem:

If f(x) be real and continuous function on [a, b], f(a) and f(b) has opposite sign and f(a) .f(b)<0, then f(x) = 0 has at least one root on (a, b)



Example of Bolzano Theorem:

Q. Use Bolzano's theorem to show that $f(x) = x^2-2 = 0$ has a root somewhere in the interval (1, 2).

Here,

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

$$f(1) = 1^2-2 = -1<0$$

Again, $f(2) = 2^2-2 = 1 >0$

Here, $f(1) . f(2) < 0$ using Bolzano's theorem , $f(a) = 0$ has at least a root in the interval (1, 2).

Checking: If $x^2-2 = 0$ Then, $x = 1.4142..$ is the solution which is also matching from the above condition.

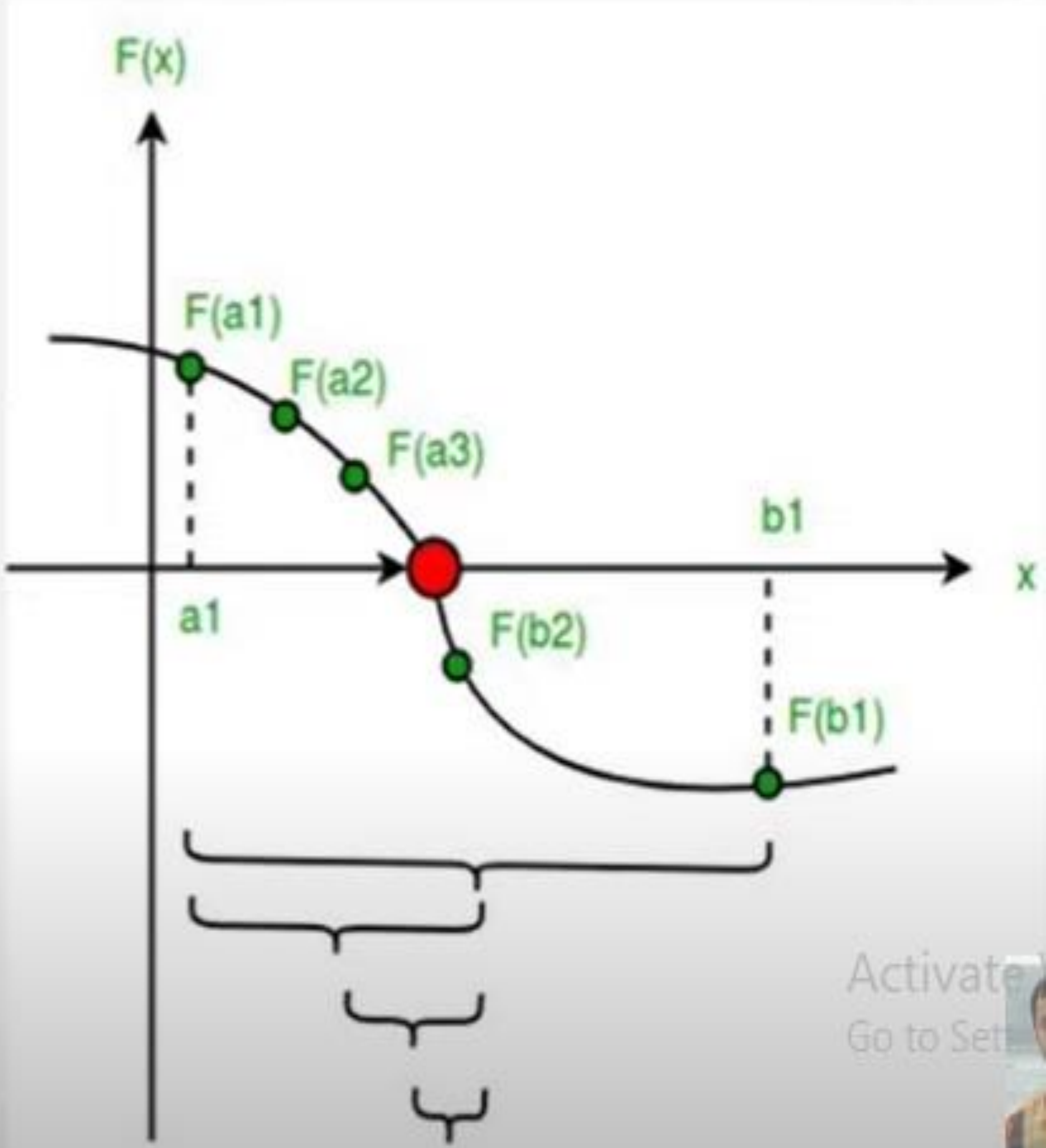
Bisection Method

The bisection method in Mathematics, is a root-finding algorithm which repeatedly bisects an interval then selects a sub-interval in which a root must lie for further processing.

It is also known as bracketing method for finding roots of equations.

It requires two initial points a and b such that f(a) and f(b) have opposite signs.

This is called a bracket of a root.



Exercise 20.1

1. a. Find the truncation error in the following series when x=1 taking first four terms

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots \quad \text{where } n! = n(n-1)(n-2)\dots3.2.1$$

b. Find the difference $\sqrt{6.37} - \sqrt{6.36}$ to three significant figures.

c. The exponential series is given by $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$

Find the value of e correct to four decimal places.

d. Compute the value of ln3 correct to five decimal places

e. The Maclaurin expansion of sin x is given by $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots$ where, x is in radius. Us the series to compute the value of sin 25^0 to an accuracy of 0.001.

2. a. Show that the function $f(x) = x^3 - x^2 + x - 1 = 0$ has three positive roots and no negative root. Fir the positive root lying in the interval [0, 2] correct to two places of decimal.

b. Show that the equation $f(x) = x^3 - x - 4 = 0$ has one positive root and using the method bisection, find the positive root correct to 3 places of decimal.

3. Apply method of successive bisection to find the

a. Square root of 2 within three places of decimal in the interval [1, 2].

b. Square root of 5 within two places of decimal in [2, 3].

c. Cube root of 26 correct to 10^{-1} in the interval [5, 6]

4. Apply method of successive bisection to find the root of the equation.

a. $x^2 - 2x - 5 = 0$ in the interval [4, 5] correct to three places of decimal.

b. $x^3 - 7x^2 + 14x - 6 = 0$ in the interval [0, 1] with accuracy 10^{-2}.

c. $x + \log_{10} x - 2 = 0$ correct to two decimal places in the interval [1, 2]

5. How many iterations do you need to get the root if you start with a = 2 and b = 3 and the tolerance is 10^{-3}?
6. Find minimum number of iterations required by the bisection method to achieve an approximation with accuracy 10^{-1} to the solution of $x^3 + x - 4 = 0$ lying in the interval [1, 4]. Find an approximation to the root with this degree of accuracy.
7. Find the root of the equation $4\,e^{-x}.\sin x - 1 = 0$ between 0 and 0.5 in three places of decimal.

☑ Answers:

1.	a. 0.0516152	b. 0.198×10^{-2}	c. 2.7183	d. 1.09861	e. 0.423
2.	a. 0.999	b. 1.797			
3.	a. 1.414	b. 2.24	c. 5.099		
4.	a. 3.450	b. 0.5859	c. 1.75		
5.	10	6. 5, 1.47	7. 0.370		

NUMERICAL COMPUTATION

1. a. Find the truncation error in the following series when x = 1 taking first four terms

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots. + \frac{x^n}{n!} + \ldots \text{ where } n! = n(n-1)(n-2)\ldots3.2.1$$

b. Find the difference $\sqrt{6.37} - \sqrt{6.36}$ to three significant figures.

c. The exponential series is given by $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$

Find the value of e correct to four decimal places.

d. Compute the value of ln3 correct to five decimal places

Solution:

a. Here, $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^3}{3!} + \ldots\ldots\ldots + \frac{x^n}{n!} + \ldots\ldots$

If x = 1, $r = e' = 1 + \frac{1}{11} + \frac{1^2}{2!} + \frac{1^3}{3!} + \ldots\ldots + \frac{1^n}{n!} + \ldots\ldots = 2.718281828$

But if we take only four terms, then

$r_1 = 1 + \frac{1}{1!} + \frac{1^2}{2!} + \frac{1^3}{3!} = 2.666667$

So Transaction error = $|r - r_1|$ = |2.718281828 – 2.666667| = |0.051615128|= 0.05161528

b. $\sqrt{6.37} - \sqrt{6.36} = 2.5238859 - 2.521904 = 1.98 \times 10^{-3} = 0.98 \times 10^{-2}$

c. Here, $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots\ldots\ldots$

Now, $e' = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots\ldots\ldots\ldots$

or, e = 1 + 1 + 0.5 + 0.16667 + 0.041667 + ... = 2.70833667

So the correct to four decimal place of e = 2.7083

d. The correct value of $\ln_3$ to five decimal place is 1.09861

2. a. Show that the function $f(x) = x^3 - x^2 + x - 1 = 0$ has three positive roots and no negative root. Find the positive root lying in the interval [0, 2] correct to two places of decimal.

b. Show that the equation $f(x) = x^3 - x - 4 = 0$ has one positive root and using the method of bisection, find the positive root correct to 3 places of decimal.

Solution:

$f(x) \quad = x^3 - x^2 + x - 1$

$\quad\quad\quad\; + \quad - \quad + \quad -$

and $f(-x) = -x^3 - x^2 - x - 1$

$\quad\quad\quad\; - \quad - \quad - \quad -$

Since f(x) has 3 changes in sign so the function has 4 positive roots and there is no changes in sign in f(–x). So there is no negative roots.

Here, $f(x) = x^3 - x^2 + x - 1$

$f(0) = -1$ & $f(2) = 2^3 - 2^2 + 2 - 1 = 8 - 4 + 1 = 5$

So, $f(0).\ f(2) = -1 \quad 5 = -5<0$

Hence, root lies between 0 and 2.

S.N.	a	b	$n = \frac{a+b}{2}$	f(a)	f(b)	f(m)	$\frac{b-a}{2}$
1	0	2	1	–1	5	0	1
2	0	1	0.5	–1	0	–0.625	0.5
3	0.5	1	0.75	–0.625	0	–0.390625	0.25
4	0.75	1	0.875	–0.390625	0	–0.2207	0.125
5	0.875	1	0.9375	–0.2207	0	–0.117432	0.0625
6	0.9375	1	0.96875	–0.117432	0	–0.09085	0.03125
7	0.96875	1	0.984375	–0.09085	0	–0.0307655	0.015625
8	0.984375	1	0.9921875				0.0078125

Since, $\frac{b-a}{2} = 0.0078125$ is very small, the required the root is 0.99 (upto 2 decimal point)

b. Similar to 'a'

3. Apply method of successive bisection to find the
   a. Square root of 2 within three places of decimal in the interval [1, 2].
   b. Square root of 5 within two places of decimal in [2, 3].
   c. Cube root of 26 correct to 10^{-1} in the interval [5, 6]

Solution:

a. Let x be the sq. root of 2.

Then,

$x = \sqrt{2}$

$x^2 = 2$

or, $x^2 - 2 = 0$

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

$f(1) = -1$

$f(2) = 2$

$f(1).\ f(2) = -1 \times 2 = -2 < D$

There exists a root in between 1 & 2 then, the calculation of square root of 2 using successive reaction method as shown in table.

a	$m = \frac{a+b}{2}$	b	f(a)	f(m)	f(b)
1	1.5	2	–1	0.25	1
1	1.25	1.5	–1	–0.437	0.5
1.25	1.375	1.5	–0.437	–0.109	0.25
1.375	1.4375	1.5	–0.109	0.0664	0.25
1.375	1.40625	1.4375	–0.109	–0.02246	0.0625
1.40625	1.421875	1.4375	–0.022	0.0217	0.03125
1.40625	1.4140625	1.421875	–0.0004	0.0106	0.007812
1.4140625	1.416015625	1.417968	–0.0004	0.0051	0.0039062
1.4140625	1.415039063	1.416015	–0.0004	0.0023	0.001953
1.4140625	1.414550782	1.415039	–0.0004	0.0095	0.00976

Here, the error 0.00976 is within 10^{-3} so the required root of the given equation of mid-point is 1.41.

b. Similar to (a)

3. Apply method of successive bisection to find the
   a. Square root of 2 within three places of decimal in the interval [1, 2].
   b. Square root of 5 within two places of decimal in [2, 3].
   c. Cube root of 26 correct to 10^{-1} in the interval [5, 6]

Solution:

a. Let x be the sq. root of 2.

Then,

$x = \sqrt{2}$

$x^2 = 2$

or, $x^2 - 2 = 0$

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

$f(1) = -1$

$f(2) = 2$

$f(1).\ f(2) = -1 \times 2 = -2 < D$

There exists a root in between 1 & 2 then, the calculation of square root of 2 using successive reaction method as shown in table.

a	$m = \frac{a+b}{2}$	b	f(a)	f(m)	f(b)
1	1.5	2	–1	0.25	1
1	1.25	1.5	–1	–0.437	0.5
1.25	1.375	1.5	–0.437	–0.109	0.25
1.375	1.4375	1.5	–0.109	0.0664	0.25
1.375	1.40625	1.4375	–0.109	–0.02246	0.0625
1.40625	1.421875	1.4375	–0.022	0.0217	0.03125
1.40625	1.4140625	1.421875	–0.0004	0.0106	0.007812
1.4140625	1.416015625	1.417968	–0.0004	0.0051	0.0039062
1.4140625	1.415039063	1.416015	–0.0004	0.0023	0.001953
1.4140625	1.414550782	1.415039	–0.0004	0.0095	0.00976

Here, the error 0.00976 is within 10^{-3} so the required root of the given equation of mid-point is 1.41.

b. Similar to (a)

4. Apply method of successive bisection to find the root of the equation.
   a. $x^2 - 2x - 5 = 0$ in the interval [4, 5] correct to three places of decimal.
   b. $x^3 - 7x^2 + 14x - 6 = 0$ in the interval [0, 1] with accuracy 10^{-2}.
   c. $x + \log_{10} x - 2 = 0$ correct to two decimal places in the interval [1, 2]

Solution:

a. Here, $f(x) = x^2 - 2x - 5 = 0$, interval = [3, 4]
So, $f(3) = 3^2 - 2 \times 3 - 5 = -2$
$f(4) = 4^2 - 2 \times 4 - 5 = 3$
Since, $f(3).\ f(4) = -2 \times 3 = -6 < 0$
The root lies between 3 & 4
Now,

n	a	b	$m = \frac{a+b}{2}$	f(a)	f(b)	f(m)	b – a
1	3	4	3.5	–2	3	0.25	1
2	3	3.5	3.25	–2	0.25	–0.9375	0.5
3	3.25	3.5	3.375	–9375	0.25	–0.359375	0.25
4	3.375	3.5	3.4375	–0.359375	0.25	0.058594	0.125
5	3.4375	3.5	3.46875	–0.058594	0.25	0.09473	0.0625
6	3.4375	3.46875	3.453125	–0.058594	0.09473	0.0178223	0.03125
7	3.4375	3.453125	3.4453	–0.058594	0.0178223	–0.192322	0.015625
8	3.4453	3.453125	3.4492125	–0.192322	0.0178223		0.007825

Since, b – a = 0.007825 is a small enough the root of $x^2 - 2x - 5 = 0$ in [3, 4] is 3.449

b. $x^3 - 7x^2 + 14x - 6 = 0$

$f(x) = x^2 - 7x^2 + 14x - 6$

$f(0) = 0^2 - 7 \times 0^2 + 14 \times 0 - 6 = -6$

$f(1) = 1^2 - 7 \times 1^2 + 14 \times 1 - 6$

$f(0) \times f(1) = -12 < 0$

So, there exist a root in the interval [0, 1]

a	$m = \frac{a+b}{2}$	b	f(0)	f(m)	f(b)	b – a
0	0.5	1	–6	–0.625	2	1
0.5	0.75	1	–0.625	0.9843	2	0.5
0.5	0.625	0.79	–0.625	0.2597	0.9843	0.25
0.5	0.5625	0.625	–0.625	–0.16126	0.2597	0.125
0.5625	0.59375	0.625	–0.16186	0.05404	0.2599	0.06
0.5625	0.578125	0.59375	–0.16186	–0.05262	0.05404	0.03
0.578125	0.5859375	0.59375	–0.0562	0.001030	0.054	0.0156
0.578125	0.58203125	0.0859375	–0.05262	–0.02571	0.00103	0.00781

The root is 0.58.

c. $x + \log_{10}x - 2 = 0$

$f(x) = x + \log_{10}x - 2$

$f(1) = -1$

$f(2) = 0.301$

$f(1).f(2) = -0.301$

There exist a root between the interval [1, 2]

a	$\mathbf{m = \frac{a+b}{2}}$	**b**	**f(a)**	**f(m)**	**f(b)**
1	1.5	2	–1	–0.32930	0.301
1.5	1.75	2	–0.32390	–0.00696	0.301
1.75	1.875	2	–0.00696	0.148001	0.301
1.75	1.8125	1.875	–0.00696	0.07077	0.148001
1.75	1.78125	1.8125	–0.0696	0.03197	0.7077
1.75	1.765625	1.78125	–0.0696	0.01252	0.08197
1.75	1.7532825	1.765625	–0.0696	–0.00286	0.01252
1.75328125	1.759453125	1.765625	–0.00286	0.004830	0.01252

The root is 1.75 up to two place of decimal.

5. How many iterations do you need to get the root if you start with a = 2 and b = 3 and the tolerance is 10^{-3}?

Solution:

Here, a = 2, b = 3, tolerance = 10^{-3} no. of iteration = ?

Since the error E should be smaller than 10^{-3}

So we have,

$$\frac{b-a}{2^i} \le Ei$$

or, $\frac{3-2}{2^i} \le 10^{-3}$

or, $\frac{1}{2^i} \le \frac{1}{10^3}$

or, $2^i \ge 10^3$

Taking log on both sides, we get,

i log2 $\ge$ 3 log 10

or, $i(0.3010) \ge 3 \times 1$

or, $i \ge \frac{3}{0.3010} = 9.96$

So, the number of iteration should be 10 or more than 10.

6. Find minimum number of iterations required by the bisection method to achieve an approximation with accuracy 10^{-1} to the solution of $x^3 + x - 4 = 0$ lying in the interval [1, 4]. Find an approximation to the root with this degree of accuracy.

Solution:

Let $f(x) = x^3 + x - 4$, internal [1, 4]

Now, $f(1) = 1^3 + 1 - 4 = -2$

$f(4) = 4^3 + 4 - 4 = 64$

and $f(1) . f(4) = -2 \times 64 = -128 < 0$

So root lies between 1 & 4

For number of iteration

$a = 1, b = 4, \sum_i = 10^{-1}$

Since, $\dfrac{|b - a|}{2^i} \leq 10^{-1}$

or, $\frac{|4 - 1|}{2^i} \leq 10^{-1}$

$\Rightarrow \quad 3 \times 10 \leq 2^i$

or, $\log 30 \leq i \log 2$

or, $1.47712 \leq i \times 0.30103$

i.e. the number of iteration should be 5 or more.

Now, the root lies in [1, 4]

So, $C_0 = \frac{1 + 4}{2} = 2.5$

& $f(2.5) = 2.5^3 + 2.5 - 4 = 14.125 > 0$

So, $f(1) . f(2.5) = -2 \times 14.125 = -28.25 < 0$

Hence, root lies in [1, 2.5]

Again, $C_1 = \frac{1 + 2.5}{2} = 1.75$

and $f(1.75) = 3.109375$

So root lies between 1 and 1.75

Now, $C_2 = \frac{1 + 1.75}{2} = 1.375$

& $f(1.375) = -0.02539$ whose absolute value is less than 10^{-1}

So the required root is 1.375.