Prof. Jelili O. Oyelade

Computational Science and Numerical Methods

Solution of Nonlinear Equations
(Root Finding Problems)

- Definitions
- Classification of Methods
  - Analytical Solutions
  - Graphical Methods
  - **Numerical Methods**
    - **Bracketing Methods (bisection, regula-falsi)**
    - **Open Methods (secant, Newton-Raphson, fixed point iteration)**
- Convergence Notations

Definitions

- **Computational science is multidisciplinary, involving researchers in the physical sciences; that is, computer scientists, software engineers, even people in the social sciences. The methods usually involve numerical simulations.**

ISSUES IN NUMERICAL ANALYSIS

- **WHAT IS NUMERICAL ANALYSIS?**
  - It is a way to do highly complicated mathematics problems on a computer.
  - It is also known as a technique widely used by scientists and engineers to solve their problems.
- **TWO ISSUES OF NUMERICAL ANALYSIS:**
  - How to compute? This corresponds to algorithmic aspects;
  - How accurate is it? That corresponds to error analysis aspects.

- **ADVANTAGES OF NUMERICAL ANALYSIS:**
  - It can obtain numerical answers of the problems that have no "analytic" solution.
  - It does NOT need special substitutions and integrations by parts. It needs only the basic mathematical operations: addition, subtraction, multiplication and division, plus making some comparisons.
- **IMPORTANT NOTES:**
  - Numerical analysis solution is always numerical.
  - Results from numerical analysis is an approximation.

- **NUMERICAL ERRORS**

When we get into the real world from an ideal world and finite to infinite, errors arise.

- SOURCES OF ERRORS:
    - Mathematical problems involving quantities of infinite precision.
    - Numerical methods bridge the precision gap by putting errors under firm control.
    - Computer can only handle quantities of finite precision.

– TYPES OF ERRORS:

- Truncation error (finite speed and time) - An example:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \left(1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!}\right) + \sum_{n=4}^{\infty} \frac{x^n}{n!}$$

$$= p_3(x) + \sum_{n=4}^{\infty} \frac{x^n}{n!}$$

- Round-off error (finite word length): All computing devices represent numbers with some imprecision, except for integers.

- Human errors: (a) Mathematical equation/model. (b) Computing tools/machines. (c) Error in original data. (d) Propagated error.

– MEASURE OF ERRORS:

Let a be a scalar to be computed and let $\bar{a}$ be its approximation.

Then, we define

- Absolute error = | true value – approximated value |.

$$e = |a - \bar{a}|$$

- Relative error = $\left| \frac{\text{true value} - \text{approximated value}}{\text{true value}} \right|$

$$e_r = \left| \frac{a - \bar{a}}{a} \right|$$

Example: Let the true value of π be 3.1415926535898 and its approximation be 3.14 as usual. Compute the absolute error and relative error of such an approximation.

The absolute error:

$$e = \left| p - \bar{p} \right| = \left| 3.1415926535898 - 3.14 \right| = 0.0015926535898$$

which implies that the approximation is accurate up to 2 decimal places.

The relative error:

$$e_r = \left| \frac{p - \bar{p}}{p} \right| = \frac{0.0015926535898}{3.1415926535898} = 0.000506957382897$$

which implies that the approximation has a accuracy of 3 significant figures.

- **STABILITY AND CONVERGENCE**
  - STABILITY in numerical analysis refers to the trend of error change iterative scheme. It is related to the concept of convergence.

    It is stable if initial errors or small errors at any time remain small when iteration progresses. It is unstable if initial errors or small errors at any time get larger and larger, or eventually get unbounded.

  - CONVERGENCE: There are two different meanings of convergence in numerical analysis:

    a. If the discretized interval is getting finer and finer after dicretizing the continuous problems, the solution is convergent to the true solution.

    b. For an iterative scheme, convergence means the iteration will get closer to the true solution when it progresses.

Solutions to Nonlinear Equations (Computing Zeros)

- **Problem:** Given a function $f(x)$, which normally is nonlinear, the problem of "computing zeros" means to find all possible points, say

$$\tilde{x}_0, \tilde{x}_1, \cdots, \tilde{x}_n$$

such that

$$f(\tilde{x}_0) = f(\tilde{x}_1) = \ldots = f(\tilde{x}_n) = 0$$

However, it is often that we are required to find a single point $\tilde{x}_0$ in certain interval, say [a,b] such that

$$\boxed{f(\tilde{x}_0) = 0}$$

Nonlinear Equation Solvers

Bracketing

- **Bisection**
- **False Position (Regula-Falsi)**

Graphical

Open Methods

- **Newton Raphson**
- **Secant**
- **Fixed point iteration**

All Iterative

Thermodynamics:

van der Waals equation; v = V/n (= volume/# moles)

Find the molecular volume v such that

$$f(v) = \left(p + \frac{a}{v^2}\right)(v - b) - RT = 0$$

p = pressure,

T = temperature,

R = universal gas constant,

a & b = empirical constants

Civil Engineering:

Find the **horizontal component of tension, *H*,** in a cable that passes through $(0,y_0)$ and (x, y)

$$f(H) = \frac{H}{w}\left[\cosh\left(\frac{wx}{H}\right) - 1\right] + y_0 - y = 0$$

w = weight per unit length of cable

Electrical Engineering :

Find the **resistance, *R*,** of a circuit such that the charge reaches *q* at specified time *t*

$$f(R) = e^{-Rt/2L} \cos\left[\left(\sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2}\right) t\right] - \frac{q}{q_0} = 0$$

L = inductance,

C = capacitance,

q_0 = initial charge

Mechanical Engineering:

Find the **value of stiffness k** of a vibrating mechanical system such that the displacement x(t) becomes zero at t= 0.5 s. The initial displacement is x_0 and the initial velocity is zero. The mass m and damping c are known, and $\lambda = c/(2m)$.

$$\mathbf{x(t) = x_0 e^{-\lambda t}\left[\cos(\mu t) + \frac{\lambda}{\mu}\sin(\mu t)\right] = 0}$$

in which

$$\mu = \sqrt{\frac{k}{m} - \frac{c^2}{4m^2}}$$

Root Finding Problems

Many problems in Science and Engineering are expressed as:

$$\boxed{\text{Given a continuous function } f(x), \text{ find the value } r \text{ such that } f(r) = 0}$$

These problems are called root finding problems.

Roots of Equations

A number ***r*** that satisfies an equation is called a root of the equation.

The equation: $x^4 - 3x^3 - 7x^2 + 15x = -18$

has four roots: $-2,\ 3,\ 3,\ and\ -1$.

i.e., $x^4 - 3x^3 - 7x^2 + 15x + 18 = (x+2)(x-3)^2(x+1)$

The equation has two simple roots (-1 and -2) and a repeated root (3) with multiplicity = 2.

Zeros of a Function

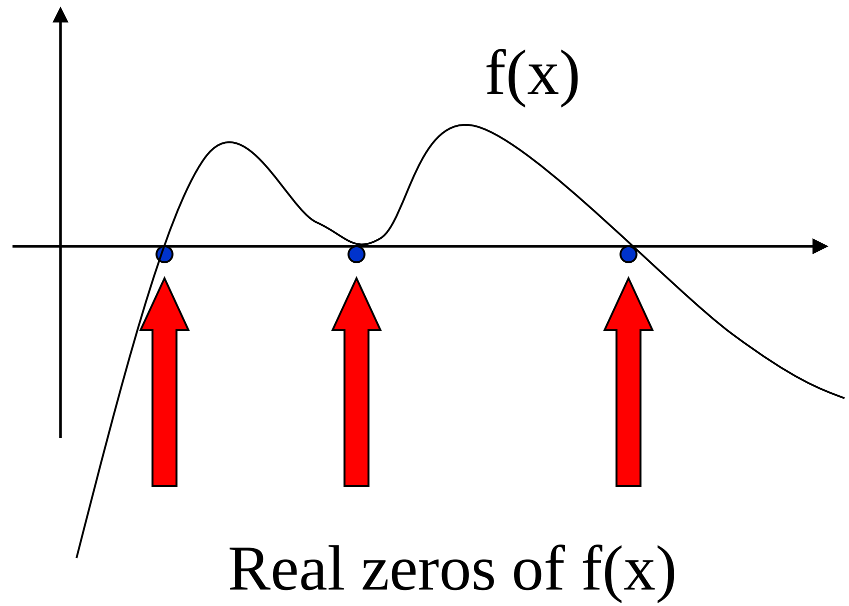
Let $f(x)$ be a real-valued function of a real variable. Any number r for which $f(r)=0$ is called a zero of the function.

Examples:

2 and 3 are zeros of the function $f(x) = (x-2)(x-3)$.

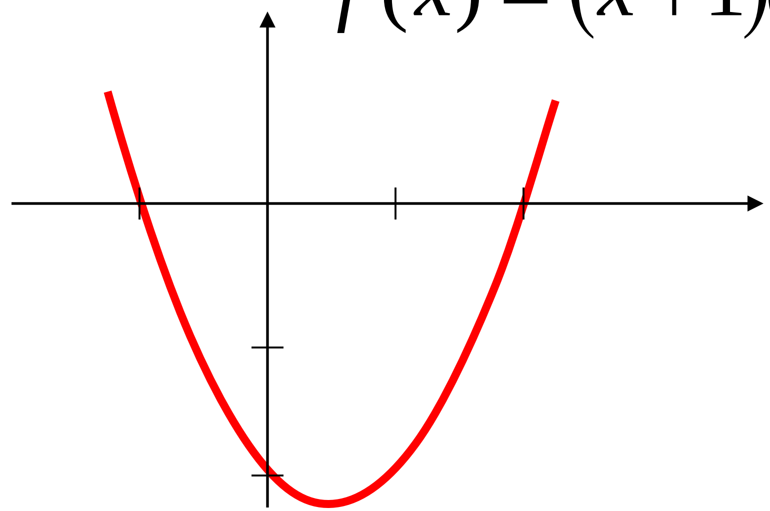
Graphical Interpretation of Zeros

- The real zeros of a function *f(x)* are the values of *x* at which the graph of the function crosses (or touches) the x-axis.



Simple Zeros

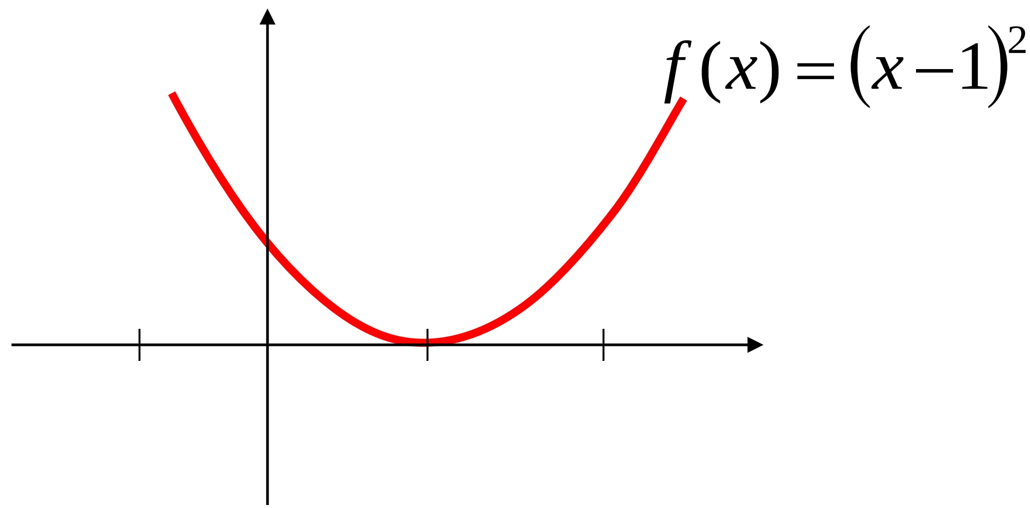
$$f(x) = (x+1)(x-2)$$



$$f(x) = (x+1)(x-2) = x^2 - x - 2$$

has two simple zeros (one at $x = 2$ and one at $x = -1$)

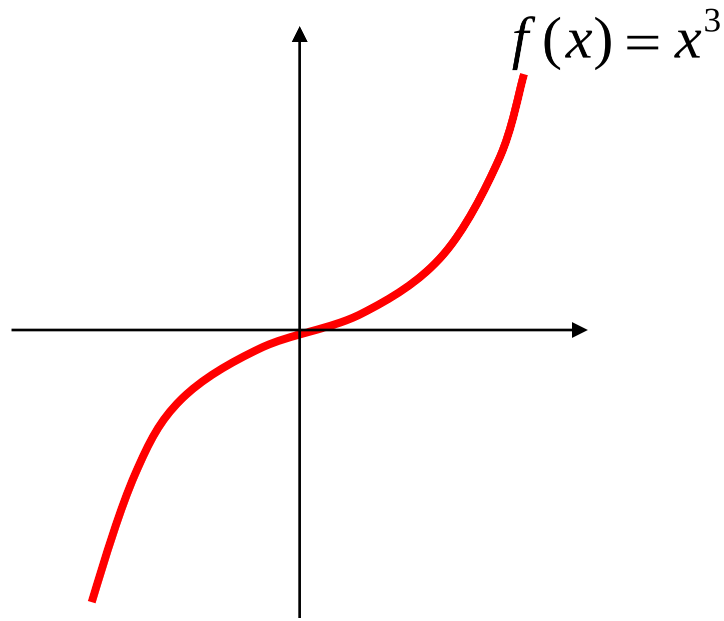
Multiple Zeros



$$f(x) = (x-1)^2 = x^2 - 2x + 1$$

has double zeros (zero with muliplicity $= 2$) at $x = 1$

Multiple Zeros



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

has a zero with muliplicity $y=3$ at $x=0$

Facts

- Any n^{th} order polynomial has exactly n zeros (counting real and complex zeros with their multiplicities).
- Any polynomial with an odd order has at least one real zero.
- If a function has a zero at $x=r$ with multiplicity m then the function and its first $(m-1)$ derivatives are zero at $x=r$ *and the* m^{th} *derivative at* r *is not zero*.

Roots of Equations & Zeros of Function

Given the equation :

$$x^4 - 3x^3 - 7x^2 + 15x = -18$$

Move all terms to one side of the equation :

$$x^4 - 3x^3 - 7x^2 + 15x + 18 = 0$$

Define $f(x)$ as :

$$f(x) = x^4 - 3x^3 - 7x^2 + 15x + 18$$

The zeros of $f(x)$ are the same as the roots of the equation $f(x) = 0$

(Which are -2, 3, 3, and -1)

Solution Methods

Several ways to solve nonlinear equations are possible:

- Analytical Solutions
  - Possible for special equations only
- Graphical Solutions
  - Useful for providing initial guesses for other methods
- Numerical Solutions
  - Open methods
  - Bracketing methods

Analytical Methods

Analytical Solutions are available for special equations only.

$$\text{Analytical solution of}: \; a\,x^2 + b\,x + c = 0$$

$$roots = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$\boxed{\text{No analytical solution is available for}: \quad x - e^{-x} = 0}$$

Graphical Methods

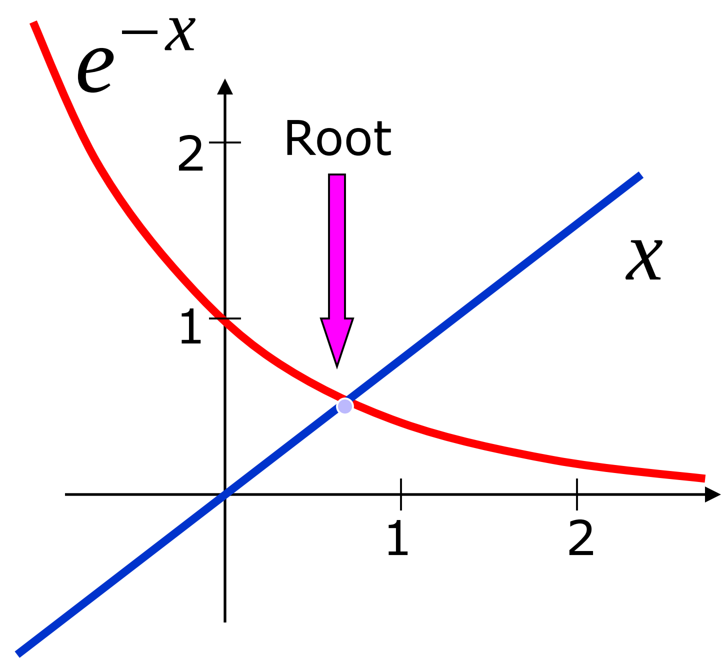
- Graphical methods are useful to provide an initial guess to be used by other methods.

$$Solve$$

$$x = e^{-x}$$

$$The\ \ root \in [0,1]$$

$$root \approx 0.6$$



Numerical Methods

Many methods are available to solve nonlinear equations:

- Bisection Method → bracketing method
- Newton's Method → open method
- Secant Method →open method
- False position Method (bracketing method)
- Fixed point iterations (open method)
- Muller's Method
- Bairstow's Method
- ……….

Bracketing Methods

- In bracketing methods, the method starts with an <u>interval</u> that contains the root and a procedure is used to obtain a smaller interval containing the root.

- Examples of bracketing methods:
  - Bisection method
  - False position method (Regula-Falsi)

Open Methods

- In the open methods, the method starts with one or more initial guess points. In each iteration, a new guess of the root is obtained.
- Open methods are usually more efficient than bracketing methods.
- They may not converge to a root.

Convergence Notation

A sequence $x_1, x_2, \ldots, x_n, \ldots$ is said to **converge** to x if to every $\varepsilon > 0$ there exists N such that:

$$\left| x_n - x \right| < \varepsilon \quad \forall n > N$$

Convergence Notation

Let $x_1, x_2, \ldots.,$ converge to x.

Linear Convergence: $$\frac{\left|x_{n+1} - x\right|}{\left|x_n - x\right|} \le C$$

Quadratic Convergence: $$\frac{\left|x_{n+1} - x\right|}{\left|x_n - x\right|^2} \le C$$

Convergence of order P: $$\frac{\left|x_{n+1} - x\right|}{\left|x_n - x\right|^p} \le C$$

Speed of Convergence

- We can compare different methods in terms of their convergence rate.
- Quadratic convergence is faster than linear convergence.
- A method with convergence order ***q*** converges faster than a method with convergence order ***p*** if ***q>p***.
- Methods of convergence order ***p>1*** are said to have super linear convergence.

Bisection Method

The Bisection Algorithm

Convergence Analysis of Bisection Method

Examples

Introduction

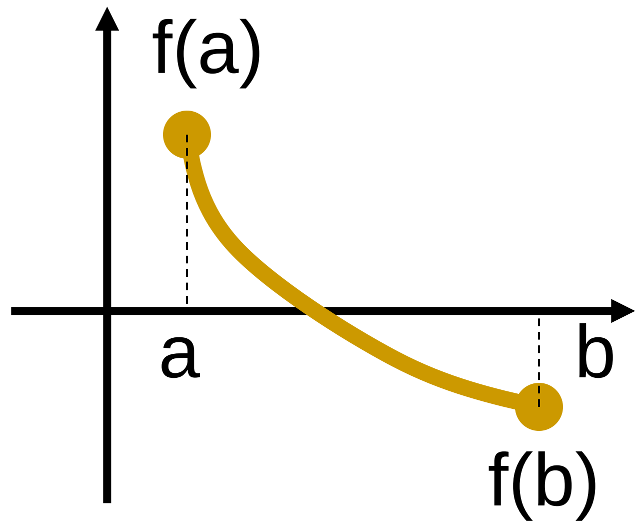
- The **Bisection method** is one of the simplest methods to find a zero of a nonlinear function.
- It is also called **interval halving** method.
- To use the Bisection method, one needs an initial interval that is known to contain a zero of the function.
- The method systematically reduces the interval. It does this by dividing the interval into two equal parts, performs a simple test and based on the result of the test, half of the interval is thrown away.
- The procedure is repeated until the desired interval size is obtained.

Intermediate Value Theorem

- Let f(x) be defined on the interval [a,b].

- Intermediate value theorem:

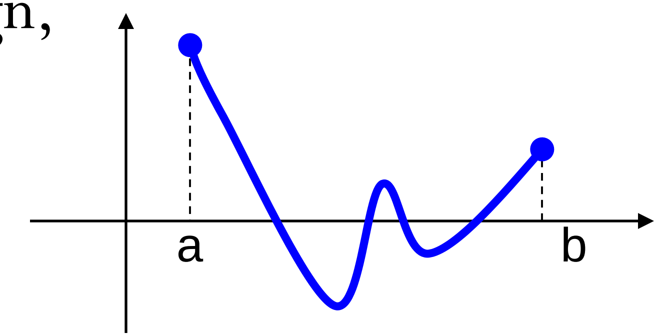
  if a function is continuous and f(a) and f(b) have different signs then the function has at least one zero in the interval [a,b].



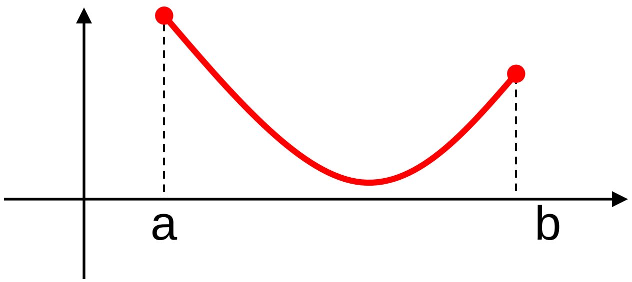
Examples

- If f(a) and f(b) have the same sign, the function may have an even number of real zeros or no real zeros in the interval [a, b].

- Bisection method can not be used in these cases.



The function has four real zeros

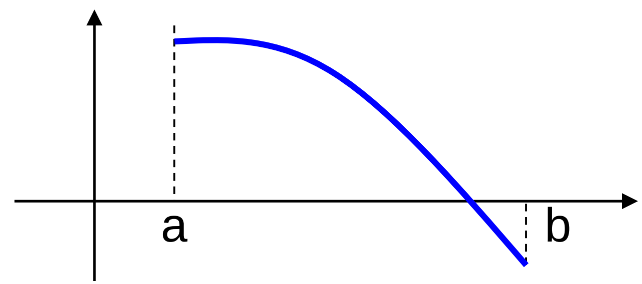


The function has no real zeros

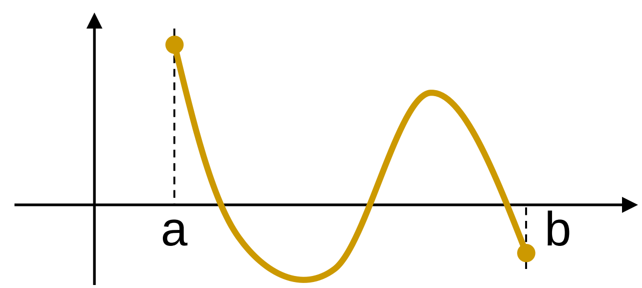
Two More Examples

- If f(a) and f(b) have different signs, the function has at least one real zero.

- Bisection method can be used to find one of the zeros.



The function has one real zero



The function has three real zeros

Bisection Method

- If the function is continuous on [a,b] and f(a) and f(b) have different signs, Bisection method obtains a new interval that is half of the current interval and the sign of the function at the end points of the interval are different.

- This allows us to repeat the Bisection procedure to further reduce the size of the interval.

Bisection Method

Assumptions:

Given an interval [a,b]

f(x) is continuous on [a,b]

f(a) and f(b) have opposite signs.

These assumptions ensure the existence of at least one zero in the interval [a,b] and the bisection method can be used to obtain a smaller interval that contains the zero.

Bisection Algorithm

Assumptions:

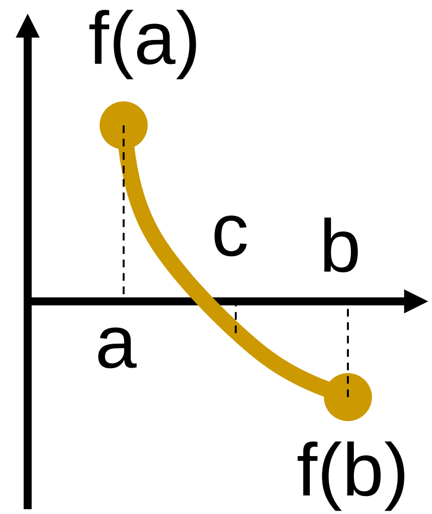
- $f(x)$ is continuous on $[a,b]$
- $f(a)\, f(b) < 0$

Algorithm:

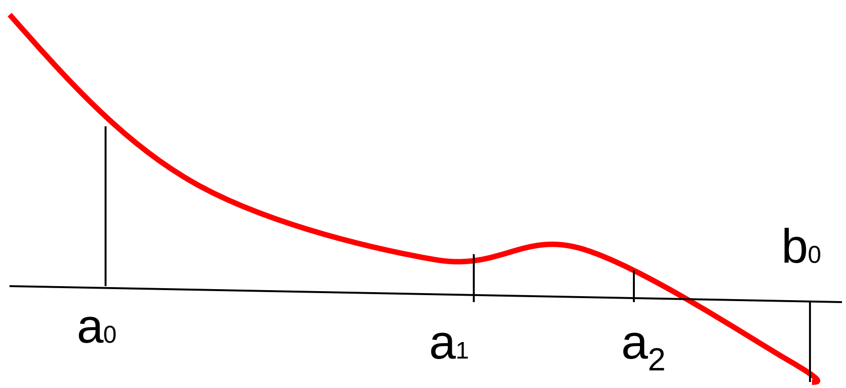
Loop

1. Compute the mid point $c=(a+b)/2$
2. Evaluate $f(c)$
3. If $f(a)\, f(c) < 0$ then new interval $[a, c]$
   If $f(a)\, f(c) > 0$ then new interval $[c, b]$

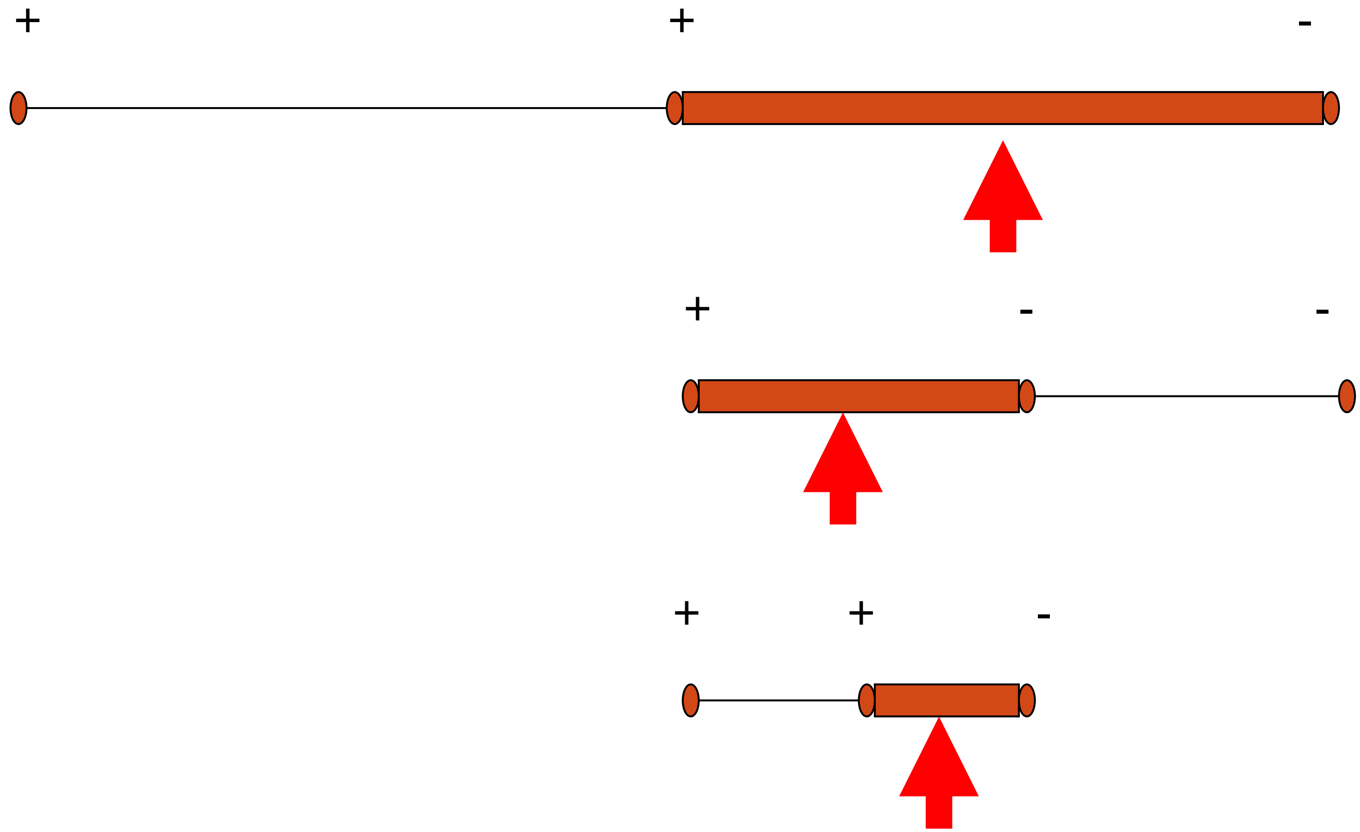
End loop



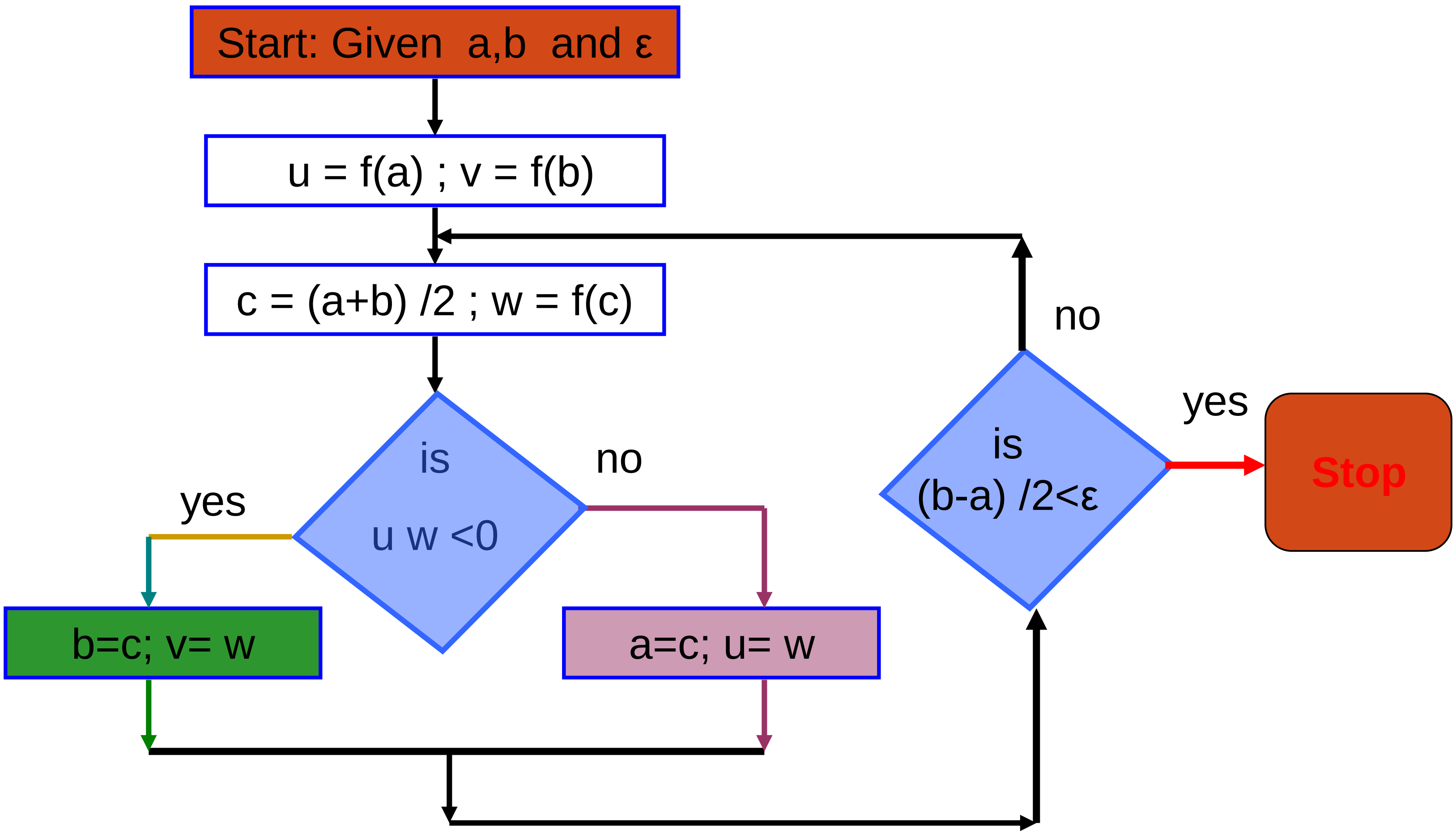
Bisection Method



Example



Flow Chart of Bisection Method



Example

Can you use Bisection method to find a zero of :

$f(x) = x^3 - 3x + 1$ in the interval [0,2]?

Answer:

$f(x)$ is continuous on [0,2]

and $\text{f}(0) * \text{f}(2) = (1)(3) = 3 > 0$

$\Rightarrow$ Assumptions are not satisfied

$\Rightarrow$ Bisection method can not be used

Example

Can you use Bisection method to find a zero of :

$f(x) = x^3 - 3x + 1$ in the interval [0,1]?

Answer:

$f(x)$ is continuous on [0,1]

and $f(0) * f(1) = (1)(-1) = -1 < 0$

$\Rightarrow$ Assumptions are satisfied

$\Rightarrow$ Bisection method can be used

Best Estimate and Error Level

Bisection method obtains an interval that is guaranteed to contain a zero of the function.

Questions:

- What is the best estimate of the zero of $f(x)$?
- What is the error level in the obtained estimate?

Best Estimate and Error Level

The best estimate of the zero of the function ***f(x)*** after the first iteration of the Bisection method is the mid point of the initial interval:

$$\text{Estimate of the zero: } r = \frac{b+a}{2}$$

$$\text{Error} \leq \frac{b-a}{2}$$

Stopping Criteria

Two common stopping criteria

1. Stop after a fixed number of iterations
2. Stop when the absolute error is less than a specified value

How are these criteria related?

Stopping Criteria

c_n : is the midpoint of the interval at the n$^{\text{th}}$ iteration

(c_n is usually used as the estimate of the root).

r: is the zero of the function.

After n iterations :

$$\left|error\right| = \left|r - c_n\right| \leq E_a^n = \frac{b-a}{2^n} = \frac{\Delta x^0}{2^n}$$

Convergence Analysis

Given $f(x)$, a, b, *and* ε

How many iterations are needed such that: $|x - r| \le \varepsilon$

where r is the zero of $f(x)$ and x is the bisection estimate (i.e., $x = c_k$)?

$$\boxed{n \ge \frac{\log(b-a) - \log(\varepsilon)}{\log(2)}}$$

Convergence Analysis – Alternative Form

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

$$n \geq \log_2\left(\frac{\text{width of initial interval}}{\text{desired error}}\right) = \log_2\left(\frac{b-a}{\varepsilon}\right)$$

Example

$$a = 6,\ \ b = 7,\ \varepsilon = 0.0005$$

How many iterations are needed such that: $\left| x - r \right| \leq \varepsilon$?

$$n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1) - \log(0.0005)}{\log(2)} = 10.9658$$

$$\Rightarrow n \geq 11$$

Example

- Use Bisection method to find a root of the equation $x = \cos(x)$ with absolute error <0.02

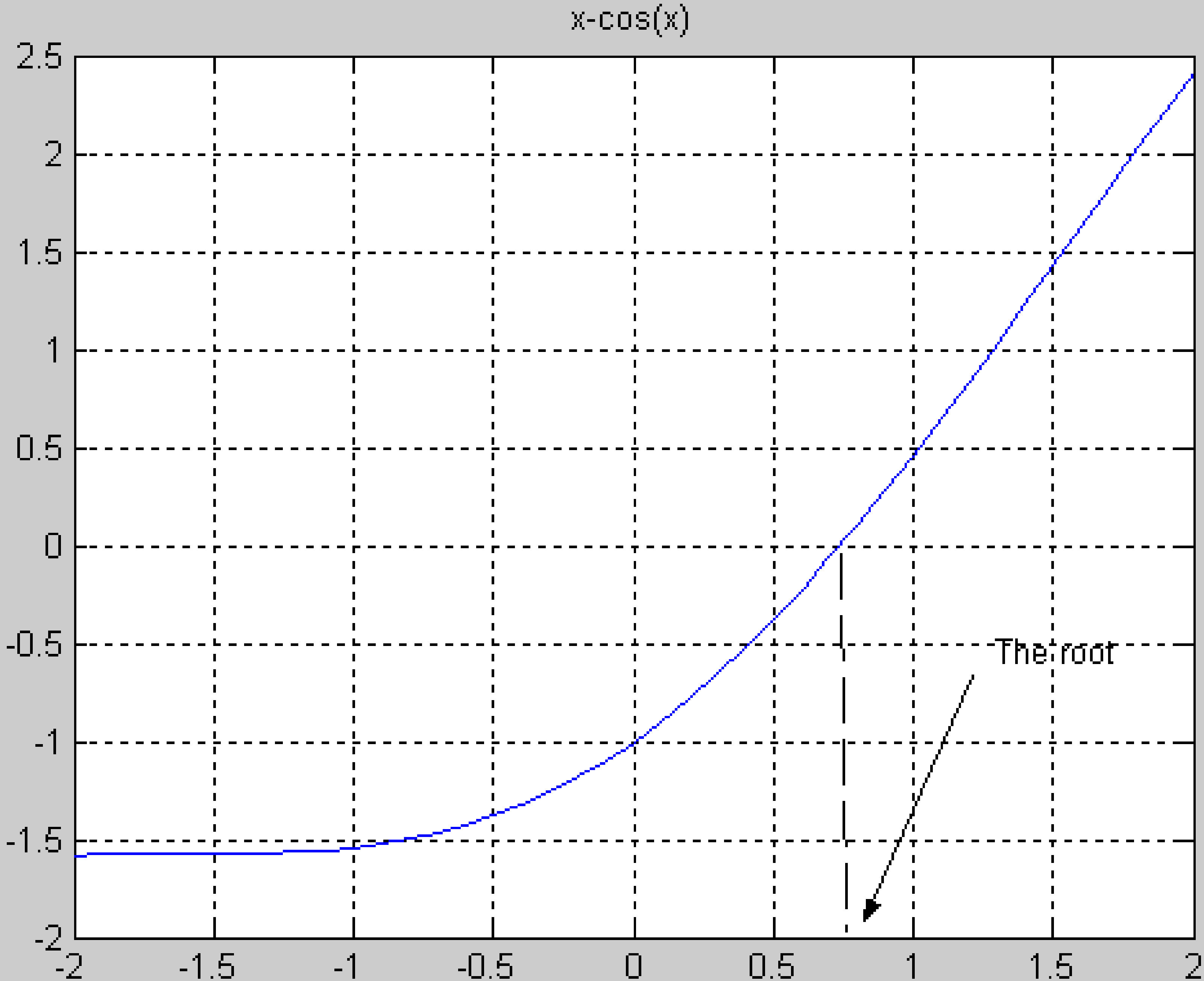
  (assume the initial interval [0.5, 0.9])

Question 1: What is $f(x)$?

Question 2: Are the assumptions satisfied ?

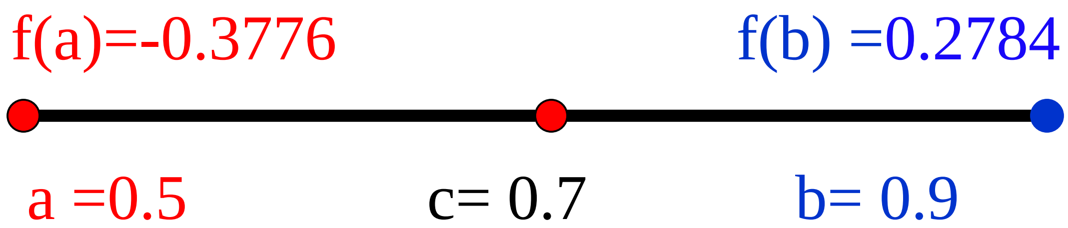
Question 3: How many iterations are needed ?

Question 4: How to compute the new estimate ?


x-cos(x)
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
The root

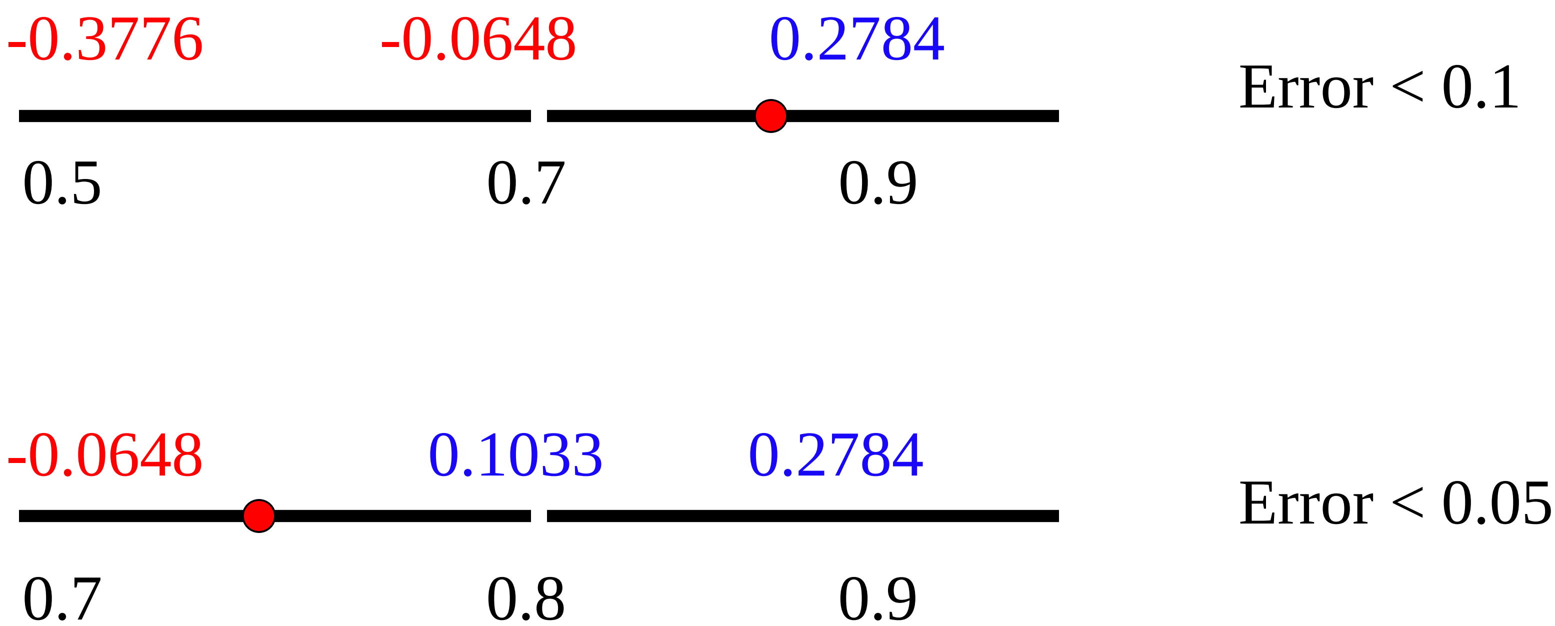
Bisection Method

Initial Interval

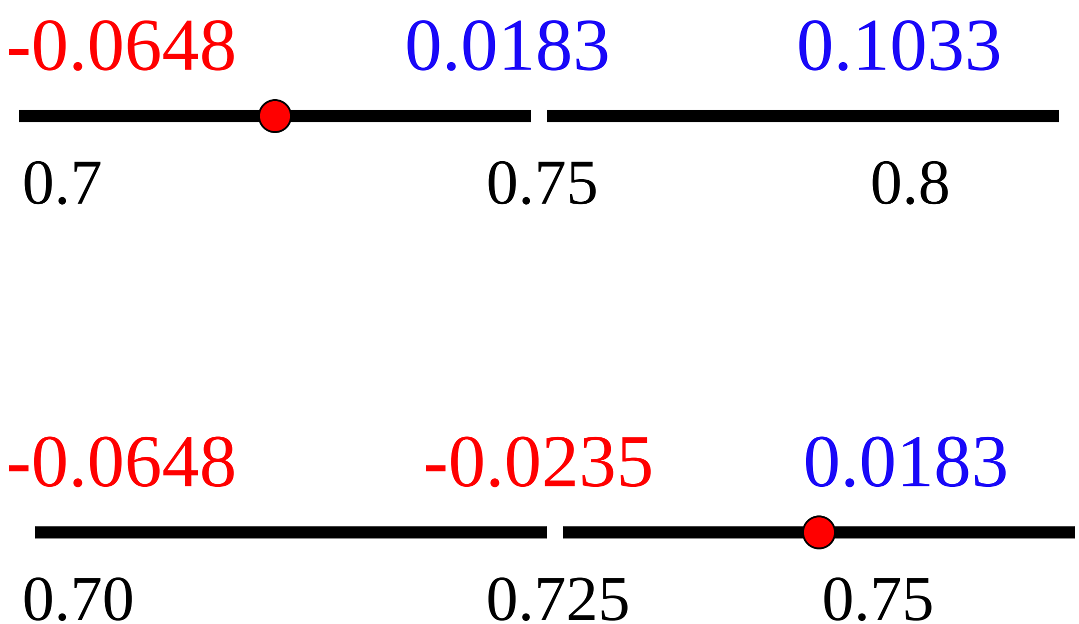


Error < 0.2

Bisection Method



Bisection Method



Error < 0.025

Error < .0125

Summary

- Initial interval containing the root: [0.5,0.9]

- After 5 iterations:
  - Interval containing the root: [0.725, 0.75]
  - Best estimate of the root is 0.7375
  - $|\text{ Error }| < 0.0125$

A Matlab Program of Bisection Method

```
a=.5; b=.9;
 u=a-cos(a);
 v=b-cos(b);
     for i=1:5
          c=(a+b)/2
          fc=c-cos(c)
          if u*fc<0
            b=c ; v=fc;
          else
            a=c; u=fc;
          end
     end
```

```
c =
   0.7000
fc =
  -0.0648
c =
   0.8000
fc =
   0.1033
c =
   0.7500
fc =
   0.0183
c =
   0.7250
fc =
  -0.0235
```

Example

Find the root of:

$$f(x) = x^3 - 3x + 1 \text{ in the interval } :[0,1]$$

* $f(x)$ is continuous

* $f(0) = 1, \; f(1) = -1 \Rightarrow f(a)\, f(b) < 0$

$\Rightarrow$ Bisection method can be used to find the root

Example

Iteration	a	b	c= $\frac{(a+b)}{2}$	f(c)	$\frac{(b-a)}{2}$
1	0	1	0.5	-0.375	0.5
2	0	0.5	0.25	0.266	0.25
3	0.25	0.5	.375	-7.23E-3	0.125
4	0.25	0.375	0.3125	9.30E-2	0.0625
5	0.3125	0.375	0.34375	9.37E-3	0.03125

Bisection Method

Advantages

- Simple and easy to implement
- One function evaluation per iteration
- The size of the interval containing the zero is reduced by 50% after each iteration
- The number of iterations can be determined a priori
- No knowledge of the derivative is needed
- The function does not have to be differentiable

Disadvantage

- Slow to converge
- Good intermediate approximations may be discarded

Bisection Method (as C function)

```
double Bisect(double xl, double xu, double es,
              int iter_max)
{
  double xr;     // Est. Root
  double xr_old; // Est. root in the previous step
  double ea;     // Est. error
  int iter = 0;  // Keep track of # of iterations
  double fl, fr; // Save values of f(xl) and f(xr)

  xr = xl;       // Initialize xr in order to
                 // calculating "ea". Can also be "xu".
  fl = f(xl);
  do {
    iter++;
    xr_old = xr;

    xr = (xl + xu) / 2;      // Estimate root
    fr = f(xr);
```

```
    if (xr != 0)
      ea = fabs((xr - xr_old) / xr) * 100;

    test = fl * fr;

    if (test < 0)
      xu = xr;
    else
    if (test > 0) {
      xl = xr;
      fl = fr;
    }
    else
      ea = 0;

  } while (ea > es && iter < iter_max);

  return xr;
}
```

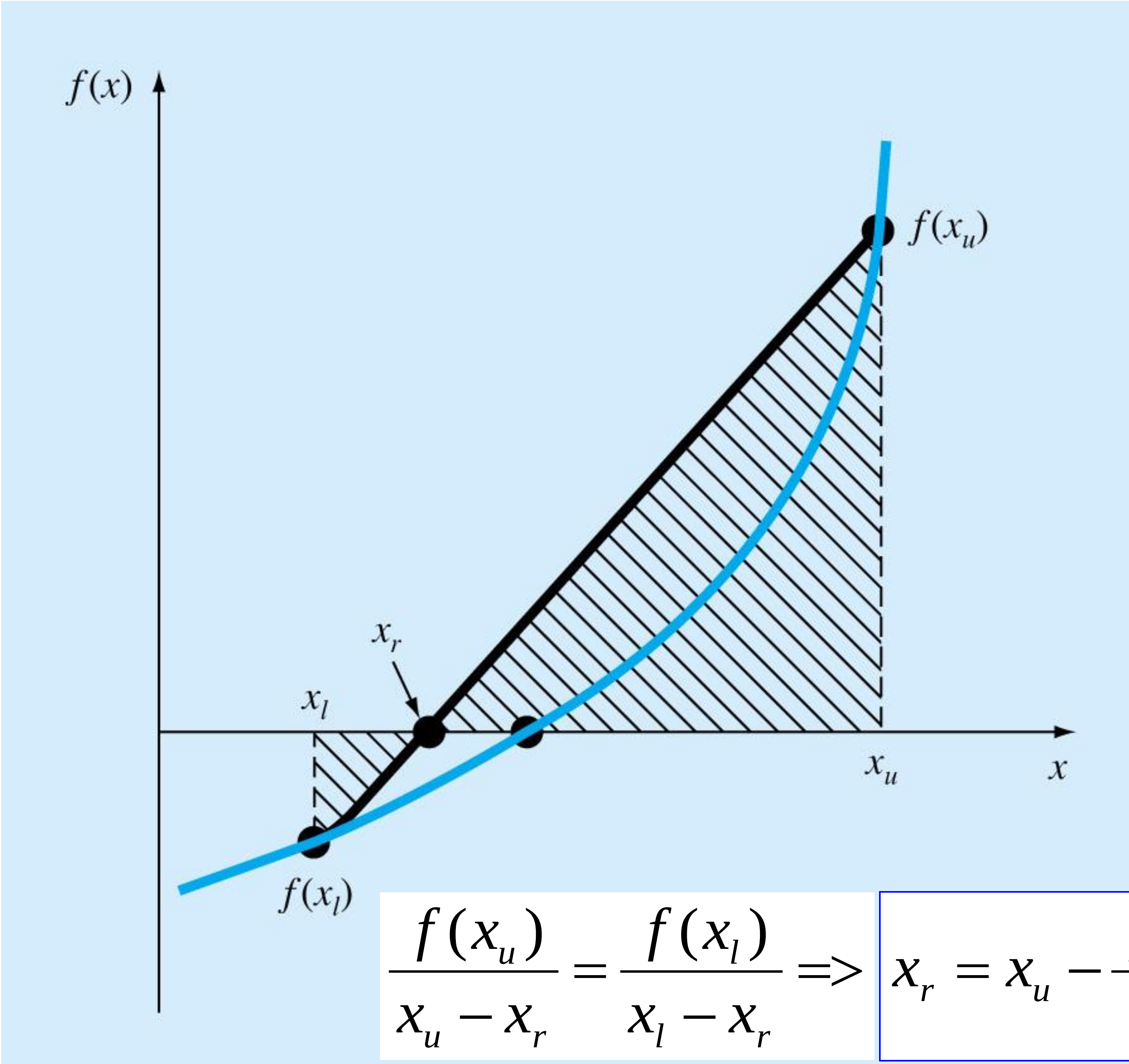
Regula – Falsi Method

Regula Falsi Method

- Also known as the false-position method, or linear interpolation method.

- Unlike the bisection method which divides the search interval by half, *regula falsi* interpolates $f(x_u)$ and $f(x_l)$ by a straight line and the intersection of this line with the x-axis is used as the new search position.

- The slope of the line connecting $f(x_u)$ and $f(x_l)$ represents the "average slope" (i.e., the value of $f'(x)$) of the points in $[x_l, x_u]$.



$$\frac{f(x_u)}{x_u - x_r} = \frac{f(x_l)}{x_l - x_r} \Rightarrow \boxed{x_r = x_u - \frac{f(x_u)(x_l - x_u)}{f(x_l) - f(x_u)}}$$

Regula Falsi Method

- The *regula falsi* method start with two point, (a, f(a)) and (b,f(b)), satisfying the condition that f(a)f(b)<0.
- The straight line through the two points (a, f(a)), (b, f(b)) is

$$y = f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

- The next approximation to the zero is the value of x where the straight line through the initial points crosses the x-axis.

$$x = a - \frac{b - a}{f(b) - f(a)} f(a) = \frac{af(b) - bf(a)}{f(b) - f(a)}$$

Example

- **Finding the Cube Root of 2 Using Regula Falsi**

- Since f(1)= -1, f(2)=6, we take as our starting bounds on the zero a=1 and b=2.
- Our first approximation to the zero is

$$x = b - \frac{b-a}{f(b)-f(a)}(f(b)) = 2 - \frac{2-1}{6+1}(6)$$

$$= 2 - 6/7 = 8/7 \approx 1.1429$$

- We then find the value of the function:

$$y = f(x) = (8/7)^3 - 2 \approx -0.5073$$

- Since f(a) and y are both negative, but y and f(b) have opposite signs

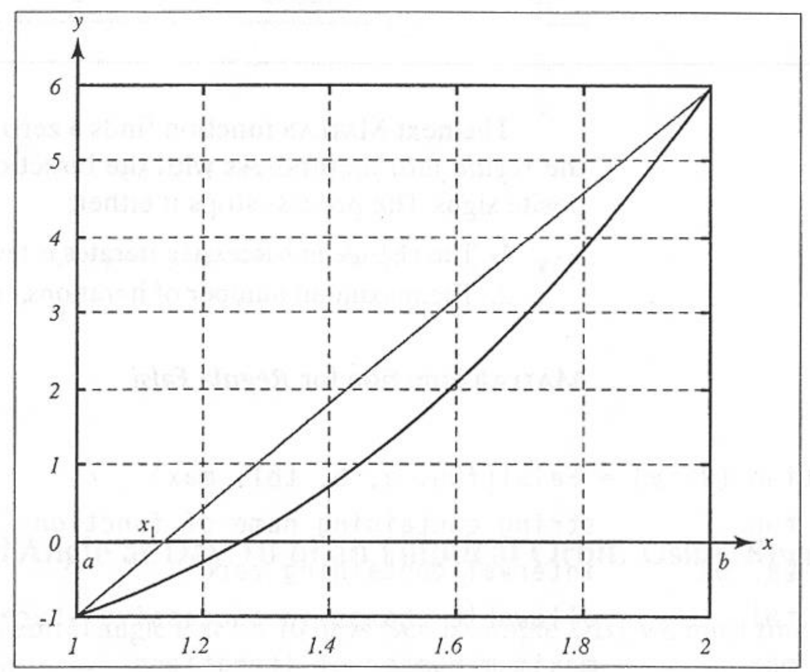


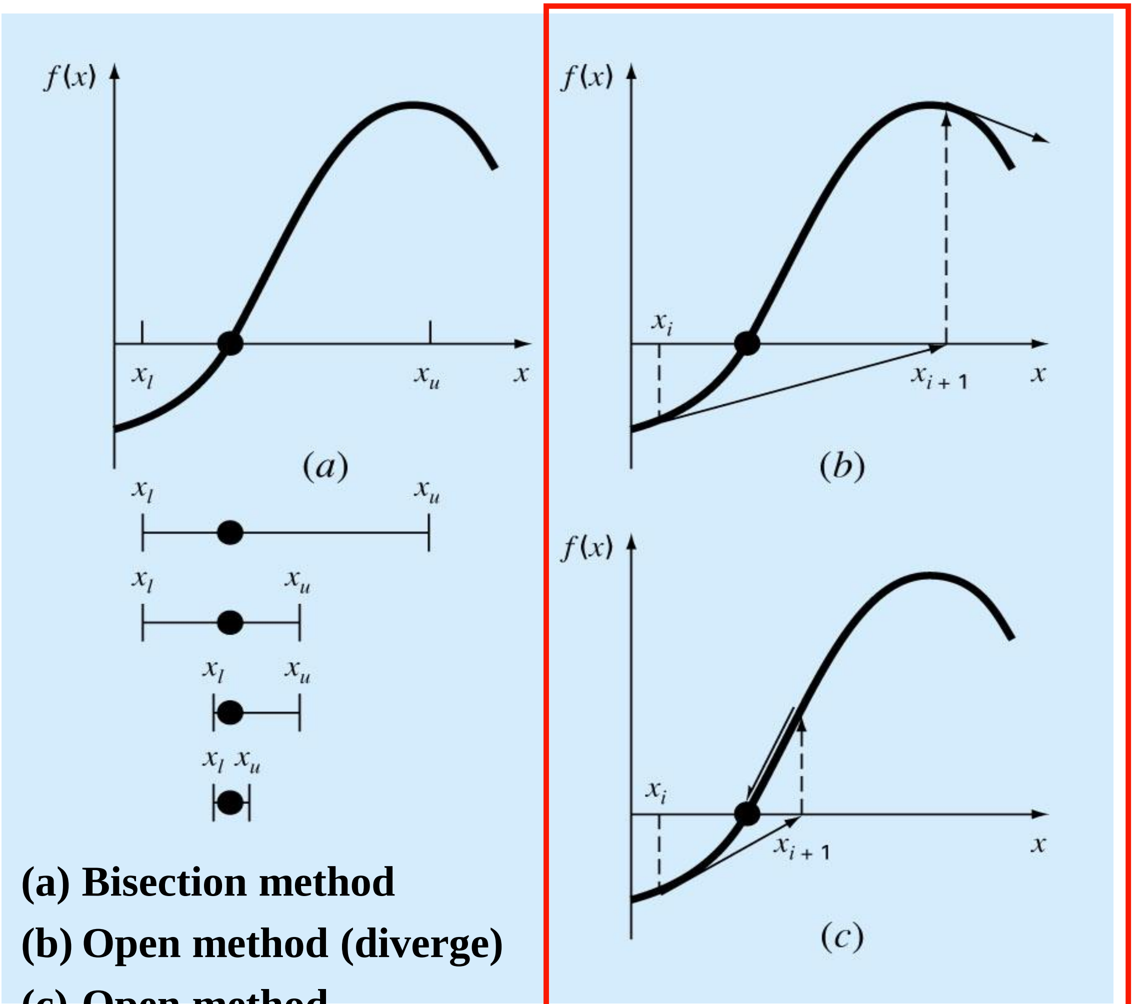
FIGURE 2.5 Graph of $y = x^3 - 2$ and approximation line on the interval [1, 2].

Example (cont.)

- Calculation of $\sqrt[3]{2}$ using *regula falsi*.

Step	a	b	x	y
1	1	2	1.1429	−0.50729
2	1.1429	2	1.2097	−0.22986
3	1.2097	2	1.2388	−0.098736
4	1.2388	2	1.2512	−0.041433
5	1.2512	2	1.2563	−0.017216
6	1.2563	2	1.2584	−0.0071239
7	1.2584	2	1.2593	−0.0029429
8	1.2593	2	1.2597	−0.0012148
9	1.2597	2	1.2598	−0.00050134
10	1.2598	2	1.2599	−0.00020687

Open Methods



(a) Bisection method

(b) Open method (diverge)

(c) Open method (converge)

To find the root for $f(x) = 0$, we construct a magic formulae

$$x_{i+1} = g(x_i)$$

to predict the root iteratively until x converge to a root. However, <u>x may diverge</u>!

What you should know about Open Methods

- How to construct the magic formulae $g(x)$?
- How can we ensure convergence?
- What makes a method converges quickly or diverge?
- How fast does a method converge?

OPEN METHOD
Newton-Raphson Method

Assumptions
Interpretation
Examples
Convergence Analysis

Newton-Raphson Method

(Also known as Newton's Method)

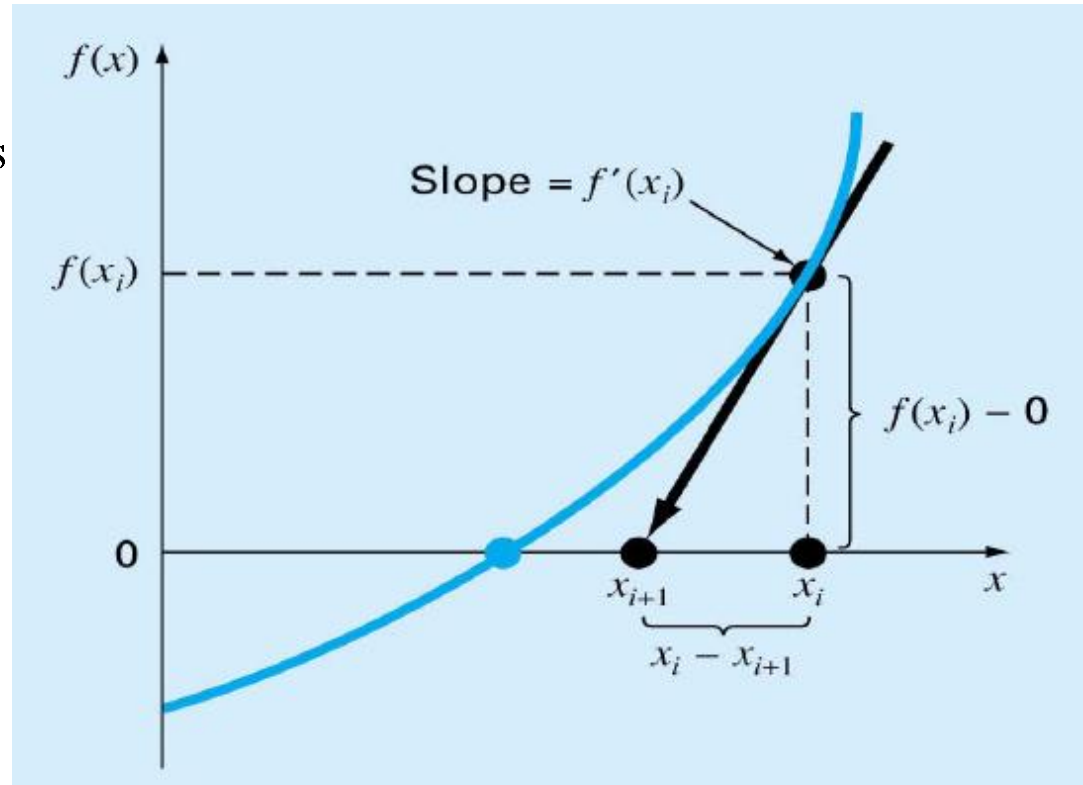
Given an initial guess of the root x_0, Newton-Raphson method uses information about the function and its derivative at that point to find a better guess of the root.

Assumptions:

- $f(x)$ is continuous and the first derivative is known
- An initial guess x_0 such that $f'(x_0) \neq 0$ is given

Newton Raphson Method
- Graphical Depiction -

- If the initial guess at the root is x_i, then a tangent to the function of x_i that is $f'(x_i)$ is extrapolated down to the *x-axis* to provide an estimate of the root at x_{i+1}.



Derivation of Newton's Method

$Given: x_i$ an initial guess of the root of $f(x)=0$

$Question:$ How do we obtain a better estimate x_{i+1}?

Taylor Therorem: $f(x+h) \approx f(x) + f'(x)h$

Find h such that $f(x+h)=0$.

$$\Rightarrow h \approx -\frac{f(x)}{f'(x)}$$

Newton – Raphson Formula

A new guess of the root: $x_{i+1} = x_i - \dfrac{f(x_i)}{f'(x_i)}$

Newton's Method

$$Given \quad f(x),\ f'(x),\ x_0$$

$$Assumption \quad f'(x_0) \neq 0$$

$$for \quad i = 0 : n$$

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

$$end$$

```
C    FORTRAN PROGRAM
      F(X) = X**3 - 3*X**2 + 1
      FP(X) = 3*X**2 - 6*X
      X = 4
      DO 10 I = 1,5
       X = X - F(X)/FP(X)
       PRINT *, X
10     CONTINUE
      STOP
      END
```

Newton's Method

$$Given \quad f(x),\ f'(x),\ x_0$$
$$Assumption \quad f'(x_0) \neq 0$$

$$for \quad i = 0:n$$
$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$
$$end$$

F.m

$$function \quad [F] = F(X)$$
$$F = X\,\hat{}\,3 - 3 * X\,\hat{}\,2 + 1$$

FP.m

$$function \quad [FP] = FP(X)$$
$$FP = 3 * X\,\hat{}\,2 - 6 * X$$

```
% MATLAB PROGRAM
X = 4
for i = 1:5
    X = X - F(X)/FP(X)
end
```

Example

Find a zero of the function $f(x) = x^3 - 2x^2 + x - 3$, $x_0 = 4$

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

Iteration 1: $x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)} = 4 - \dfrac{33}{33} = 3$

Iteration 2: $x_2 = x_1 - \dfrac{f(x_1)}{f'(x_1)} = 3 - \dfrac{9}{16} = 2.4375$

Iteration 3: $x_3 = x_2 - \dfrac{f(x_2)}{f'(x_2)} = 2.4375 - \dfrac{2.0369}{9.0742} = 2.2130$

Example

k (Iteration)	x_k	$f(x_k)$	$f'(x_k)$	x_{k+1}	$\lvert x_{k+1} - x_k \rvert$
0	**4**	**33**	**33**	**3**	1
1	**3**	**9**	**16**	**2.4375**	0.5625
2	**2.4375**	**2.0369**	**9.0742**	**2.2130**	0.2245
3	**2.2130**	**0.2564**	**6.8404**	**2.1756**	0.0384
4	**2.1756**	**0.0065**	**6.4969**	**2.1746**	0.0010

Convergence Analysis

Theorem:

Let $f(x), f'(x)$ *and* $f''(x)$ be continuous at $\mathrm{x} \approx \mathrm{r}$

where $f(r) = 0$. *If* $f'(r) \neq 0$ then there exists $\delta > 0$

such that $\left|x_0 - r\right| \leq \delta \Rightarrow \dfrac{\left|x_{k+1} - r\right|}{\left|x_k - r\right|^2} \leq C$

$$C = \frac{1}{2} \frac{\max\limits_{\left|x_0 - r\right| \leq \delta} \left|f''(x)\right|}{\min\limits_{\left|x_0 - r\right| \leq \delta} \left|f'(x)\right|}$$

Convergence Analysis
Remarks

When the guess is close enough to a simple root of the function then Newton's method is guaranteed to converge quadratically.

Quadratic convergence means that the number of correct digits is nearly doubled at each iteration.

Error Analysis of Newton-Raphson Method

Using an iterative process we get x_{k+1} from x_k and other info.

We have $x_0, x_1, x_2, \ldots, x_{k+1}$ as the estimation for the root α.

Let $\delta_k = \alpha - x_k$

Error Analysis of Newton-Raphson Method

By definition

$$\delta_i = \alpha - x_i \qquad (1)$$

$$\delta_{i+1} = \alpha - x_{i+1} \qquad (2)$$

Newton-Raphson method

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

$$\Rightarrow f(x_i) = f'(x_i)(x_i - x_{i+1})$$

$$\Rightarrow f(x_i) + f'(x_i)(-x_i) = f'(x_i)(-x_{i+1})$$

$$\Rightarrow f(x_i) + f'(x_i)(\alpha - x_i) = f'(x_i)(\alpha - x_{i+1}) \qquad (3)$$

Error Analysis of Newton-Raphson Method

Suppose α is the true value (i.e., $f(\alpha) = 0$).

Using Taylor's series

$$f(\alpha) = f(x_i) + f'(x_i)(\alpha - x_i) + \frac{f''(c)}{2}(\alpha - x_i)^2$$

$$\Rightarrow 0 = f(x_i) + f'(x_i)(\alpha - x_i) + \frac{f''(c)}{2}(\alpha - x_i)^2$$

$$\Rightarrow 0 = f'(x_i)(\alpha - x_{i+1}) + \frac{f''(c)}{2}(\alpha - x_i)^2 \qquad \text{(from (3))}$$

$$\Rightarrow 0 = f'(x_i)(\delta_{i+1}) + \frac{f''(c)}{2}(\delta_i)^2 \qquad \text{(from (1) and (2))}$$

$$\Rightarrow \delta_{i+1} = \frac{-f''(c)}{2f'(x_i)}\delta_i^2 \cong \frac{-f''(\alpha)}{2f'(\alpha)}\delta_i^2$$

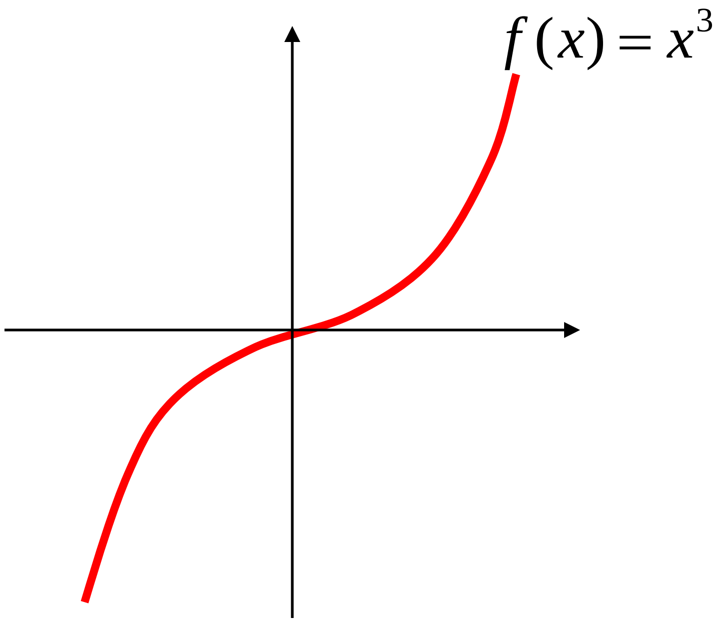
When x_i and α are very close to each other, c is between x_i and α.

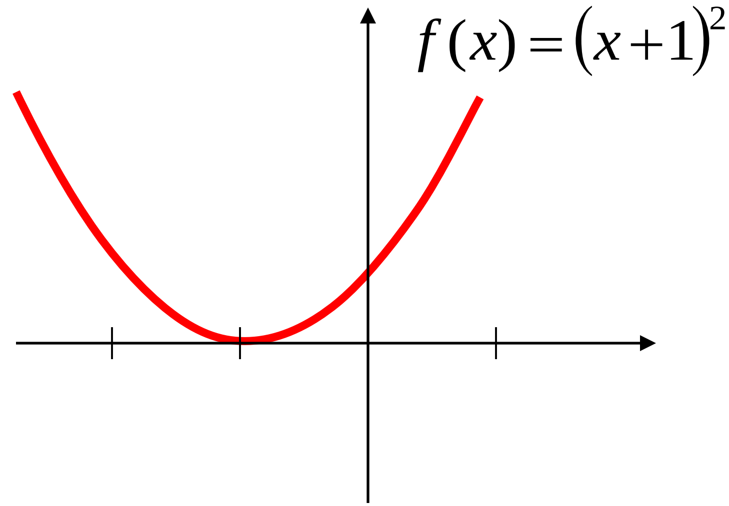
The iterative process is said to be of second order.

Problems with Newton's Method

- If the initial guess of the root is far from the root the method may not converge.
- Newton's method converges linearly near multiple zeros { $f(r) = f'(r) = 0$ }. In such a case, modified algorithms can be used to regain the quadratic convergence.

Multiple Roots

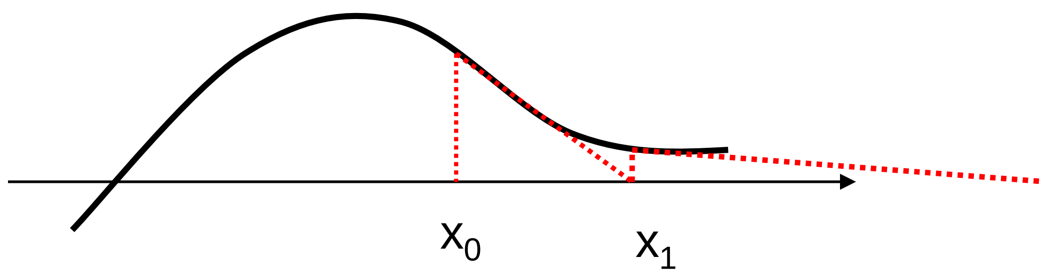




f(x) has three zeros at x = 0

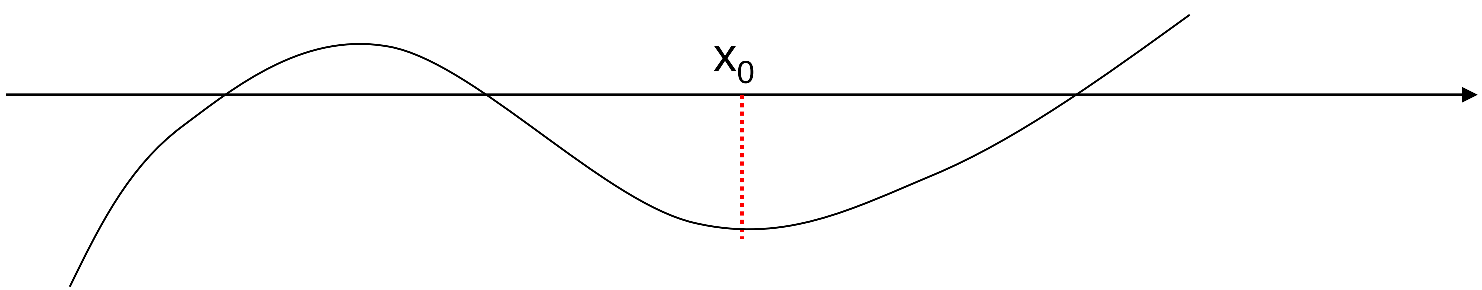
f(x) has two zeros at x = -1

Problems with Newton's Method
- Runaway -



The estimates of the root is going away from the root.

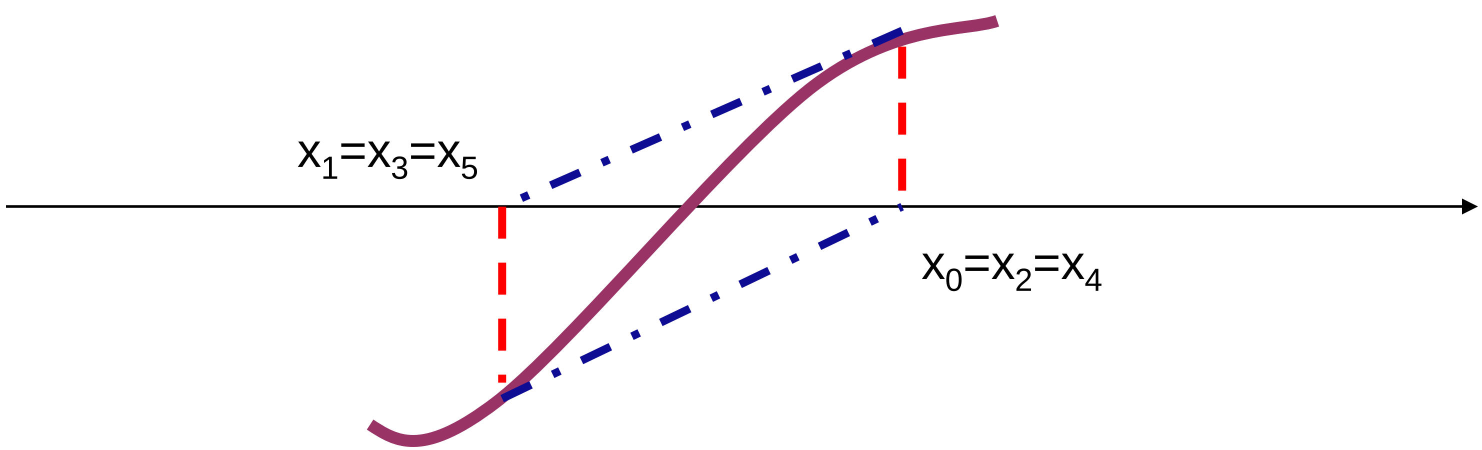
Problems with Newton's Method
- Flat Spot -



The value of $f'(x)$ is zero, the algorithm fails.

If $f'(x)$ is very small then x_1 will be very far from x_0.

Problems with Newton's Method
- Cycle -



The algorithm cycles between two values x_0 and x_1

Secant Method

- Secant Method
- Examples
- Convergence Analysis

Newton's Method (Review)

$$Assumptions: \quad f(x),\ f'(x),\ x_0 \quad are\ available,$$

$$f'(x_0) \neq 0$$

$$Newton's\ \ Method\ \ new\ estimate:$$

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

Problem:

$f'(x_i)$ is not available,

or difficult to obtain analytically.

The Secant Method

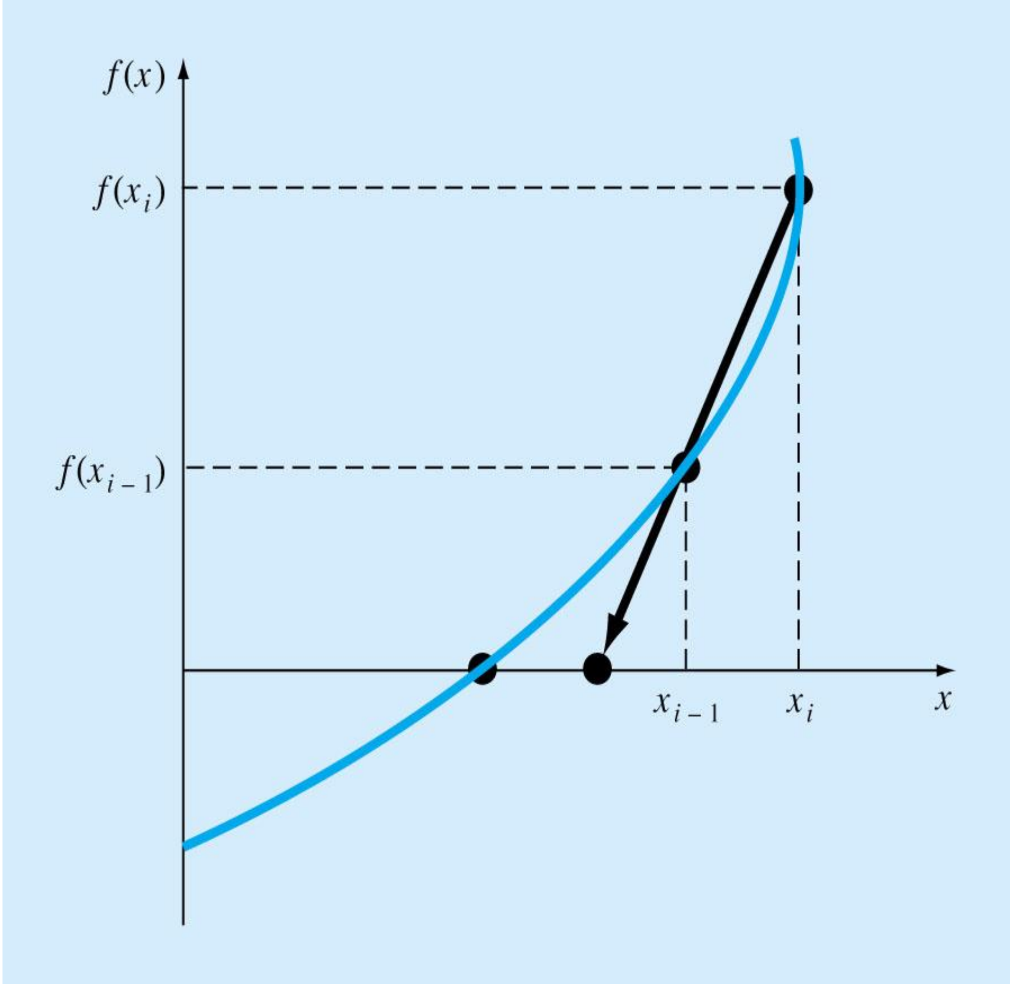
$$x_{i+1} = x_i - f(x_i)\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})}$$

- Requires two initial estimates x_o, x_1.

  However, it is not a "bracketing" method.

- *The Secant Method* has the same properties as *Newton*'s method.

  Convergence is not guaranteed for all x_o, f(x).



Secant Method

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

if x_i *and* x_{i-1} are two initial points :

$$f'(x_i) = \frac{f(x_i) - f(x_{i-1})}{(x_i - x_{i-1})}$$

$$\boxed{x_{i+1} = x_i - \frac{f(x_i)}{\dfrac{f(x_i) - f(x_{i-1})}{(x_i - x_{i-1})}} = \quad x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}}$$

- **Notice that this is very similar to the false position method in form**
- **Still requires two initial estimates**
- **This method requires two initial estimates of *x* but does not require an analytical expression of the derivative.**
- **But it doesn't bracket the root at all times - there is no sign test**

Newton-Raphson: $$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$$

$$f'(x_i) = \frac{df}{dx} \cong \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}}$$

Secant: $$x_{i+1} = x_i - f(x_i)\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})} \qquad i = 1, 2, 3, \ldots$$

Secant Method

Assumptions:

Two initial points x_i *and* x_{i-1}

$such\ that\ f(x_i) \neq f(x_{i-1})$

New estimate (Secant Method):

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Secant Method

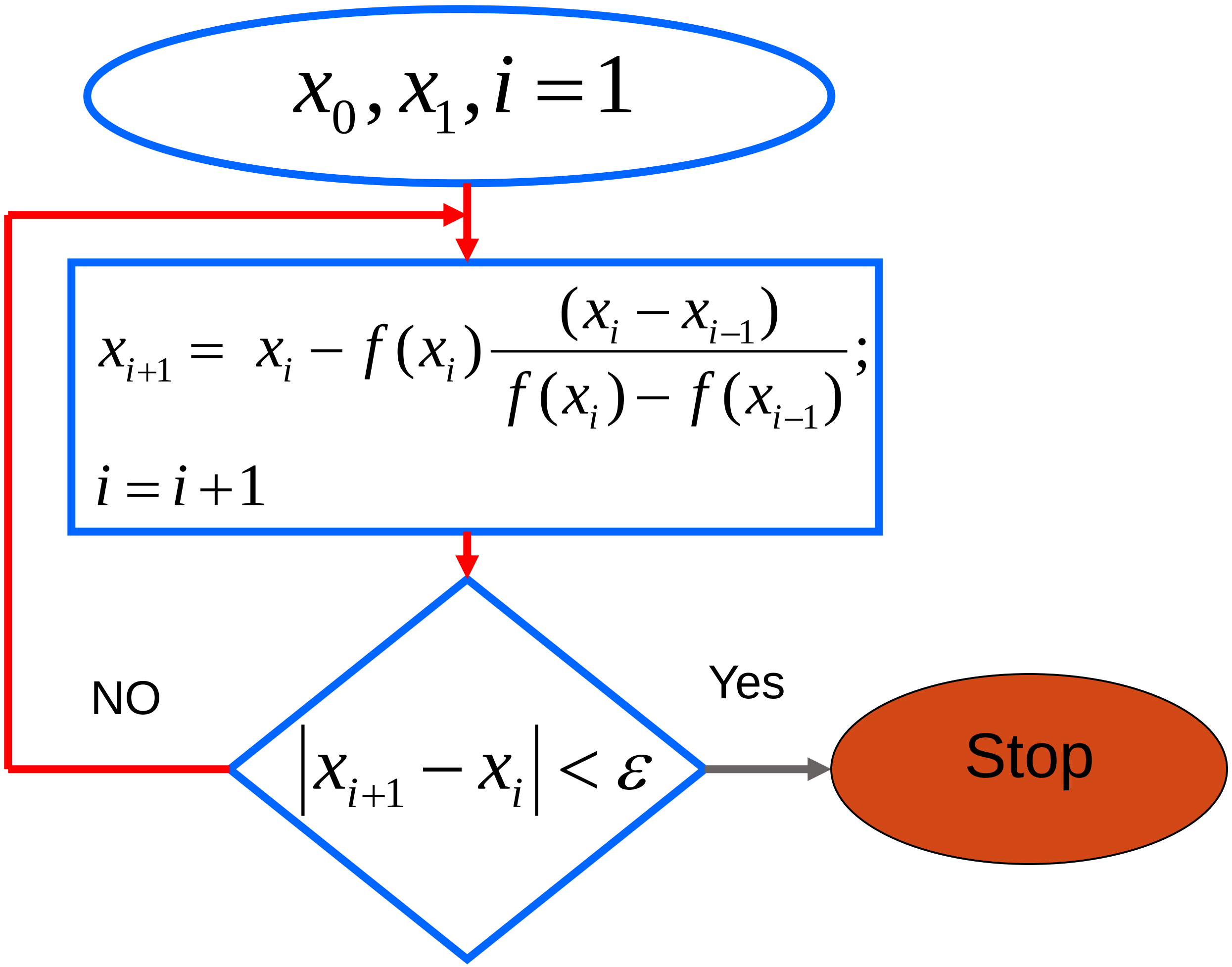
$$f(x) = x^2 - 2x + 0.5$$

$$x_0 = 0$$

$$x_1 = 1$$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Secant Method - Flowchart



Modified Secant Method

In this modified Secant method, only one initial guess is needed :

$$f'(x_i) \approx \frac{f(x_i + \delta\, x_i) - f(x_i)}{\delta\, x_i}$$

$$x_{i+1} = x_i - \frac{f(x_i)}{\dfrac{f(x_i + \delta\ x_i) - f(x_i)}{\delta\ x_i}} = \quad x_i - \frac{\delta\, x_i\, f(x_i)}{f(x_i + \delta\ x_i) - f(x_i)}$$

Problem : How to select δ ?

If not selected properly, the method may diverge .

Example

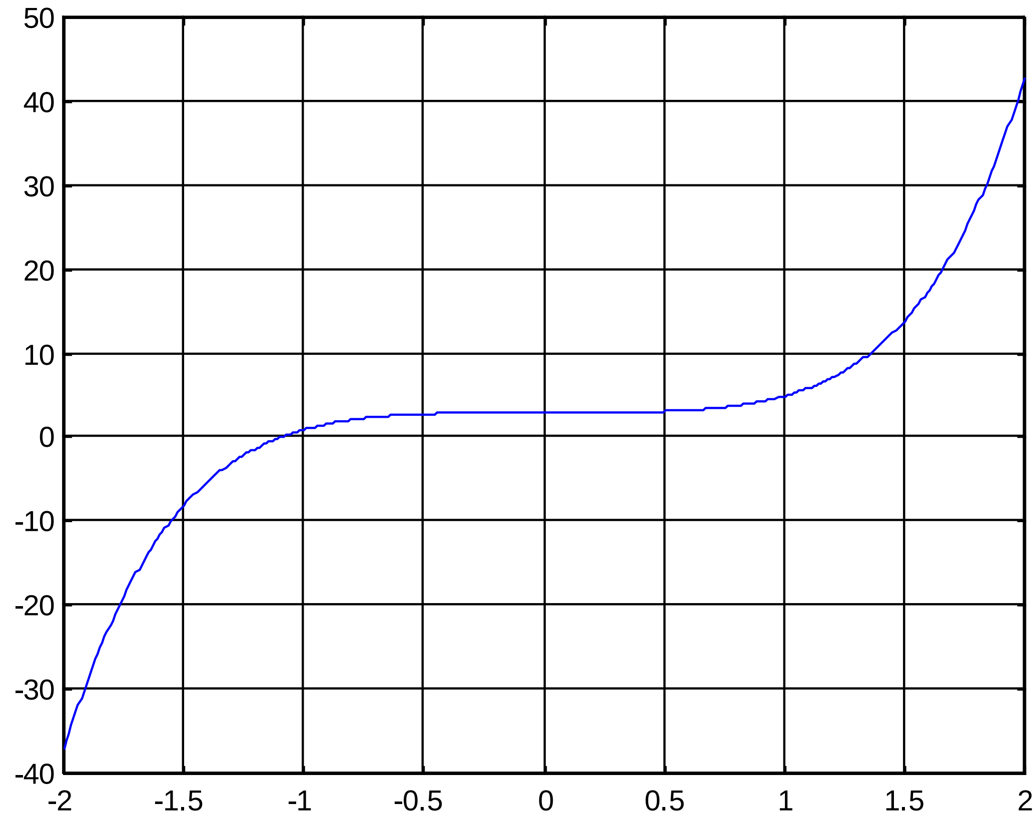
Find the roots of :

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

Initial points

$$x_0 = -1 \; and \; x_1 = -1$$

$$with \;\; error < 0.001$$



Example

x(i)	f(x(i))	x(i+1)	|x(i+1)-x(i)|
-1.0000	1.0000	-1.1000	0.1000
-1.1000	0.0585	-1.1062	0. 0062
-1.1062	0.0102	-1.1052	0.0009
-1.1052	0.0001	-1.1052	0.0000

Convergence Analysis

- The rate of convergence of the Secant method is super linear:

$$\frac{|x_{i+1} - r|}{|x_i - r|^{\alpha}} \le C, \qquad \alpha \approx 1.62$$

$r : root \qquad x_i :$ estimate of the root at the i^{th} iteration.

- It is better than Bisection method but not as good as Newton's method.

OPEN METHOD

Fixed Point Iteration

Fixed Point Iteration

- Also known as **one-point iteration** or **successive substitution**

- To find the root for $f(x) = 0$, we reformulate $f(x) = 0$ so that there is an x on one side of the equation.

$$f(x) = 0 \quad \Leftrightarrow \quad g(x) = x$$

- If we can solve $g(x) = x$, we solve $f(x) = 0$.
  - x is known as the fixed point of $g(x)$.
- We solve $g(x) = x$ by computing

$$x_{i+1} = g(x_i) \quad \text{with } x_0 \text{ given}$$

until x_{i+1} converges to x.

Fixed Point Iteration – Example

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

$$x^2 + 2x - 3 = 0 \quad \Rightarrow \quad 2x = 3 - x^2 \quad \Rightarrow \quad x = \frac{3 - x^2}{2}$$

$$\Rightarrow x_{i+1} = g(x_i) = \frac{3 - x_i^2}{2}$$

Reason: If x converges, i.e. $x_{i+1} \rightarrow x_i$

$$x_{i+1} = \frac{3 - x_i^2}{2} \rightarrow x_i = \frac{3 - x_i^2}{2}$$

$$\Rightarrow x_i^2 + 2x_i - 3 = 0$$

Example

Find root of $f(x) = e^{-x} - x = 0$.

(Answer: $\alpha = 0.56714329$)

We put $x_{i+1} = e^{-x_i}$

i	x_i	ε_a (%)	ε_t (%)
0	0		100.0
1	1.000000	100.0	76.3
2	0.367879	171.8	35.1
3	0.692201	46.9	22.1
4	0.500473	38.3	11.8
5	0.606244	17.4	6.89
6	0.545396	11.2	3.83
7	0.579612	5.90	2.20
8	0.560115	3.48	1.24
9	0.571143	1.93	0.705
10	0.564879	1.11	0.399

- There are infinite ways to construct $g(x)$ from $f(x)$.

For example, $f(x) = x^2 - 2x - 3 = 0$ (ans: $x = 3$ or -1)

Case a:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x^2 = 2x + 3$$
$$\Rightarrow x = \sqrt{2x+3}$$
$$\Rightarrow g(x) = \sqrt{2x+3}$$

Case b:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x(x-2) - 3 = 0$$
$$\Rightarrow x = \frac{3}{x-2}$$
$$\Rightarrow g(x) = \frac{3}{x-2}$$

Case c:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow 2x = x^2 - 3$$
$$\Rightarrow x = \frac{x^2-3}{2}$$
$$\Rightarrow g(x) = \frac{x^2-3}{2}$$

So which one is better?

Case a

$$x_{i+1} = \sqrt{2x_i + 3}$$

1. $x_0 = 4$
2. $x_1 = 3.31662$
3. $x_2 = 3.10375$
4. $x_3 = 3.03439$
5. $x_4 = 3.01144$
6. $x_5 = 3.00381$

Converge!

Case b

$$x_{i+1} = \frac{3}{x_i - 2}$$

1. $x_0 = 4$
2. $x_1 = 1.5$
3. $x_2 = -6$
4. $x_3 = -0.375$
5. $x_4 = -1.263158$
6. $x_5 = -0.919355$
7. $x_6 = -1.02762$
8. $x_7 = -0.990876$
9. $x_8 = -1.00305$

Converge, but slower

Case c

$$x_{i+1} = \frac{x_i^2 - 3}{2}$$

1. $x_0 = 4$
2. $x_1 = 6.5$
3. $x_2 = 19.625$
4. $x_3 = 191.070$

Diverge!

Fixed Point Iteration Impl. (as C function)

```c
// x0: Initial guess of the root
// es: Acceptable relative percentage error
// iter_max: Maximum number of iterations allowed
double FixedPt(double x0, double es, int iter_max) {
  double xr = x0;  // Estimated root
  double xr_old;   // Keep xr from previous iteration
  int iter = 0;    // Keep track of # of iterations

  do {
    xr_old = xr;
    xr = g(xr_old);  // g(x) has to be supplied
    if (xr != 0)
      ea = fabs((xr - xr_old) / xr) * 100;

    iter++;
  } while (ea > es && iter < iter_max);

  return xr;
}
```

Comparison of Root Finding Methods

- Advantages/disadvantages
- Examples

Summary

Method	Pros	Cons
Bisection	- Easy, Reliable, Convergent - One function evaluation per iteration - No knowledge of derivative is needed	- Slow - Needs an interval [a,b] containing the root, i.e., $f(a)f(b)<0$
Newton	- Fast (if near the root) - Two function evaluations per iteration	- May diverge - Needs derivative and an initial guess x0 such that $f'(x0)$ is nonzero
Secant	- Fast (slower than Newton) - One function evaluation per iteration - No knowledge of derivative is needed	- May diverge - Needs two initial points guess x0, x1 such that $f(x0)- f(x1)$ is nonzero

Example

Use Secant method to find the root of :

$$f(x) = x^6 - x - 1$$

Two initial points $x_0 = 1$ *and* $x_1 = 1.5$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Solution

k	x_k	$f(x_k)$
0	1.0000	-1.0000
1	1.5000	8.8906
2	1.0506	-0.7062
3	1.0836	-0.4645
4	1.1472	0.1321
5	1.1331	-0.0165
6	1.1347	-0.0005

Example

Use Newton's Method to find a root of:

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

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $\left|x_{k+1} - x_k\right| < 0.001$, or

if $\left|f(x_k)\right| < 0.0001$.

Five Iterations of the Solution

k	x_k	$f(x_k)$	$f'(x_k)$	ERROR
0	1.0000	-1.0000	2.0000	
1	1.5000	0.8750	5.7500	0.1522
2	1.3478	0.1007	4.4499	0.0226
3	1.3252	0.0021	4.2685	0.0005
4	1.3247	0.0000	4.2646	0.0000
5	1.3247	0.0000	4.2646	0.0000

Example

Use Newton's Method to find a root of:

$f(x) = e^{-x} - x$

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

Example

Use Newton's Method to find a root of:

$$f(x) = e^{-x} - x, \qquad f'(x) = -e^{-x} - 1$$

x_k	$f(x_k)$	$f'(x_k)$	$\dfrac{f(x_k)}{f'(x_k)}$
1.0000	-0.6321	-1.3679	0.4621
0.5379	0.0461	-1.5840	-0.0291
0.5670	0.0002	-1.5672	-0.0002
0.5671	0.0000	-1.5671	-0.0000

Example

Estimates of the root of: x-cos(x)=0.

0.60000000000000		**Initial guess**
0.74401731944598	1	correct digit
0.73909047688624	4	correct digits
0.73908513322147	10	correct digits
0.73908513321516	14	correct digits

Example

In estimating the root of: ***x-cos(x)=0***, to get more than 13 correct digits:

- 4 iterations of Newton ($x_0=0.8$)
- 43 iterations of Bisection method (initial interval [0.6, 0.8])
- 5 iterations of Secant method ($x_0=0.6$, $x_1=0.8$)

Question

Given a floating ball with a specific gravity of 0.6 and has a radius of 5.5 cm. You are asked to find the depth to which the ball is submerged when floating in water.

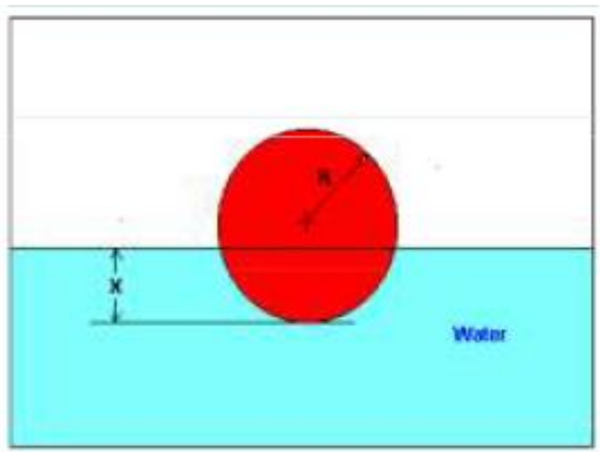


Diagram of the floating ball

The equation that gives the depth x to which the ball is submerged under water is given by:

$$x^3 - 0.165x^2 + 3.993x10^{-4} = 0$$

1. Use the Bisection method of finding roots of equations to find the depth x to which the ball is submerged under water. Conduct three iterations to estimate the root of the above equation.
2. Find the absolute relative approximate error at the end of each iteration.
3. Use both false position and newton methods to solve the roots of the equations.

Hint: From the Physics point of view, the ball would be submerged between x = 0 and x = 2R, where R = radius of the ball.

That is, $0 \le x \le 2R \implies 0 \le x \le 2(0.055) \implies 0 \le x \le 0.11$

CSC 431
Interpolation

Tutorials

Interpolation Methods

- Lagrange Polynomial
- Newton's Divided difference

Problem: Given a set of measured data, say $n+1$ pairs,

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$$

the problem of interpolation is to find a function $f(x)$ such that

$$f(x_i) = y_i, \quad i = 0, 1, \ldots, n$$

- x_i is called nodes;
- $f(x)$ is said to interpolate the data and is called interpolation function.
- $f(x)$ is said to approximate $g(x)$ if the data are from a function $g(x)$.
- It is called interpolate (or extrapolate) if $f(x)$ gives values within (or outside) $[x_0, x_n]$.

A simple choice for $f(x)$ is a polynomial of degree n:

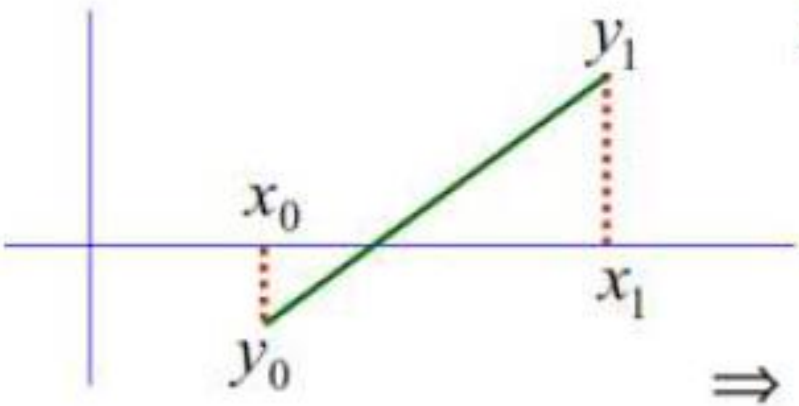
$$f(x) = p_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

The existence and uniqueness have been verified. There is only one polynomial existing for the interpolation.

LAGRANGIAN POLYNOMIALS

(a) Fitting Two Points

Fit the linear polynomial for two given points (x_0, y_0) and (x_1, y_1).



$$p_1(x) = L_0(x)y_0 + L_1(x)y_1$$

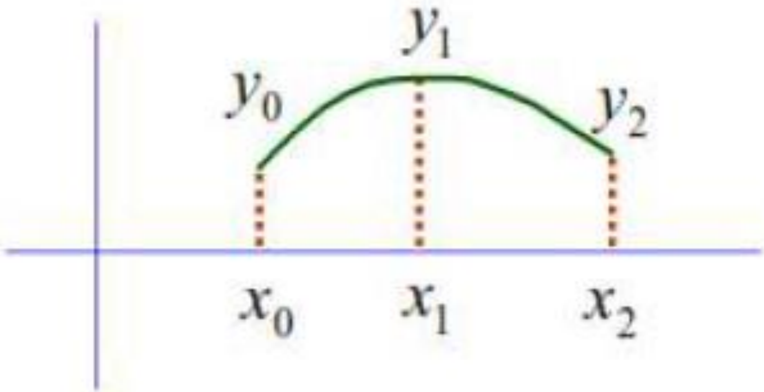
$$= \left(\frac{x - x_1}{x_0 - x_1}\right)y_0 + \left(\frac{x - x_0}{x_1 - x_0}\right)y_1$$

$$\Rightarrow \quad p_1(x_0) = y_0, \quad p_1(x_1) = y_1$$

27

(b) Fitting Three Points

Fit the quadratic polynomial for three given points.



$$\Rightarrow \quad \begin{aligned} p_2(x_0) &= y_0 \\ p_2(x_1) &= y_1 \\ p_2(x_2) &= y_2 \end{aligned}$$

$$\begin{aligned} p_2(x) &= \left(\frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)} \right) y_0 + \left(\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)} \right) y_1 + \left(\frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)} \right) y_2 \\ &= L_0(x) y_0 + L_1(x) y_1 + L_2(x) y_2 \end{aligned}$$

$$L_0(x) = \begin{cases} 1, & x = x_0 \\ 0, & x = x_1 \\ 0, & x = x_2 \end{cases} \qquad L_1(x) = \begin{cases} 0, & x = x_0 \\ 1, & x = x_1 \\ 0, & x = x_2 \end{cases} \qquad L_2(x) = \begin{cases} 0, & x = x_0 \\ 0, & x = x_1 \\ 1, & x = x_2 \end{cases}$$

(c) Fitting n+1 Points

Lagrangian polynomial for fitting $n+1$ given points

$$(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n).$$

is given by

$$p_n(x) = \sum_{k=0}^{n} L_k(x) y_k = \sum_{k=0}^{n} \frac{l_k(x)}{l_k(x_k)} y_k$$

where

$$l_k(x) = \frac{1}{x - x_k}(x - x_0)(x - x_1)(x - x_2)\cdots(x - x_n);$$

$$L_k(x) = \begin{cases} 1, & x = x_k \\ 0, & x = x_j, \quad j \neq k \end{cases}$$

The degree of the polynomial required is less than the number of data points. In general a data set of m points will require a polynomial of degree $m-1$ or less.

Lagrange Interpolating Polynomial

The Lagrange Form of an interpolating polynomial makes use of elimination of terms and cancellation to fit the data set.

Data Set	**Polynomial**
$\{(x_1, y_1)\}$	$p(x) = y_1$
$\{(x_1, y_1), (x_2, y_2)\}$	$p(x) = y_1 \frac{(x-x_2)}{(x_1-x_2)} + y_2 \frac{(x-x_1)}{(x_2-x_1)}$
$\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$	$p(x) = y_1 \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + y_2 \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + y_3 \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)}$
$\{(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)\}$	$p(x) = y_1 \frac{(x-x_2)(x-x_3)\cdots(x-x_m)}{(x_1-x_2)(x_1-x_3)\cdots(x_1-x_m)} + y_2 \frac{(x-x_1)(x-x_3)\cdots(x-x_m)}{(x_2-x_1)(x_2-x_3)\cdots(x_2-x_m)} + \cdots + y_m \frac{(x-x_1)(x-x_2)\cdots(x-x_{m-1})}{(x_m-x_1)(x_m-x_2)\cdots(x_m-x_{m-1})}$ $p(x) = \sum_{i=1}^{m} y_i \prod_{\substack{j=1 \\ j\neq i}}^{m} \frac{(x-x_j)}{(x_i-x_j)}$

Example 1

Compute ln 9.2 from ln 9.0 = 2.1972 and ln 9.5 = 2.2513 by the Lagrange interpolation and determine the error from a = ln 9.2 = 2.2192.

Solution. Given (x_0, f_0) and (x_1, f_1) we set

$$L_0(x) = \frac{x - x_1}{x_0 - x_1}, \quad L_1(x) = \frac{x - x_0}{x_1 - x_0},$$

which gives the Lagrange polynomial

$$p_1(x) = L_0(x)f_0 + L_1(x)f_1 = \frac{x - x_1}{x_0 - x_1}f_0 + \frac{x - x_0}{x_1 - x_0}f_1.$$

$\chi_0 = 9.0$, $\chi_1 = 9.5$, $f_0 = 2.1972$, and $f_1 = 2.2513$.

$$L_0(9.2) = \frac{9.2 - 9.5}{9.0 - 9.5} = 0.6, \quad L_1(9.2) = \frac{9.2 - 9.0}{9.5 - 9.0} = 0.4,$$

$$\ln 9.2 \approx \tilde{a} = p_1(9.2) = L_0(9.2) f_0 + L_1(9.2) f_1$$

$$= 0.6 \cdot 2.1972 + 0.4 \cdot 2.2513 = 2.2188.$$

The error is $\epsilon = a - \tilde{a} = 2.2192 - 2.2188 = 0.0004.$

Example 2

Using the following table,

(a) Explicitly construct the Lagrange Interpolating polynomial $P_3(x)$.

(b) Interpolate $f(3)$.

0	7
1	13
2	21
4	43

Degree of the interpolating polynomial: $n = 3$.

Nodes: $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 4$.

Functional values: $f_0 = 7, f_1 = 13, f_2 = 21, f_3 = 43$.

$$P_3(x) = L_0(x)f_0 + L_1(x)f_1 + L_2(x)f_2$$

$$L_0(x) = \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_1 - x_2)(x_0 - x_3)}$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)}$$

$$L_2(x) = \frac{(x - x_0)(x - x_1)(x - x_2)}{(x_2 - x_0)(x_2 - x_2)(x_2 - x_3)}$$

$$L_0(x) = \frac{(x-1)(x-2)(x-4)}{(-1)(-2)(-4)} = \frac{(x-1)(x-2)(x-4)}{-8}$$

$$L_1(x) = \frac{(x-0)(x-2)(x-4)}{1\cdot(-1)(-3)} = \frac{x(x-2)(x-4)}{3}$$

$$L_2(x) = \frac{(x-0)(x-1)(x-4)}{2\cdot 1\cdot(-2)} = \frac{x(x-1)(x-4)}{-4}$$

$$L_3(x) = \frac{(x-0)(x-1)(x-2)}{4\cdot 3\cdot 2} = \frac{x(x-1)(x-2)}{24}$$

$L_0(3) = \frac{1}{4}$, $L_1(3) = -1$, $L_2(3) = \frac{3}{2}$, $L_3(3) = \frac{1}{4}$.

So, $P_3(3) = 7L_0(3) + 13L_1(3) + 21L_2(3) + 43L_3(3) = 31$.

Interpolated value of $f(3) = P_3(3) = 31$.

Accuracy check: Note that $f(x)$ in this case is $f(x) = \chi^2 + 5\chi + 7$ and the exact value of $f(x)$ at $\chi = 3$ is 31. Thus the interpolation error for this example is zero. This is because the function to be interpolated is itself a polynomial of degree 2, and the interpolating polynomial $P_3(x)$ is of degree 3.

Example 3

Given

i	x_i	$f(x_i)$
0	2	$\frac{1}{2}$
1	2.5	$\frac{1}{2.5}$
2	4	$\frac{1}{4}$

(a) Explicitly find the Lagrange interpolating polynomial $P_2(x)$,

(b) Use $P_2(x)$ to interpolate $f(3)$.

Degree of the interpolating polynomial: $n = 2$.

Nodes: $x_0 = 2, x_1 = 2.5, x_2 = 4$

Functional values: $f_0 = \frac{1}{2}, f_1 = \frac{1}{2.5}, f_2 = \frac{1}{4}$

$$L_0(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 2.5)(x - 4)}{(-0.5)(-2)} = (x - 2.5)(x - 4) = x^2 - 6.5x + 10$$

$$L_1(x) = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 2)(x - 4)}{(0.5)(-1.5)} = -\frac{1}{0.75}(x - 2)(x - 4) = \frac{x^2 - 6x + 8}{0.75}$$

$$L_2(x) = \frac{(x - x_1)(x - x_0)}{(x_2 - x_1)(x_2 - x_0)} = \frac{1}{3}(x - 2.5)(x - 2) = \frac{x^2 - 4.5x + 5}{3}$$

$$P_2(x) = f(x_0)L_0(x) + f(x_1)L_1(x) + f(x_2)L_2(x) = \frac{1}{2}L_0(x) + \frac{1}{2.5}L_1(x) + \frac{1}{4}L_2(x)$$

$$= \frac{1}{2}(x^2 - 6.5x + 10) + \frac{1}{2.5}\frac{(x^2 - 6x + 8)}{0.75} + \frac{1}{4}\frac{(x^2 - 4.5x + 5)}{3}$$

Interpolated value of $f(3) = P_2(3) = 0.3250$.

Accuracy Check:

(i) The value of $f(x)$ at $x = 3$ is $\frac{1}{3} \approx 0.3333$.

Absolute Error: $|f(3) - P_2(3)| = |0.3333 - 0.3250| = 0.0083$.

Example 4

Compute $\ln 9.2$ from $\ln 9.0 = 2.1972$, $\ln 9.5 = 2.2513$, and $\ln 11.0 = 2.3979$ by the Lagrange interpolation and determine the error from $\ln 9.2 = 2.2192$.

Given (x_0, f_0), (x_1, f_1), and (x_2, f_2) we set

$$L_0(x) = \frac{l_0(x)}{l_0(x_0)} = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)},$$

$$L_1(x) = \frac{l_1(x)}{l_1(x_1)} = \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)},$$

$$L_2(x) = \frac{l_2(x)}{l_2(x_2)} = \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)},$$

$$p_2(x) = L_0(x)f_0 + L_1(x)f_1 + L_2(x)f_2.$$

χ_0= 9.0, χ_1= 9.5, χ_2= 11.0, f_0= 2.1972, and f_1= 2.2513, f_2= 2.3979.

$$L_0(x) = \frac{(x-9.5)(x-11.0)}{(9.0-9.5)(9.0-11.0)} = x^2 - 20.5x + 104.5, \quad L_0(9.2) = 0.5400;$$

$$L_1(x) = \frac{(x-9.0)(x-11.0)}{(9.5-9.0)(9.5-11.0)} = \frac{1}{0.75}(x^2 - 20x + 99), \quad L_1(9.2) = 0.4800;$$

$$L_2(x) = \frac{(x-9.0)(x-9.5)}{(11.0-9.0)(11.0-9.5)} = \frac{1}{3}(x^2 - 18.5x + 85.5), \quad L_2(9.2) = -0.0200$$

$$\ln 9.2 \approx p_2(9.2) = L_0(9.2)f_0 + L_1(9.2)f_1 + L_2(9.2)f_2 =$$

$$0.5400 \cdot 2.1972 + 0.4800 \cdot 2.2513 - 0.0200 \cdot 2.3979 = 2.2192,$$

with an actual error

$$\epsilon = 2.21920 - 2.21885 = 0.00035$$

Exercise 1

Find the Lagrange Interpolating polynomial P_3 through the points (0, 3), (1, 2), (2, 7), and (4, 59), and then approximate value $f(3)$ by $P_3(3)$.

$$p_3(\mathrm{x}) = \chi 2 \; - 2\chi + 3$$

$$f(3) \approx \; p_3(3) = 27 - 6 + 3 = 24$$

Exercise 2

Given that $f(-2) = 46$, $f(-1) = 4$, $f(1) = 4$, $f(3) = 156$, and $f(4) = 484$, use Lagrange's interpolation formula to estimate the value of $f(0)$.

Newton’s Divided difference

Example 1

- Given the following table of values:

x	1	1.5	2.5
$f(x)$	0	0.4055	0.9163

a) Compute the coe … erpolating polynomial $P_2(x)$ using divided differences.

b) Interpolate f(1.9).

Nodes: $x_0 = 1, x_1 = 1.5, x_2 = 2.5$

Functional Values: $f_0 = 0, f_1 = 0.4055, f_2 = 0.9163$

$P_0 = f_0 = 0.$

$$
\begin{aligned}
a_0 &= f[x_0] = f_0 = 0 \\
a_1 &= f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{x_1 - x_0} = \frac{f_1 - f_0}{x_1 - x_0} = 0.8118 \\
a_2 &= f[x_0, x_1, x_2] = \frac{f[x_1, x_2] - f[x_2, x_1]}{x_2 - x_0} \\
&= \frac{\dfrac{f[x_2] - f[x_1]}{x_2 - x_1} - \dfrac{f[x_1] - f[x_0]}{x_1 - x_0}}{x_2 - x_0} \\
&= \frac{\dfrac{f_2 - f_1}{x_2 - x_1} - \dfrac{f_1 - f_0}{x_1 - x_0}}{x_2 - x_0} = \frac{0 - 5108 - 0.81108}{1.5} = -0.2002
\end{aligned}
$$

Interpolation of $f(1.9)$ using $P_2(x)$:

$$\text{Let } P_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1)$$

$$= 0.6585$$

Accuracy Check: The tabulated values of $f(x)$ correspond to $f(x) = \ln(x)$.

Exact Value of $\ln(1.5) = 0.6419$.

Absolute Error: $|P_2(1.9) - \ln(1.5)| = |0.6585 - 0.6419| = 0.0166$.

Example 2

Compute cosh(0.56) from the given values and estimate the error.

Construct the table of forward differences

0.5	1.127626			
		0.057839		
0.6	1.185645		0.011865	
		0.069704		0.000697
0.7	1.255169		0.012562	
		0.082266		
0.8	1.337435			

$$\cosh(0.56) \approx p_3(0.56) =$$

$$1.127626+0.6\cdot 0.057839+\frac{0.6(-0.4)}{2}\cdot 0.011865+\frac{0.6(-0.4)(-1.4)}{6}\cdot 0.000697 =$$

$$1.127626 + 0.034703 - 0.001424 + 0.000039 = 1.160944.$$

$$\epsilon_3(0.56) = \cosh(0.56) - p_3(0.56)$$

LINEAR INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2) and x.

$$y = y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1}$$

QUADRATIC INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), and x.

$$y = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

CUBIC INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), and x.

$$y = y_1 \frac{(x - x_2)(x - x_3)(x - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)} + y_2 \frac{(x - x_1)(x - x_3)(x - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)} +$$

$$y_3 \frac{(x - x_1)(x - x_2)(x - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)} + y_4 \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)}$$

Nth INTERPOLATION: GENERAL SYSTEM OF INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$y = y_1 \frac{(x-x_2)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)\dots(x_1-x_{n-1})(x_1-x_n)}$$
$$+ \; y_2 \frac{(x-x_1)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)\dots(x_2-x_{n-1})(x_2-x_n)}$$
$$+ \; y_3 \frac{(x-x_1)(x-x_2)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)\dots(x_3-x_{n-1})(x_3-x_n)}$$
$$+ \; y_{n-1} \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_n)}{(x_n-x_1)(x_n-x_2)(x_n-x_3)\dots(x_n-x_{n-2})(x_n-x_{n-1})}$$
$$+ \; y_n \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)(x_n-x_3)\dots(x_n-x_{n-2})(x_n-x_{n-1})}$$

LAGRANGE METHOD

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$y = y\ (x_1)L_1(x) + y\ (x_2)L_2(x) + y\ (x_3)L_3(x) + y\ (x_4)L_4(x)$$
$$+ \; \dots\dots + y\ (x_{n-1})L_{n-1}(x) + y\ (x_n)L_n(x)$$

$$y = \sum_{i>0}^{n} y\ (x_i)L_i(x)$$

Find the velocity at t=16 seconds using the Lagrangian method for

a) Linear interpolation
b) Quadratic interpolation
c) Cubic interpolation

Methods for finding Roots of Algebraic and Transcendental equations

In all scientific fields, there's always the need to find the root of an equation, equivalently the zero of a function. Numerical methods allow for more complicated cases of handling roots of quadratic and polynomial equations.

Bisection method

As the name implies, we obtain the points x_1 and x_2, such that $f(x_2)\, f(x_1) < 0$, meaning that the value of f has opposite signs at the two points, which points to the fact that a root exists between x_1 and x_2. We approximate this root by the average of the two, i.e., $(x_1 + x_2) / 2$. Let this be x_3. Then we evaluate $f(x_3)$. x_3 is then combined with x_1 or x_2, depending on the one at which the sign of the function is opposite $f(x_3)$. This gives x_4. This process is repeated until f(x) attains the prescribed tolerance. The convergence of the Bisection method is slow and steady.

Bisection Method

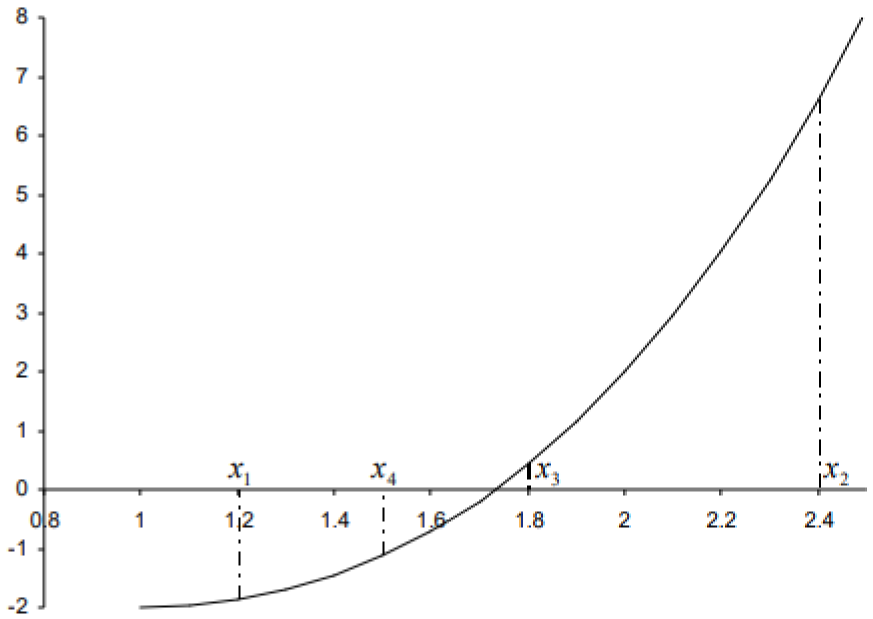


Fig.

- The **Bisection method** is one of the simplest methods to find a zero of a nonlinear function.
- It is also called **interval halving** method.
- To use the Bisection method, one needs an initial interval that is known to contain a zero of the function.
- The method systematically reduces the interval. It does this by dividing the interval into two equal parts, performs a simple test and based on the result of the test, half of the interval is thrown away.
- The procedure is repeated until the desired interval size is obtained.

Bisection Algorithm

Assumptions:

- f(x) is continuous on [a,b]
- $f(a)\,f(b) < 0$

Loop

1. Compute the mid point $c=(a+b)/2$
2. Evaluate f(c)
3. If $f(a)\,f(c) < 0$ then new interval [a, c]

   If $f(a)\,f(c) > 0$ then new interval [c, b]

End loop

Question 5a

Can you use Bisection method to find a zero of:
$f(x)=x^3-3x+1$ in the interval [0,2]?

Answer:

$f(x)$ is continuous on [0,2]
and $\mathrm{f}(0) * \mathrm{f}(2) = (1)(3) = 3 > 0$
$\Rightarrow$ Assumptions are not satisfied
$\Rightarrow$ Bisection method can not be used

Question 5b

Can you use Bisection method to find a zero of:
$f(x)=x^3-3x+1$ in the interval [0,1]?

Answer:

$f(x)$ is continuous on [0,1]
and $\mathrm{f}(0) * \mathrm{f}(1) = (1)(-1) = -1 < 0$
$\Rightarrow$ Assumptions are satisfied
$\Rightarrow$ Bisection method can be used

Question 5c

Find the root of:

$$f(x) = x^3 - 3x + 1 \text{ in the interval } :[0,1]$$

* $f(x)$ is continuous

* $f(0) = 1,\ f(1) = -1 \Rightarrow f(a)\ f(b) < 0$

$\Rightarrow$ Bisection method can be used to find the root

Iteration	a	b	c= (a+b) / 2	f(c)	(b-a) / 2
1	0	1	0.5	-0.375	0.5
2	0	0.5	0.25	0.266	0.25
3	0.25	0.5	.375	-7.23E-3	0.125
4	0.25	0.375	0.3125	9.30E-2	0.0625
5	0.3125	0.375	0.34375	9.37E-3	0.03125

Question 5d

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ between the points 1.4 and 1.7, using the bisection method. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(1.4) = -0.192$$

$$f(1.7) = 0.756$$

$$x_3 = \frac{1.4 + 1.7}{2} = 1.55$$

$$f(1.55) = 1.4025 \times 10^{-1}$$
$$x_4 = \frac{1.55 + 1.4}{2} = 1.475$$
$$f(1.475) = -0.0588$$
$$x_5 = \frac{1.55 + 1.475}{2} = 1.5125$$

This confirm that the Table for Bisection method is indeed true

n	x	$f(x)$
1	1.55	0.14025
2	1.475	-5.88E-02
3	1.5125	3.22E-02
4	1.49375	-1.54E-02
5	1.503125	7.87E-03
6	1.498437	-3.89E-03
7	1.500781	1.96E-03
8	1.499609	-9.76E-04
9	1.500195	4.89E-04
10	1.499902	-2.44E-04
11	1.500049	1.22E-04
12	1.499976	-6.10E-05

CONVERGENCE ANALYSIS OF BISECTION METHOD

Bisection method obtains an interval that is guaranteed to contain a zero of the function.

Questions:

- What is the best estimate of the zero of $f(x)$?
- What is the error level in the obtained estimate?

The best estimate of the zero of the function *f(x)* after the first iteration of the Bisection method is the mid point of the initial interval:

$$Estimate\ of\ the\ zero:\ \ r = \frac{b+a}{2}$$

$$Error \le \frac{b-a}{2}$$

$$Given\ \ f(x),\ a,\ b,\ and\ \ \varepsilon$$

How many iterations are needed such that: $|x - r| \le \varepsilon$ where *r* is the zero of *f(x)* and *x* is the bisection estimate (i.e., $x = c_k$)?

$$\boxed{n \ge \frac{\log(b-a) - \log(\varepsilon)}{\log(2)}}$$

$$\boxed{n \ge \log_2\left(\frac{\text{width of initial interval}}{\text{desired error}}\right) = \log_2\left(\frac{b-a}{\varepsilon}\right)}$$

Two common stopping criteria

1. Stop after a fixed number of iterations
2. Stop when the absolute error is less than a specified value

After n iterations :

$$|error| = |r - c_n| \le E_a^n = \frac{b-a}{2^n} = \frac{\Delta x^0}{2^n}$$

Question 5e

$a = 6,\ b = 7,\ \varepsilon = 0.0005$

How many iterations are needed such that: $|x - r| \le \varepsilon$?

$$n \ge \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1) - \log(0.0005)}{\log(2)} = 10.9658$$

$$\Rightarrow n \ge 11$$

Question 5f

- Use Bisection method to find a root of the equation x = cos(x) with absolute error <0.02
  (assume the initial interval [0.5, 0.9])

What is *f (x)* ?

Are the assumptions satisfied ?

How many iterations are needed ?

How to compute the new estimate ?

Question 5f (i) – what is f(x)?

x = cos(x)

f(x) = x – cos(x)

Question 5f (ii) – Are the assumptions satisfied?

Assuming interval [0.5, 0.9]

f(0.5) = 0.5 – cos(0.5) = -0.3776; This is a negative value

f(0.9) = 0.9 – cos(0.9) = 0.2784; This is a positive value

f(0.5)*f(0.9) = -0.3776 * 0.2764 < 0; Assumption is therefore satisfied.

Bisection method can be used.

Question 5f (iii) – How many iterations are needed?

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

a = 0.5, b = 0.9, ε = 0.02

n $\geq$ [log(0.9 – 0.5) – log(0.02)] / log(2)

n $\geq$ [-0.3979 - -1.6990] / 0.3010

n $\geq$ 1.3011 / 0.3010

n $\geq$ 4.3226

n $\geq$ 5

Question 5f (iii) – How to compute the new estimate?

$$Estimate\ of\ the\ zero:\ r=\frac{b+a}{2},\quad Error \leq \frac{b-a}{2}$$

r1 = (0.9 + 0.5) / 2 = 0.7; **Error < (0.9 - 0.5)/2 $\leq$ 0.2;**

f(0.7) = 0.7 – cos(0.7) = 0.7 – 0.9999 = -0.2999

f(0.5) = -0.3776; f(0.9) = 0.2784; f(0.7) = -0.2999

r2 = (0.7 + 0.9) / 2 = 0.8, **Error < (0.9 - 0.7) / 2 ≤ 0.1**

f(0.8) = 0.8 – cos(0.8) = 0.8 – 0.9999 = -0.1999

f(0.7) = -0.2999; f(0.9) = 0.2784; f(0.8) = -0.1999

r3 = (0.8 + 0.9) / 2 = 0.85, **Error < (0.9 - 0.8) / 2 ≤ 0.5**

f(0.85) = 0.85 – cos(0.85) = 0.85 – 0.9999 = -0.1499

f(0.8) = -0.1999; f(0.9) = 0.2784; f(0.85) = -0.1499

r4 = (0.85 + 0.9) / 2 = 0.875, **Error < (0.9 - 0.85) / 2 ≤ 0.025**

f(0.875) = 0.875 – cos(0.875) = 0.875 – 0.9999 = -0.1249

f(0.85) = -0.1499; f(0.9) = 0.2784; f(0.875) = -0.1249

r5 = (0.875 + 0.9) / 2 = 0.8875, **Error < (0.9 - 0.875) / 2 ≤ 0.02**

f(0.8875) = 0.8875 – cos(0.8875) = 0.8875 – 0.9999 = -0.1124

f(0.875) = -0.1249; f(0.9) = 0.2784; f(0.8875) = -0.1124

Newton-Raphson Method

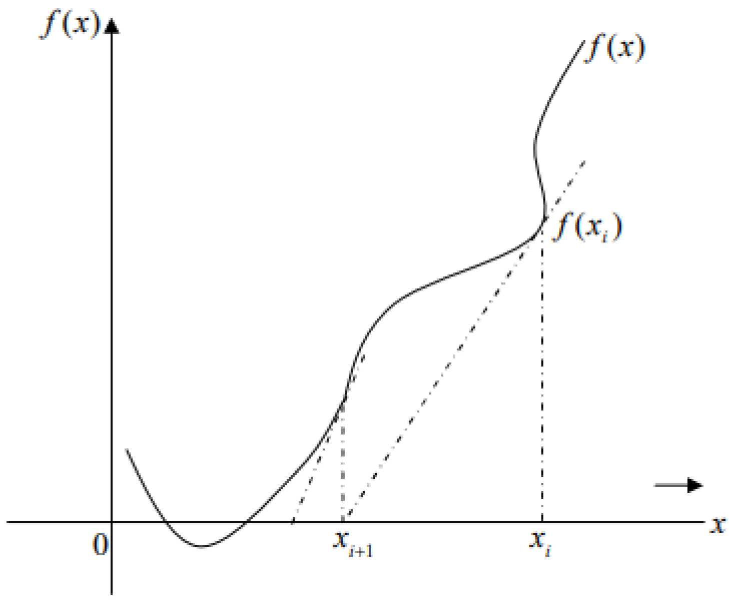
It is quite clear that the function $f(x)$ must be differentiable for you to be able apply the Newton-Raphson method.

More generally,

$$x_{i+1} = x_i + \Delta x = x_i - \frac{f(x_i)}{f'(x_i)}$$

With an initial guess of x_0, we can then get a sequence x_1, x_2, ..., which we expect to converge to the root of the equation.

Newton-Raphson method is equivalent to taking the slope of the function f(x) at the i^{th} iterative point, and the next approximation is the point where the slope intersects the x axis.



Question 5b

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ starting with the point 1.4, using the **Newton-Raphson Method**. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(x) = 2x^3 - 3x^2 - 2x + 3$$

$$f'(x) = 6x^2 - 6x - 2$$

$$x_0 = 1.4$$

$$\begin{aligned} x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} &= \frac{x_0 f'(x_0) - f(x_0)}{f'(x_0)} \\ &= \frac{6x_0^{\,3} - 6x_0^{\,2} - 2x_0 - 2x_0^{\,3} + 3x_0^{\,2} + 2x_0 - 3}{6x_0^{\,2} - 6x_0 - 2} \\ &= \frac{4x_0^{\,3} - 3x_0^{\,2} - 3}{6x_0^{\,2} - 6x_0 - 2} \\ &= \frac{4(1.4)^3 - 3(1.4)^2 - 3}{6(1.4)^2 - 6(1.4) - 2} \\ &= 1.5412 \end{aligned}$$

$x_1 = 1.5412$, $|x_1 - x_0| = 0.1412$

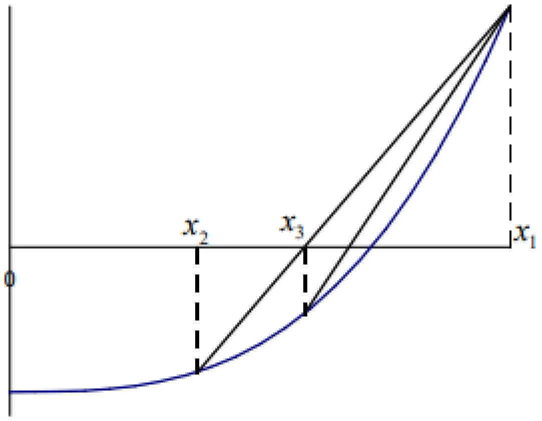
$x_2 = 1.5035$, $|x_2 - x_1| = 0.0377$

$x_3 = 1.5$, $|x_3 - x_2| = 0.0035$

$x_4 = 1.5$, $|x_4 - x_3| = 0$

Regula-falsi method

A regula-falsi or a method of false position assumes a test value for the solution of the equation.



Then, for an arbitrary x and the corresponding y,

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

gives the equation of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Setting $y = 0$, that is, where the chord crosses the x-axis,

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

Then, we evaluate $f(x_3)$. Just as in the case of root-bisection, if the sign is opposite that of $f(x_1)$, then a root lies in-between x_1 and x_3. Then, we replace x_2 by x_3 in equation

In just the same way, if the root lies between x_1 and x_3, we replace x_2 by x_1. We shall repeat this procedure until we are as close to the root as desired.

Question 5c

Find a zero of the function $f(x) = 2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the regula-falsi method.

$f(1.4) = -0.192$, $f(1.7) = 0.756$

A solution lies between $x = 1.4$ and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.4607595$

$f(1.4607595) = -0.088983$

The root lies between 1.46076 and 1.7. Let $x_1 = 1.46076$ and $x_2 = 1.7$.

$$x_4 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4607595 - (-0.088983)\frac{1.7 - 1.46076}{0.756 - (-0.088983)}$$
$$= 1.485953$$

Table for Regula-falsi method

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.495149	-0.011985
4	1.498346	-0.004118
5	1.499439	-0.001401
6	1.499810	-0.000475
7	1.499936	-0.000161
8	1.499978	-0.000055

Secant Method

In the case of the secant method, it is not necessary that the root lie between the two initial points. As such, the condition $f(x_1)f(x_2) < 0$ is not needed. Following the same analysis with the case of the regula-falsi method,

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

Setting $y = 0$ gives

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

Thus, having found x_n, we can obtain x_{n+1} as,

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})},\ n = 2, 3, \ldots$$

By inspection, if $f(x_n) - f(x_{n-1}) = 0$, the sequence does not converge, because the formula fails to work for x_{n+1}. The regula-falsi scheme does not have this problem as the associated sequence always converges.

Question 5d

Find the roots of the equation f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the secant method.

$x_1 = 1.4,\ x_2 = 1.7$

$f(1.4) = -0.192$, $f(1.7) = 0.756$

A solution lies between $x = 1.4$ and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$
$$= 1.460759$$

$$f(x_3) = -0.088983$$

$$x_4 = x_3 - f(x_3)\frac{x_3 - x_2}{f(x_3) - f(x_2)} = 1.460759 - (-0.088983) \times \frac{1.460759 - 1.7}{-0.088983 - 0.756}$$
$$= 1.485953$$

If the scheme continues, the table for secant method will be

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.501487	0.003730
4	1.499949	-0.000129
5	1.500000	0.000000

0 INTRO TO COMPUTATIONAL SCIENCE AND NUMERICAL METHODS

Numerical analysis ---> Numerical methods -----> Computational numerical methods ----> Computational numerical analysis

Numerical Analysis

Numerical Analysis is the study of numerical methods. Numerical analysis finds application in all fields of engineering and the physical sciences, and in the 21st century also the life and social sciences, medicine, business and even the arts. The GOAL of numerical analysis is the design and analysis of techniques/METHODS to give approximate but accurate solutions to hard problems.

Numerical methods

Numerical methods are mathematical attempts at finding approximate solutions of problems rather than the exact ones.

Computational numerical methods

Before modern computers, numerical methods often relied on hand formulas, using data from large printed tables. Since the mid20th century, computers calculate the required functions instead, but many of the same formulas continue to be used in software algorithms.

Computational numerical analysis

Current growth in computing power has enabled the use of more complex numerical analysis, providing detailed and realistic mathematical models in science and engineering. Numerical analysis continues this long tradition: rather than giving exact symbolic answers translated into digits and applicable only to real-world measurements, approximate solutions within specified error bounds are used.

1 Methods of Approximations

- Rounding off to significant figures
- Rounding off to decimal places
    - Working with arithmetic precision

2 Methods of Errors

    - Sources of errors
- Rounding errors
- Inherent errors
- Truncation errors
- True errors
- Relative true errors
- Absolute errors
- Relative absolute errors
- Approximate errors
- Relative approximate errors
- Absolute relative errors
- Percentage errors
    - Propagation of errors

3 Methods of Drawing the Lines of best fit

- Linearization
- Least squares curve fitting
- Group averages grouped averages curve fitting

4 Methods of solving Linear Systems of Equations

- Gaussian elimination
- Gauss-Jordan elimination
- LU decomposition
- Jacobi iteration
- Gauss-Seidal iteration

5 Methods of finding the roots of Algebraic and Transcendental Equations

- Bisection
- Newton-Raphson
- Regula-falsi
- Secant
- Modified Secant
- Fixed point Iteration
    - Zeros of functions, convergence analysis

6 Methods of Finite Differences

- First forward difference
- First backward difference
- First central difference

7 Methods of Interpolation

- Lagrange Interpolation; linear, quadratic, cubic, …
- Newton’s divided difference interpolation; first, second, third, …
- Newton’s forward interpolation formula
- Newton’s backward interpolation formula

8 Methods of Numerical Integration

- Newton-coates Quadrature
- Trapezoidal rule
- Simpson’s one-third rule
- Simpson’s three-eight rule
- Romberg’s method

9 Methods of Solving First Order Ordinary Differential Equations

- Picard’s Method
- Euler Method
- Modified Euler Method
- Runge-Kutta first order method
- Runge-Kutta second order method
- Runge-Kutta third order method
- Runge-Kutta fourth order method

1 Methods of Approximations

Question 1

Without calling any in-built library or function, write a new function from scratch to

a. round-off any number to a stated precision of decimal place
b. return the absolute value of any number
c. find the natural log of any number
d. approximate any number to its nearest_integer
e. take the magnitude of any number
f. round off any number to a stated amount of significant figures
g. re-write ‘f’ using built-in libraries and functions.

Solution to Question 1f

ALGORITHM - To round off numbers to certain amount of significant figures

```
Given any ‘number’, with the ‘num_sig_figs’ to approximate the number to
If ‘number’ == 0,
        Return 0.0
Else
        Take the absolute value of ‘number’
        Take the natural log of the absolute value
        Take ‘number’ to its nearest_integer
        Take the magnitude of the natural log
        Calculate rounding_factor = 10** (num_sig_figs – magnitude - 1)
        Find rounded_number = nearest_integer / rounding_factor
        If ‘number’ > 0
                Return rounded_number
        If ‘number’ < 0
                Return –rounded_number
```

PSEUDOCODE - To round off numbers to certain amount of significant figures

Without calling any in-built library or function, this is a new function from scratch to round off numbers to certain amount of significant figures

```
def round_to_significant_figures(number, num_sig_figs):
    if number == 0:
        return 0.0

    # Calculate the absolute value of 'number'
    def absolute_value(number):
        if number < 0:
            return -number
        else:
            return number
    abs_number = absolute_value(number)

    # Calculate the natural logarithm of the absolute value
    def custom_ln(number, num_terms=100):
        if number == 1:
            return 0.0
        elif number < 1:
            number = 1 / number
            num_terms = -num_terms
        nat_log = 0.0
        for n in range(1, num_terms + 1):
            term = ((number - 1) ** n) / n
            if n % 2 == 0:
                nat_log -= term
            else:
                nat_log += term
        return nat_log

    # Calculate the magnitude of the natural log
```

```
    def custom_floor(nat_log):
        if nat_log >= 0:
            return int(nat_log)
        else:
            integer_part = int(nat_log)
            if integer_part == nat_log:
                return integer_part
            else:
                return integer_part - 1
    magnitude = custom_floor(custom_ln(abs_number))

    # Calculate rounding_factor
    rounding_factor = 10 ** (num_sig_figs - magnitude - 1)

    # Use rounding factor to round number to the specified significant figures
    def custom_roundoff(number):
        decimal_part = number - int(number)
        if decimal_part < 0.5:
            return int(number)
        else:
            return int(number) + 1
    rounded_number = custom_roundoff(abs_number * rounding_factor) / rounding_factor

    # Restore the sign
    if number > 0:
        return rounded_number
    else:
        return -rounded_number

#round_to_significant_figures(number, num_sig_figs)
```

Arithmetic precision

It might be necessary to round off numbers to make them useful for numerical computation, moreso as it would require an infinite computer memory to store an unending number.

The precision of a number is an indication of the number of digits that have been used to express it. In scientific computing, it is the number of significant digits or numbers, while in management and financial systems, it is the number of decimal places. We are quite aware that most currencies in the world are quoted to two decimal places.

Arithmetic precision (often referred to simply as precision) is the specified number of significant figures or digits to which the number of interest is to be rounded.

2 Methods of errors

Rounding Errors

These are errors incurred by truncating a sequence of digits representing a number, as we saw in the case of representing the rational number 3/7 by 2.3333, instead of 2.3333….., which is an unending number. Apart from being unable to write this number in an exact form by hand, our instruments of calculation, be it the calculator or the computer, can only handle a finite string of digits. Rounding errors can be reduced if we change the calculation procedure in such a way as to avoid the subtraction of nearly equal numbers or division by a small number. It can also be reduced by retaining at least one more significant figure at each step than the one given in the data, and then rounding off at the last step.

Inherent Errors

As the name implies, these are errors that are inherent in the statement of the problem itself. This could be due to the limitations of the means of calculation, for instance, the calculator or the computer. This error could be reduced by using a higher precision of calculation.

Truncation Errors

If we truncate Taylor's series, which should be an infinite series, then some error is incurred. This is the error associated with truncating a sequence or by terminating an iterative process. This kind of error also results when, for instance, we carry out numerical differentiation or integration, because we are replacing an infinitesimal process with a finite one. In either case, we would have required that the elemental value of the independent variable tend to zero in order to get the exact value.

Absolute Error, Relative True Error, Relative Approximate Error and Percentage Error etc.

Question 2

a. A student measured the length of a string of actual length 72.5 cm as 72.4 cm.
   i. Calculate the absolute error and the percentage error
   ii. Write a function that accepts measured length and actual length to output absolute error and the percentage error.

b. The derivative of a function f(x) can be approximated by the equation

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}, \qquad \text{if } f(x) = 7e^{0.5x}, \text{ and } h = 0.3,$$

   i. Find the true value, the approximate value, true error, and relative error of f'(2)
   ii. If true values are not known or are very difficult to obtain, then Approximate error (E_a) = Present Approximation – Previous approximation. For f(x) = $7e^{0.5x}$ at x = 2, find
      - f'(2) using h = 0.3
      - f'(2) using h = 0.15

- approximate error, and relative approximate error of f'(2)

iii. Write a function that takes in any value of x for the derivative of a function f(x) approximated by the equation f'(x) = [f(x + h) - f(x)] / h for f(x) = 7e^0.5x, h1 = 0.3, h2 = 0.15, and returns true value, approximate value, true error, relative error, approximate error, and relative approximate error of f'(x)

Solution to Question 2a

(2ai)

Absolute error = | actual value – measured value |
Relative absolute error = | actual value – measured value | / actual value
The percentage error = Relative absolute error x 100
Absolute error = | 72.5 – 72.4 |= 0.01.
The percentage error = (0.1 / 72.5) x 100 = 0.1379

(2aii)

```
def calculate_errors(actual_length, measured_length):
    absolute_error = abs(actual_length - measured_length)
    relative_error = absolute_error / actual_length
    percentage_error = relative_error * 100

    return {
        "Absolute Error": absolute_error,
        "Relative Absolute Error": relative_error,
        "Percentage Error": percentage_error
    }

calculate_errors(actual_length, measured_length)
```

Solution to Question 2b

(2bi)
Approximate value of f'(x), for x = 2, and h = 0.3

$$f'(2) \approx \frac{f(2+0.3)-f(2)}{0.3}$$
$$= \frac{f(2.3)-f(2)}{0.3}$$
$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$
$$= \frac{22.107-19.028}{0.3} \quad = 10.263$$

(2bi cont'd)
The exact or true value of f'(2) can be found by using our knowledge of differential calculus

$$f(x) = 7e^{0.5x}$$
$$f'(x) = 7 \times 0.5 \times e^{0.5x}$$
$$= 3.5e^{0.5x}$$

$$f'(2) = 3.5e^{0.5(2)}$$
$$= 9.5140$$

(2bi cont'd)

True Error = True Value – Approximate Value
E_t = 9.5140 – 10263 = -0.722

Relative true error = (True value – Approximate value) / True value

= (9.5140 – 10263) / 9.5140= -0.722 / 9.5140

(2bii)

For x = 2, and h = 0.3

- Approximate value of f'(x) = 10.263

For x = 2, and h = 0.15

- Approximate value of f'(x) =

$$f^{'}(2) \approx \frac{f(2+0.15) - f(2)}{0.15}$$

$$= \frac{f(2.15) - f(2)}{0.15}$$

$$= \frac{7e^{0.5(2.15)} - 7e^{0.5(2)}}{0.15}$$

$$= \frac{20.50 - 19.028}{0.15} = 9.8800$$

Approximate error (E_a) = Present Approximation – Previous approximation

= 9.8800 – 10.263

= -0.38300

Relative approximate error = Approximate error / Previous approximation

$$= \frac{-0.38300}{9.8800} = -0.038765$$

(Question 2biii)

```
import math
def f(x):
    return 7 * math.exp(0.5 * x)
def derivative_of_f(x):
    return 3.5 * math.exp(0.5 * x)

def calculate_derivative_error(x, first_h, f, second_h=None):
    true_value = derivative_of_f(x)
    first_approximate_value = (f(x + first_h) - f(x)) / first_h
    true_error = abs(true_value - first_approximate_value)
    relative_true_error = true_error / true_value
    if second_h is not None:
        second_approximate_value = (f(x + second_h) - f(x)) / second_h
        approximate_error = abs(second_approximate_value - first_approximate_value)
        relative_approximate_error = approximate_error / second_approximate_value
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "Second Approximate Value": second_approximate_value,
            "Approximate Error": approximate_error,
            "Relative Approximate Error": relative_approximate_error
        }
    else:
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "True Error": true_error,
            "Relative True Error": relative_true_error
        }
#calculate_derivative_error(x, first_h, f)
#calculate_derivative_error(x, first_h, f, second_h)
```

Propagation of errors

In numerical methods, the calculations are not made with exact numbers. How do these inaccuracies propagate through the calculations?

Question 2c

Find the bounds for the propagation in adding two numbers. For example if one is calculating $X+Y$ where

$$X = 1.5 \pm 0.05$$

$$Y = 3.4 \pm 0.04$$

Solution

Maximum possible value of $X = 1.55$

Maximum possible value of $Y = 3.44$

Maximum possible value of $X + Y = 1.55 + 3.44 = 4.99$

Minimum possible value of $X = 1.45$.

Minimum possible value of $Y = 3.36$.

Minimum possible value of $X + Y = 1.45 + 3.36 = 4.81$

Hence

$$4.81 \leq X + Y \leq 4.99.$$

Propagation of Errors In Formula

If f is a function of several variables $X_1, X_2, X_3, \ldots\ldots., X_{n-1}, X_n$

then the maximum possible value of the error in f is

$$\Delta f \approx \left|\frac{\partial f}{\partial X_1}\Delta X_1\right| + \left|\frac{\partial f}{\partial X_2}\Delta X_2\right| + \ldots\ldots. + \left|\frac{\partial f}{\partial X_{n-1}}\Delta X_{n-1}\right| + \left|\frac{\partial f}{\partial X_n}\Delta X_n\right|$$

Question 2d

The strain in an axial member of a square cross-section is given by

$$\in = \frac{F}{h^2 E}$$

Given F = 72

h = 4 x 10^{-3}

E = 70 x 10^{9}

Find the maximum possible error in the measured strain.

Solution

$$\in = \frac{72}{(4\times10^{-3})^2(70\times10^9)}$$
$$= 64.286\times10^{-6}$$
$$= 64.286\mu$$

$$\Delta \in = \left|\frac{\partial \in}{\partial F}\Delta F\right| + \left|\frac{\partial \in}{\partial h}\Delta h\right| + \left|\frac{\partial \in}{\partial E}\Delta E\right|$$

$$\frac{\partial \in}{\partial F} = \frac{1}{h^2E} \qquad \frac{\partial \in}{\partial h} = -\frac{2F}{h^3E} \qquad \frac{\partial \in}{\partial E} = -\frac{F}{h^2E^2}$$

Thus

$$\Delta E = \left|\frac{1}{h^2E}\Delta F\right| + \left|\frac{2F}{h^3E}\Delta h\right| + \left|\frac{F}{h^2E^2}\Delta E\right|$$
$$= \left|\frac{1}{(4\times10^{-3})^2(70\times10^9)}\times0.9\right| + \left|\frac{2\times72}{(4\times10^{-3})^3(70\times10^9)}\times0.0001\right|$$
$$+ \left|\frac{72}{(4\times10^{-3})^2(70\times10^9)^2}\times1.5\times10^9\right|$$

Hence

$\in = (64.286\mu \pm 5.3955\mu)$

Question 2e

Subtraction of numbers that are nearly equal can create unwanted inaccuracies. Using the formula for error propagation, show that this is true.

Solution

Let z = x – y, Then

$$
\begin{aligned}
|\Delta z| &= \left|\frac{\partial z}{\partial x}\Delta x\right| + \left|\frac{\partial z}{\partial y}\Delta y\right| \\
&= |(1)\Delta x| + |(-1)\Delta y| \\
&= |\Delta x| + |\Delta y|
\end{aligned}
$$

So the relative error or relative change is

$$\left|\frac{\Delta z}{z}\right| = \frac{|\Delta x| + |\Delta y|}{|x-y|}$$

Check: $x = 2 \pm 0.001$

$$y = 2.003 \pm 0.001$$

$$\left|\frac{\Delta z}{z}\right| = \frac{|0.001| + |0.001|}{|2 - 2.003|}$$

$= 0.667$ Percentage error = 66.67%

Taylor series

Some examples of common Taylor series

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

The general form of Taylor series is given as

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots$$

provided that all derivatives of f(x) are continuous and exist in the interval [x,x+h]

What does this mean in plain English?

As Archimedes would have said, "*Give me the value of the function at a single point, and the (first, second, and so on) values of all its derivatives at that single point, and I can give you the value of the function at any other point*"

Question 2f

Find the value of f(6) given that f(4) =125, f'(4) = 74, f''(4) = 30, f'''(4) = 6 and all other higher order derivatives of f(x) at x=4 are zero.

Solution

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + \cdots$$

x = 4

h = 6 – 4 = 2

Since the higher order derivatives are zero,

$$f(4+2) = f(4) + f'(4)2 + f''(4)\frac{2^2}{2!} + f'''(4)\frac{2^3}{3!}$$

$$f(6) = 125 + 74(2) + 30\left(\frac{2^2}{2!}\right) + 6\left(\frac{2^3}{3!}\right)$$

= 125 + 148 + 60 + 8

= 341

Note that to find f(6) exactly, we only need the value of the function and all its derivatives at some other point, in this case x = 4.

Error in Taylor series

The Taylor polynomial of order n of a function f(x) with (n+1) continuous derivatives in the domain [x,x+h] is given by

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + \cdots + f^{(n)}(x)\frac{h^n}{n!} + R_n(x)$$

where the remainder is given by

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

Such that

$$x < c < x+h$$

that is, c is some point in the domain [x,x+h]

Derivation for Maclaurin Series for e^x

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

The Maclaurin series is simply the Taylor series about the point x=0

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + f''''(x)\frac{h^4}{4} + f'''''(x)\frac{h^5}{5} + \cdots$$

$$f(0+h) = f(0) + f'(0)h + f''(0)\frac{h^2}{2!} + f'''(0)\frac{h^3}{3!} + f''''(0)\frac{h^4}{4} + f'''''(0)\frac{h^5}{5} + \cdots$$

Since f(x) = e^x, f'(x) = e^x, f''(x) = e^x, $f^n(x) = e^x$, and $f^n(0) = e^0 = 1$;

The Maclaurin series is then:

$$f(h) = (e^0) + (e^0)h + \frac{(e^0)}{2!}h^2 + \frac{(e^0)}{3!}h^3 \ldots$$

$$= 1 + h + \frac{1}{2!}h^2 + \frac{1}{3!}h^3 \ldots$$

Therefore

$$f(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

It can be seen that as the number of terms used increases, the error bound decreases and hence a better estimate of the function can be found.

Question 2g

How many terms would it require to get an approximation of e^1 within a magnitude of true error of less than $10^{-6.}$

Solution

Using (n + 1) terms of Taylor seriesgives error bound of

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c) \qquad x = 0, h = 1, f(x) = e^x$$

$$R_n(0) = \frac{(0-1)^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

$$= \frac{(-1)^{n+1}}{(n+1)!} e^c$$

Since

x < c < x+h

0 < c < 0+1

0 < c < 1

$$\frac{1}{(n+1)!} < |R_n(0)| < \frac{e}{(n+1)!}$$

So if we want to find out how many terms it would require to get an approximation of e^1 magnitude of true error of less than 10^6,

$$\frac{e}{(n+1)!} < 10^{-6}$$

$$(n+1)! > 10^6 e$$

$$(n+1)! > 10^6 \times 3$$

$$n \geq 9$$

So 9 terms or more are needed to get a true error less than 10^{-6}.

3 Drawing line of best fit

The process of fitting a curve to a set of data is called curve-fitting.

Linearisation

A nonlinear relationship can be linearised and the resulting graph analysed to bring out the relationship between variables.

$y = ix + j$ ----------> linear or straight line graph, i=slope, j=intercept

$y = ix^2 + jx + k$ --------> quadratic graph or curve

$y = ix^n + jx + k$: $n >= 3$ --------> polynomial or sinusoidal wave form graph

$y = ie^x$ ---------------> ?? non-linear graph

$y = 2\log_x i3$ --------> ?? non-linear graph

Remember:

ln(x) is the natural logarithm to the base 'e' ≈ 2.71828, often referred to simply as "log."

$\log_{10}(x)$ is the common logarithm to the base 10, often referred to simply as "log."

In mathematical notation, the distinction is clear:

$\ln(x) = \log_e(x)$, where 'e' is the base of the natural logarithm.

$\log(x) = \log_{10}(x)$ where 10 is the base of common logarithm.

Case 1: $y = ae^x$.

(i) We could take the logarithm of both sides to base e:

$$\ln y = \ln(ae^x) = \ln a + \ln e^x = x + \ln a,$$

since $\ln e^x = x$. Thus, a plot of $\ln y$ against x gives a linear graph with slope unity and a y-intercept of ln a.

(ii) We could also have plotted y against e^x. The result is a linear graph through the origin, with slope equal to a.

Case 2: $T = 2\pi\sqrt{\dfrac{l}{g}}$

We can write this expression in three different ways:

(i) $\ln T = \ln(2\pi) + \dfrac{1}{2}\ln\left(\dfrac{l}{g}\right) = \ln(2\pi) + \dfrac{1}{2}(\ln l - \ln g)$.

Rearranging, we obtain,

$$\ln T = \frac{1}{2}\ln l + \left(\ln(2\pi) - \frac{1}{2}\ln g\right)$$

writing this in the form $y = mx + c$, we see that a plot of $\ln T$ against $\ln l$ gives a slope of 0.5 and a $\ln T$ intercept of $\left(\ln(2\pi) - \dfrac{1}{2}\ln g\right)$. Once the intercept is read of the graph, you can then calculate the value of g.

(ii) $T = \dfrac{2\pi}{\sqrt{g}}\sqrt{l}$

A plot of T versus $\sqrt{l}$ gives a linear graph through the origin (as the intercept is zero). The slope of the graph is $\dfrac{2\pi}{\sqrt{g}}$, from which the value of g can be recovered.

Case 3: $N = N_0 e^{-\lambda t}$

The student can show that a plot of $\ln N$ versus t will give a linear graph with slope $-\lambda$, and $\ln N$ intercept is $\ln N_0$.

What other functions of N and t could you plot in order to get λ and N_0?

Case 4: $\dfrac{1}{f} = \dfrac{1}{u} + \dfrac{1}{v}$

We rearrange the equation:

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$$

A plot of v^{-1} (y-axis) versus u^{-1} (x-axis) gives a slope of -1 and a vertical intercept of $\dfrac{1}{f}$.

Question 3a

A student obtained the following reading with a mirror in the laboratory.

u	10	20	30	40	50
v	-7	-10	-14	-15	-17

Linearise the relationship $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$. Plot the graph of v^{-1} versus u^{-1} and draw the line of best fit. Hence, find the focal length of the mirror. All distances are in cm.

Solution

u	v	1/u	1/v
10	-7	0.1	-0.14286
20	-10	0.05	-0.1
30	-14	0.033333	-0.07143
40	-15	0.025	-0.06667
50	-17	0.02	-0.05882

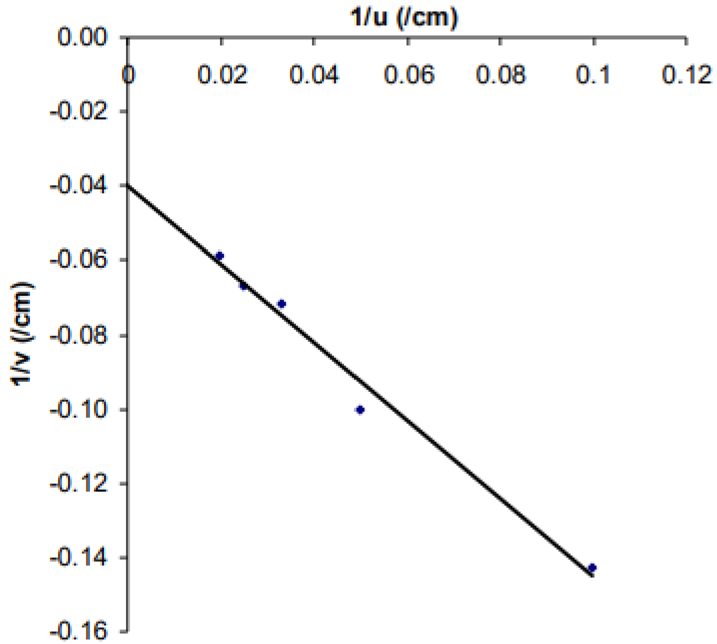


Fig. 1.1: Linear graph of the function $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$

The slope is -1.05 and the intercept -0.04. From $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$, we see that the intercept is $\frac{1}{f} = -0.04$, or $f = -\frac{1}{0.04} = -25$ cm.

Method of least squares curve fitting

The least square method entails minimizing the sum of the squares of the difference between the measured value and the one predicted by the assumed equation.

$$m = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2}$$

$$c = \bar{y} - m\bar{x}$$

Question 3b

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

Thus, for the following set of readings:

t	5	12	19	26	33
x	23	28	32	38	41

The table can be extended to give

t	5	12	19	26	33	Σ=95	$\bar{t}$ =19
x	23	28	32	38	41	Σ=162	$\bar{x}$ =32.4
tx	115	336	608	988	1353	Σ=3400	$\overline{tx}$=680
t^2	25	144	361	676	1089	Σ=2295	$\overline{x^2}$ =459

$$m = \frac{\overline{tx} - \bar{t}\,\bar{x}}{\overline{t^2} - \bar{t}^2} = \frac{680 - 19 \times 32.4}{459 - 19^2} = 0.6571$$

$$c = \bar{x} - m\bar{t} = 32.4 - 0.6571 \times 19 = 19.9151$$

Hence, the relationship between x and t is,

$$x = 0.6571t + 19.9151$$

Method of group averages curve fitting

As the name implies, a set of data is divided into two groups, each of which is assumed to have a zero sum of residuals.

$$\bar{y}_1 = m\bar{x}_1 + c$$
$$\bar{y}_2 = m\bar{x}_2 + c$$

Subtracting,

$$\bar{y}_1 - \bar{y}_2 = m(\bar{x}_1 - \bar{x}_2)$$
$$m = \frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1 - \bar{x}_2}$$

and

$$c = \bar{y}_1 - m\bar{x}_1$$

Question 3c

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

First divide the data into 2 groups

t	5	12	19
x	23	28	32

and

t	26	33
x	38	41

The tables can be extended to give, for Table 3:

t	5	12	19	Σ=36	$\bar{t}_1$=12
x	23	28	32	Σ=83	$\bar{x}_1$=27.666667

and for Table 4:

t	26	33	Σ=59	$\bar{t}_2$=29.5
x	38	41	Σ=79	$\bar{x}_2$=39.5

$$m = \frac{\bar{x}_1 - \bar{x}_2}{\bar{t}_1 - \bar{t}_2} = \frac{27.666667 - 39.5}{12 - 29.5} = 0.67619$$

and

$$c = \bar{x}_1 - m\bar{t}_1 = 27.666667 - (0.67619 \times 12)$$
$$= 19.552387$$

Thus, the equation of best fit is,

$$x = 0.67619t + 19.552387$$

4 Methods of Linear Systems of Equation

Let us consider a linear system of equations

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

This can be written in the form

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

Gaussian Elimination involving 2 variables

Question 4a: find x and y given that

$2x + 3y = 13$

$x – y = -1$

The augmented matrix representing our system of two equations is

$$\left[\begin{array}{cc|c} 2 & 3 & 13 \\ 1 & -1 & -1 \end{array}\right]$$

By Gaussian elimination, we seek to make every entry below the main diagonal zero. This we achieve by reducing 1 to zero, making use of the first row.

Thus,

$$5y = 15 \Rightarrow y = 3$$

Substituting this in the first row gives

$$2x + 3(3) = 13$$

from which we obtain $x = 2$.

The process of reducing every element below the main diagonal to zero (row echelon form) is called Gaussian Elimination. That of substituting obtained values to calculate other variables is called Back Substitution.

Gaussian Elimination involving 2 variables

The same process can be carried over to the case of a system of three equations.

Question 4b:

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

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

This yields (by Gaussian elimination)

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right] \xrightarrow[(iii)'=(i)-(2/3)(iii)]{(ii)'=(i)-2(ii)} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 7/3 & 5/3 & 3 \end{array}\right]$$

$$\xrightarrow{(iii)''=(ii)'+(15/7)(iii)'} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

Upon back substitution,

$$-10z = 10 \text{ or } z = -1$$

$$z = -1;\ y + z = 1 \Rightarrow y = 2;\ 2x + y - z = 5 \Rightarrow x = 1$$

Traditionally, in Mathematics, it is usual to use indices such as x_1, x_2, etc. instead of x, y, z. Do you have any idea why this is so? It is because if we stay with the alphabets, we shall soon run out of symbols. Bear in mind that not all the alphabets can be employed as variables; as an example, a, b, c are commonly used as constants. In addition, it makes it easy to associate the coefficients a_{11}, a_{12}, etc. with x_1, x_2, etc. respectively. More importantly in numerical work, it makes programming easier. For instance for our system of three equations, we could use the more general notation:

The general Gaussian elimination for linear system of 3 variables is thus given as:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] \begin{array}{c} \xrightarrow{(ii)'=a_{12}(i)-a_{11}(ii)} \\ \xrightarrow{(iii)=a_{11}(i)-a_{11}(iii)} \end{array} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & a_{32}' & a_{33}' & a_{34}' \end{array}\right]$$

$$\xrightarrow{(iii)''=a_{32}(ii)'+a_{22}(iii)'} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & 0 & a_{33}'' & a_{34}'' \end{array}\right]$$

Question 4b-2:

$$\begin{aligned} -3x1 + 2x2 - x3 &= -1, \\ 6x1 - 6x2 + 7x3 &= -7, \\ 3x1 - 4x2 + 4x3 &= -6. \end{aligned}$$

First write out the augmented matrix

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 6 & -6 & 7 & -7 \\ 3 & -4 & 4 & -6 \end{pmatrix}$$

Perform row reduction by multiplying the first row by 2 (the lcm of all x1's), then add first row to both second and third row

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & -2 & 3 & -7 \end{pmatrix}$$

Perform row reduction by multiplying the second row by -1 (the lcm of x2's in rows 2 and 3), then add second row to third row

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

$$-2x3 = 2 \rightarrow x3 = -1$$
$$-2x2 = -9 - 5x3 = -4 \rightarrow x2 = 2,$$
$$-3x1 = -1 - 2x2 + x3 = -6 \rightarrow x1 = 2.$$

Therefore, we have:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_n \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$$

Gauss-Jordan Elimination

This entails eliminating in addition to the entries below the major diagonal, the entries above it, so that the main matrix is a diagonal matrix. In that case, the solution to the system is given by dividing the element in the augmented part of the matrix by the diagonal element for that row.

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

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

In solving the same problem using Gauss-Jordan elimination, we continue from completion of the Gaussian elimination part.

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right] \xrightarrow[(ii)'=(iii)-2(ii)]{(i)'=(iii)-10(i)} \left[\begin{array}{ccc|c} -20 & 10 & 0 & -40 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

$$\xrightarrow{(i)''=(ii)'+(i)'} \left[\begin{array}{ccc|c} -20 & 0 & 0 & -20 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

It follows that,

$-20x = -20$ or $x = 1$; $-10y = 20$ or $y = 2$; and $-10z = 10$ or $z = -1$

(lower and upper echelon) - LU decomposition

Suppose we could write the matrix

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

This implies that

$$l_{11}u_{11} = a_{11},\ l_{11}u_{12} = a_{12},\ l_{11}u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11},\ a_{22} = l_{21}u_{12} + l_{22}u_{22},\ a_{23} = l_{21}u_{13} + l_{22}u_{23}$$
$$a_{31} = l_{31}u_{11},\ a_{32} = l_{31}u_{12} + l_{32}u_{22},\ a_{33} = l_{31}u_{13} + l_{32}u_{23} + l_{33}u_{33}$$

Without loss of generality, we could set the diagonal elements of the L matrix equal to 1.

Then,

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

Multiplying out the right side of equation 3.19,

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

From the equality of matrices, this requires that,

$$u_{11} = a_{11}$$

$$u_{12} = a_{12}$$
$$u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11} \qquad \Rightarrow l_{21} = a_{21}/u_{11} = a_{21}/a_{11}$$
$$a_{31} = l_{31}u_{11} \qquad \Rightarrow l_{31} = a_{31}/u_{11} = a_{31}/a_{11}$$
$$a_{22} = l_{21}u_{12} + u_{22} \text{ , or } u_{22} = a_{22} - l_{21}u_{12} = a_{22} - \frac{a_{21}}{u_{11}}u_{12}$$
$$\Rightarrow u_{22} = a_{22} - \frac{a_{21}}{a_{11}}a_{12}$$
$$a_{23} = l_{21}u_{13} + u_{23}\text{, or } u_{23} = a_{23} - l_{21}u_{13} = a_{23} - \frac{a_{21}}{u_{11}}u_{13}$$
$$\Rightarrow u_{23} = a_{23} - \frac{a_{21}}{a_{11}}a_{13}$$
$$l_{32} = \frac{a_{32} - l_{31}u_{12}}{u_{22}} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{u_{11}}u_{12}\right]$$
$$\Rightarrow l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}}a_{12}\right]$$
$$a_{32} = l_{31}u_{12} + l_{32}u_{22}$$
$$a_{33} = l_{31}u_{13} + l_{32}u_{23} + u_{33}$$
$$\Rightarrow u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23}$$

You can see that we have determined all the nine elements of the two matrices in terms of the elements of the original matrix.

Once we have obtained L and U, then we can write the original equation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

as

$$LU\mathbf{x} = \mathbf{y}$$

where $\mathbf{x}$ and $\mathbf{y}$ are column vectors.

We shall write $\mathbf{w} = U\mathbf{x}$

Then,

$$L\mathbf{w} = \mathbf{y}$$

Now we continue to solving **Question5b** again using LU decomposition

The corresponding matrix is

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

$$u_{11} = a_{11} = 2$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = -1$$
$$l_{21} = a_{21} / a_{11} = 1/2$$
$$l_{31} = a_{31} / a_{11} = 3/2$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 3 - \frac{1}{2}(1) = 5/2$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - \frac{1}{2}(-1) = 2 + \frac{1}{2} = 5/2$$
$$l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}} a_{12}\right] = \frac{1}{5/2}\left[-2 - \frac{3}{2}(1)\right] = -7/5$$
$$u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = -4 - (3/2)(-1) - (-7/5)(5/2) = 1$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 5/2 & 5/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 2 \\ 3 & -2 & -4 \end{bmatrix}$$

The above decomposition is correct as the multiplication of L and U gives the original matrix.

The original equation is equivalent to

$$LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$$

$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 3 \end{bmatrix}$$

Solving,

$$w_1 = \quad 5$$

$$\frac{1}{2}w_1 + w_2 = 5 \text{ or } w_2 = 5 - \frac{1}{2}w_1 = 5 - \frac{1}{2}(5) = \frac{5}{2}$$

$$\frac{3}{2}w_1 - \frac{7}{5}w_2 + w_3 = 3\text{, or } w_3 = 3 + \frac{7}{5}w_2 - \frac{3}{2}w_1 = 3 + \frac{7}{5}\left(\frac{5}{2}\right) - \frac{3}{2}(5) = -1$$

$U\mathbf{x} = \mathbf{w}$ implies:

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

By back substitution,

$$x_3 = -1$$

$$\frac{5}{2}x_2 + \frac{5}{2}x_3 = \frac{5}{2} \Rightarrow \frac{5}{2}x_2 = \frac{5}{2} - \frac{5}{2}x_3 = \frac{5}{2} - \frac{5}{2}(-1) = 5$$

$$x_2 = 2$$

$$2x_1 + x_2 - x_3 = 5$$

$$x_1 = \frac{5 - x_2 + x_3}{2} = \frac{5 - 2 + (-1)}{2} = 1$$

The solution set is therefore,

$$x_1 = 1,\ y = 2,\ z = -1.$$

Question 4c

Solve the system of linear equations $x+y+z=-1$, $x+2y+2z=-4$, $9x+6y+z=7$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

(i) Gaussian elimination

Initial augmented matrix

1	1	1	-1
1	2	2	-4
9	6	1	7

First round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	-3	-8	16

Second round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

Last matrix for Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

First round of Jordan elimination

5	5	0	2
0	5	0	-8
0	0	-5	7

Second round of Jordan elimination

-25	0	0	-50
0	5	0	-8
0	0	-5	7

(iii) LU decomposition

$x+y+z=-1$

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

$9x+6y+z=7$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7, \text{ or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$$x_3 = -7/5 = -1.4$$

$$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$$

$$x_2 = -8/5 = -1.6$$

$$x_1 + x_2 + x_3 = -1$$

$$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \frac{10}{5} = 2$$

The solution set is therefore,

$$x_1 = 2,\ y = -1.6,\ z = -1.4.$$

Question 4d

Solve the system of linear equations $x + 2y + 2z = -2$, $2x + 2y + z = -4$, $9x + 6y + 2z = -14$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

Gaussian elimination

Initial augmented matrix

1	2	2	-2
2	2	1	-4
9	6	2	-14

First round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	-12	-16	4

Second round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

Answers

x	0
y	-3
z	2

Gauss-Jordan elimination

Last matrix for Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

First round of Jordan elimination

4	8	0	-24
0	-8	0	24
0	0	-4	-8

Second round of Jordan elimination

32	0	0	0
0	-8	0	24
0	0	-4	-8

LU decomposition

$$x + 2y + 2z = -4$$
$$9x + 6y + z = 7$$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

W got the decomposition right, as the multiplication of the L and U gives the original matrix.

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7, \text{ or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$$x_3 = -7/5 = -1.4$$
$$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$$
$$x_2 = -8/5 = -1.6$$
$$x_1 + x_2 + x_3 = -1$$
$$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \frac{10}{5} = 2$$

The solution set is therefore,

$$x_1 = 2,\ y = -1.6,\ z = -1.4.$$

It is usually a good practice to revert to fractions to avoid incurring rounding errors.

5 Methods for finding Roots of Algebraic and Transcendental equations

In all scientific fields, there's always the need to find the root of an equation, equivalently the zero of a function. Numerical methods allow for more complicated cases of handling roots of quadratic and polynomial equations.

Bisection method

As the name implies, we obtain the points x_1 and x_2, such that $f(x_2)\, f(x_1) < 0$, meaning that the value of f has opposite signs at the two points, which points to the fact that a root exists between x_1 and x_2. We approximate this root by the average of the two, i.e., $(x_1 + x_2) / 2$. Let this be x_3. Then we evaluate $f(x_3)$. x_3 is then combined with x_1 or x_2, depending on the one at which the sign of the function is opposite $f(x_3)$. This gives x_4. This process is repeated until f(x) attains the prescribed tolerance. The convergence of the Bisection method is slow and steady.

Bisection Method

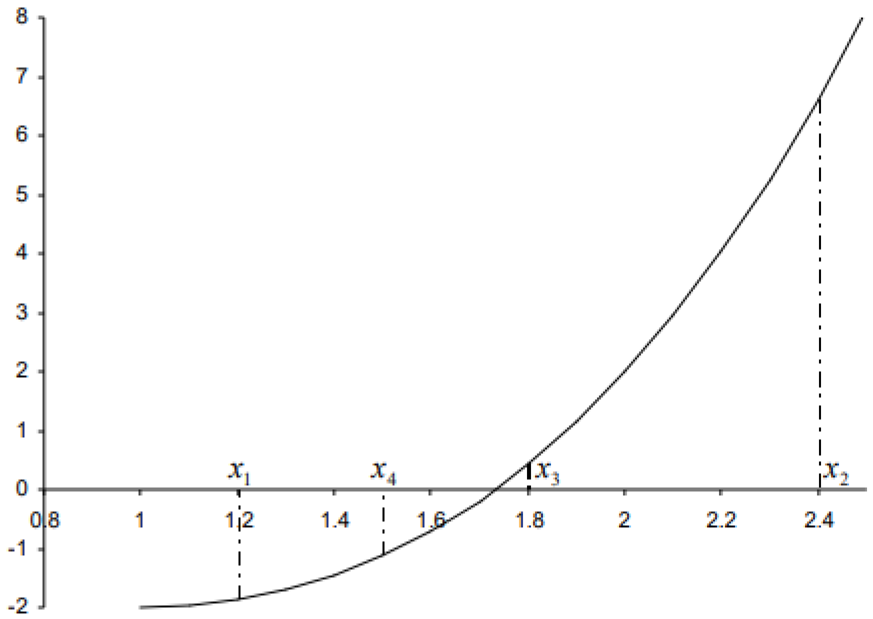


Fig.

- The **Bisection method** is one of the simplest methods to find a zero of a nonlinear function.
- It is also called **interval halving** method.
- To use the Bisection method, one needs an initial interval that is known to contain a zero of the function.
- The method systematically reduces the interval. It does this by dividing the interval into two equal parts, performs a simple test and based on the result of the test, half of the interval is thrown away.
- The procedure is repeated until the desired interval size is obtained.

Bisection Algorithm

Assumptions:

- f(x) is continuous on [a,b]
- $f(a)\,f(b) < 0$

Loop

1. Compute the mid point $c=(a+b)/2$
2. Evaluate f(c)
3. If $f(a)\,f(c) < 0$ then new interval [a, c]
   If $f(a)\,f(c) > 0$ then new interval [c, b]

End loop

Question 5a

Can you use Bisection method to find a zero of:
$f(x) = x^3 - 3x + 1$ in the interval [0,2]?

Answer:

$f(x)$ is continuous on [0,2]
and $\text{f}(0) * \text{f}(2) = (1)(3) = 3 > 0$
$\Rightarrow$ Assumptions are not satisfied
$\Rightarrow$ Bisection method can not be used

Question 5b

Can you use Bisection method to find a zero of:
$f(x) = x^3 - 3x + 1$ in the interval [0,1]?

Answer:

$f(x)$ is continuous on [0,1]
and $\text{f}(0) * \text{f}(1) = (1)(\text{-}1) = -1 < 0$
$\Rightarrow$ Assumptions are satisfied
$\Rightarrow$ Bisection method can be used

Question 5c

Find the root of:

$$f(x) = x^3 - 3x + 1 \quad \text{in the interval} : [0,1]$$

* $f(x)$ is continuous
* $f(0) = 1, \; f(1) = -1 \Rightarrow f(a)\, f(b) < 0$

$\Rightarrow$ Bisection method can be used to find the root

Iteration	a	b	c= $\frac{(a+b)}{2}$	f(c)	$\frac{(b-a)}{2}$
1	0	1	0.5	-0.375	0.5
2	0	0.5	0.25	0.266	0.25
3	0.25	0.5	.375	-7.23E-3	0.125
4	0.25	0.375	0.3125	9.30E-2	0.0625
5	0.3125	0.375	0.34375	9.37E-3	0.03125

Question 5d

Find a zero of the function $f(x) = 2x^3 - 3x^2 - 2x + 3$ between the points 1.4 and 1.7, using the bisection method. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(1.4) = -0.192$$

$$f(1.7) = 0.756$$

$$x_3 = \frac{1.4 + 1.7}{2} = 1.55$$

$$f(1.55) = 1.4025 \times 10^{-1}$$

$$x_4 = \frac{1.55 + 1.4}{2} = 1.475$$

$$f(1.475) = -0.0588$$

$$x_5 = \frac{1.55 + 1.475}{2} = 1.5125$$

This confirm that the Table for Bisection method is indeed true

n	x	$f(x)$
1	1.55	0.14025
2	1.475	-5.88E-02
3	1.5125	3.22E-02
4	1.49375	-1.54E-02
5	1.503125	7.87E-03
6	1.498437	-3.89E-03
7	1.500781	1.96E-03
8	1.499609	-9.76E-04
9	1.500195	4.89E-04
10	1.499902	-2.44E-04
11	1.500049	1.22E-04
12	1.499976	-6.10E-05

CONVERGENCE ANALYSIS OF BISECTION METHOD

Bisection method obtains an interval that is guaranteed to contain a zero of the function.

Questions:

- What is the best estimate of the zero of $f(x)$?
- What is the error level in the obtained estimate?

The <u>best estimate</u> of the zero of the function $f(x)$ after the first iteration of the Bisection method is the mid point of the initial interval:

$$\textit{Estimate of the zero}: \ r = \frac{b+a}{2}$$

$$\textit{Error} \leq \frac{b-a}{2}$$

$$\textit{Given} \quad f(x),\ a,\ b,\ \textit{and} \ \varepsilon$$

How many iterations are needed such that: $|x - r| \leq \varepsilon$ where r is the zero of $f(x)$ and x is the bisection estimate (i.e., $x = c_k$)?

$$\boxed{n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)}}$$

$$\boxed{n \geq \log_2\left(\frac{\text{width of initial interval}}{\text{desired error}}\right) = \log_2\left(\frac{b-a}{\varepsilon}\right)}$$

Two common stopping criteria

1. Stop after a fixed number of iterations
2. Stop when the absolute error is less than a specified value

After n iterations :

$$|error| = |r - c_n| \leq E_a^n = \frac{b-a}{2^n} = \frac{\Delta x^0}{2^n}$$

Question 5e

$a = 6,\ b = 7,\ \varepsilon = 0.0005$

How many iterations are needed such that: $|x - r| \leq \varepsilon$?

$$n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1) - \log(0.0005)}{\log(2)} = 10.9658$$

$$\Rightarrow n \geq 11$$

Question 5f

- Use Bisection method to find a root of the equation x = cos(x) with absolute error <0.02
  (assume the initial interval [0.5, 0.9])

What is *f (x)* ?

Are the assumptions satisfied ?

How many iterations are needed ?

How to compute the new estimate ?

Question 5f (i) – what is f(x)?

x = cos(x)

f(x) = x – cos(x)

Question 5f (ii) – Are the assumptions satisfied?

Assuming interval [0.5, 0.9]

f(0.5) = 0.5 – cos(0.5) = -0.3776; This is a negative value

f(0.9) = 0.9 – cos(0.9) = 0.2784; This is a positive value

f(0.5)*f(0.9) = -0.3776 * 0.2764 < 0; Assumption is therefore satisfied.

Bisection method can be used.

Question 5f (iii) – How many iterations are needed?

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

a = 0.5, b = 0.9, ε = 0.02

n $\geq$ [log(0.9 – 0.5) – log(0.02)] / log(2)

n $\geq$ [-0.3979 - -1.6990] / 0.3010

n $\geq$ 1.3011 / 0.3010

n $\geq$ 4.3226

n $\geq$ 5

Question 5f (iii) – How to compute the new estimate?

$$\textit{Estimate of the zero}: \; r = \frac{b+a}{2}, \quad Error \leq \frac{b-a}{2}$$

r1 = (0.9 + 0.5) / 2 = 0.7; **Error < (0.9 - 0.5)/2 $\leq$ 0.2;**

f(0.7) = 0.7 – cos(0.7) = 0.7 – 0.9999 = -0.2999

f(0.5) = -0.3776; f(0.9) = 0.2784; f(0.7) = -0.2999

r2 = (0.7 + 0.9) / 2 = 0.8, **Error < (0.9 - 0.7) / 2 $\leq$ 0.1**

f(0.8) = 0.8 – cos(0.8) = 0.8 – 0.9999 = -0.1999

f(0.7) = -0.2999; f(0.9) = 0.2784; f(0.8) = -0.1999

r3 = (0.8 + 0.9) / 2 = 0.85, **Error < (0.9 - 0.8) / 2 $\leq$ 0.5**

f(0.85) = 0.85 – cos(0.85) = 0.85 – 0.9999 = -0.1499

f(0.8) = -0.1999; f(0.9) = 0.2784; f(0.85) = -0.1499

r4 = (0.85 + 0.9) / 2 = 0.875, **Error < (0.9 - 0.85) / 2 $\leq$ 0.025**

f(0.875) = 0.875 – cos(0.875) = 0.875 – 0.9999 = -0.1249

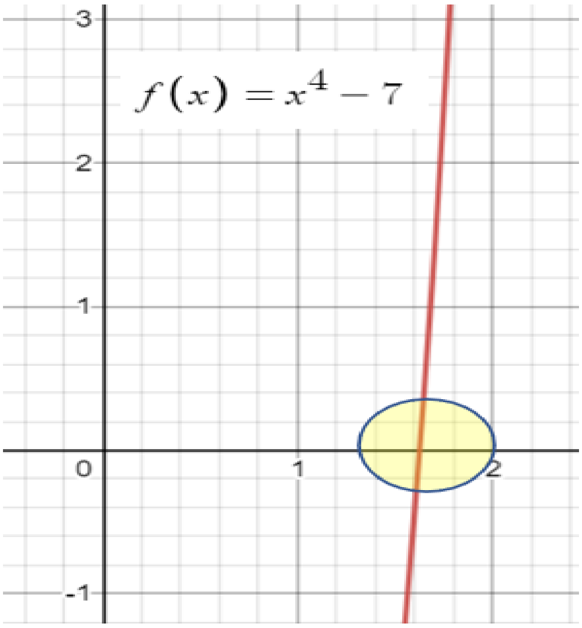
f(0.85) = -0.1499; f(0.9) = 0.2784; f(0.875) = -0.1249

r5 = (0.875 + 0.9) / 2 = 0.8875, **Error < (0.9 - 0.875) / 2 $\leq$ 0.02**

f(0.8875) = 0.8875 – cos(0.8875) = 0.8875 – 0.9999 = -0.1124

f(0.875) = -0.1249; f(0.9) = 0.2784; f(0.8875) = -0.1124

Question 5g



Find the 3rd approximation of the root of f(x) = x^4–7 using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(left)	f(mid)	f(right)	New Interval	Midpoint
f(1) = -6	f(1.5) = -2	f(2) = 9	(1.5, 2)	1.75
f(1.5) = -2	f(1.75) = 2.4	f(2) = 9	(1.5, 1.75)	1.625
f(1.5) = -2	f(1.625) = -.03	f(1.75) = 2.4	(1.625, 1.75)	1.6875

$f(x) = x^4 - 7$

$f(2) = (2)^4 - 7 = 9$; this is positive

$f(1) = (1)^4 - 7 = -6$; this is negative

$f(2)*f(1) = 9 * -6 < 0$; Assumption is therefore satisfied. Bisection method can be used.

for;

Starting interval (1, 2)

mid x = [2+1] / 2 = 1.5; Initial estimate

$f(mid) = f(1.5) = (1.5)^4 - 7 = 5.0625 - 7 = -1.9375$.

for;

f(2) = 9, f(1) = -6, **f(1.5) = -1.9375**

Next interval (2, 1.5)

mid x = [2+1.5]/2 = 1.75; first approximation

$f(mid) = f(1.75) = (1.75)^4 - 7 = 9.3789 - 7 = 2.3789$.

for;

f(2) = 9**,** **f(1.5) = -1.9375,** **f(1.75) = 2.3789**

Next interval (1.75, 1.5)

mid x = [1.75+1.5]/2 = 1.625; second approximation

$f(mid) = f(1.625) = (1.625)^4 - 7 = 6.9729 - 7 = -0.0271$.

for;

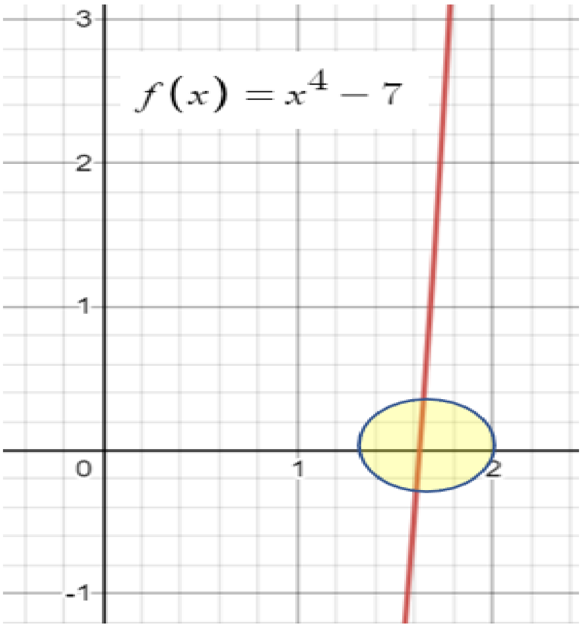
f(1.75) = 2.3789, f(1.5) = -1.9375, **f(1.625) = -0.0271**

Next interval (1.75, 1.625)

mid x = [1.75+1.625]/2 = 1.6875; third approximation

$f(mid) = f(1.6875) = (1.6875)^4 - 7 = 8.1091 - 7 = 1.1091$. **Stop.**

Question 5h



Find the 3rd approximation of the root of $f(x) = 10 – x^2$ using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

$f(x) = 10 – x^2$

> **$f(2) = 10 - (2)^2 = 6$; this is positive**
>
> **$f(1) = 10 - (1)^2 = 9$; this is also positive**
>
> **$f(2)*f(1) = 6 * 9 \ !< 0$;** Assumption is NOT satisfied. Bisection method cannot be used.

Question 5h

Given a floating ball with a specific gravity of 0.6 and has a radius of 5.5 cm. You are asked to find the depth to which the ball is submerged when floating in water.

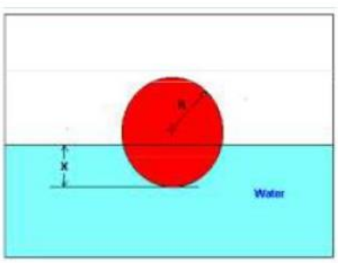


Diagram of the floating ball

The equation that gives the depth x to which the ball is submerged under water is given by:

$$x^3 - 0.165x^2 + 3.993x10^{-4} = 0$$

1. Use the Bisection method of finding roots of equations to find the depth x to which the ball is submerged under water. Conduct three iterations to estimate the root of the above equation.
2. Find the absolute relative approximate error at the end of each iteration.
3. Use both false position and newton methods to solve the roots of the equations.

Hint: From the Physics point of view, the ball would be submerged between $x = 0$ and $x = 2R$, where R = radius of the ball.

That is, $0 \leq x \leq 2R$ ==> $0 \leq x \leq 2(0.055)$ ==> $0 \leq x \leq 0.11$

Newton-Raphson Method

(Also known as Newton's Method)

Given an initial guess of the root x_0, Newton-Raphson method uses information about the function and its derivative at that point to find a better guess of the root.

Assumptions:

- $f(x)$ is continuous and the first derivative is known
- An initial guess x_0 such that $f'(x_0) \neq 0$ is given

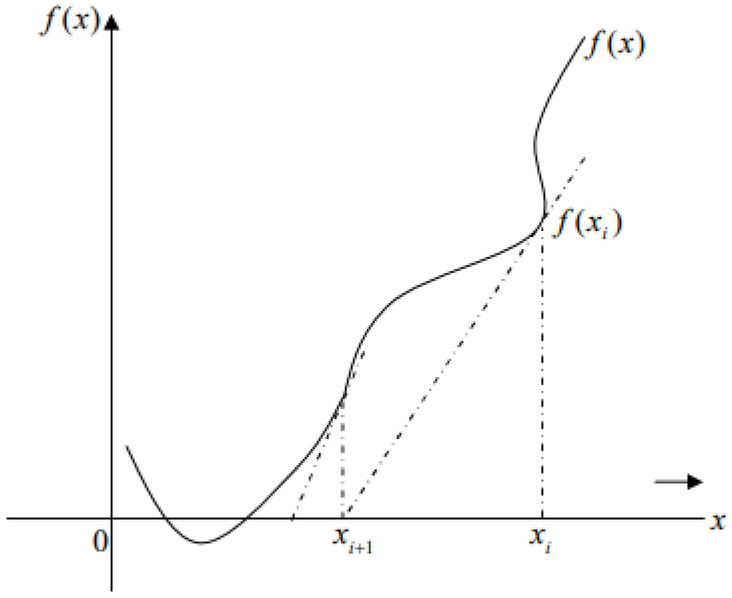
It is quite clear that the function $f(x)$ must be differentiable for you to be able apply the Newton-Raphson method.

More generally,

$$x_{i+1} = x_i + \Delta x = x_i - \frac{f(x_i)}{f'(x_i)}$$

With an initial guess of x_0, we can then get a sequence x_1, x_2, …, which we expect to converge to the root of the equation.

Newton-Raphson method is equivalent to taking the slope of the function f(x) at the i[th] iterative point, and the next approximation is the point where the slope intersects the x axis.



Question 5i

Find a zero of the function $f(x) = 2x^3 – 3x^2 – 2x + 3$ starting with the point 1.4, using the **Newton-Raphson Method**. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(x) = 2x^3 - 3x^2 - 2x + 3$$
$$f'(x) = 6x^2 - 6x - 2$$
$$x_0 = 1.4$$
$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = \frac{x_0 f'(x_0) - f(x_0)}{f'(x_0)}$$
$$= \frac{6x_0^3 - 6x_0^2 - 2x_0 - 2x_0^3 + 3x_0^2 + 2x_0 - 3}{6x_0^2 - 6x_0 - 2}$$
$$= \frac{4x_0^3 - 3x_0^2 - 3}{6x_0^2 - 6x_0 - 2}$$
$$= \frac{4(1.4)^3 - 3(1.4)^2 - 3}{6(1.4)^2 - 6(1.4) - 2}$$
$$= 1.5412$$

$x_1 = 1.5412$, $|x_1 - x_0| = 0.1412$
$x_2 = 1.5035$, $|x_2 - x_1| = 0.0377$
$x_3 = 1.5$, $|x_3 - x_2| = 0.0035$
$x_4 = 1.5$, $|x_4 - x_3| = 0$

Question 5j-2

Find a zero of the function $f(x) = x^3 - 2x^2 + x - 3$, $x_0 = 4$

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

Iteration 1: $x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)} = 4 - \dfrac{33}{33} = 3$

Iteration 2: $x_2 = x_1 - \dfrac{f(x_1)}{f'(x_1)} = 3 - \dfrac{9}{16} = 2.4375$

Iteration 3: $x_3 = x_2 - \dfrac{f(x_2)}{f'(x_2)} = 2.4375 - \dfrac{2.0369}{9.0742} = 2.2130$

k (Iteration)	$\mathbf{x_k}$	$\mathbf{f(x_k)}$	$\mathbf{f'(x_k)}$	$\mathbf{x_{k+1}}$	$\mathbf{\lvert x_{k+1} - x_k \rvert}$
0	**4**	**33**	**33**	**3**	1
1	**3**	**9**	**16**	**2.4375**	0.5625
2	**2.4375**	**2.0369**	**9.0742**	2.2130	0.2245
3	2.2130	0.2564	6.8404	2.1756	0.0384
4	2.1756	0.0065	6.4969	2.1746	0.0010

Question 5j-2

Use Newton's Method to find a root of:

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

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

Five Iterations of the Solution

k	x_k	$f(x_k)$	$f'(x_k)$	ERROR
0	1.0000	-1.0000	2.0000	
1	1.5000	0.8750	5.7500	0.1522
2	1.3478	0.1007	4.4499	0.0226
3	1.3252	0.0021	4.2685	0.0005
4	1.3247	0.0000	4.2646	0.0000
5	1.3247	0.0000	4.2646	0.0000

Question 5j-3

Use Newton's Method to find a root of:

$f(x) = e^{-x} - x$

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

x_k	$f(x_k)$	$f'(x_k)$	$\frac{f(x_k)}{f'(x_k)}$
1.0000	-0.6321	-1.3679	0.4621
0.5379	0.0461	-1.5840	-0.0291
0.5670	0.0002	-1.5672	-0.0002
0.5671	0.0000	-1.5671	-0.0000

Question 5k

Estimates of the root of: x-cos(x)=0.

0.600000000000000		**Initial guess**
0.744017319445598	1	correct digit
0.739090476886624	4	correct digits
0.739085133322147	10	correct digits
0.739085133321516	14	correct digits

Question 5k-2

Given the equation: $\boldsymbol{f(x) = x^3 - 10 = 0}$ which root lies between 2 and 3. Find the real root using the Newton Raphson method (Up to 3 iterations and correct to 4 decimal places).

Taking $x_0 = 2$.

Let

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

and

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

Given $x_0 = 2$

$$\boldsymbol{x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}} \qquad (1)$$

Where i=0,1,2,3…

Let $f(x_i) = f(x)$ and $f'(x_i) = f'(x)$

First Iteration:

Hence, substituting $i = 0$ into equation (1) to get the first approximation of the root

We have

$$\boldsymbol{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}} \qquad \mathbf{(2)}$$

$x_0 = 2;\ f(x_0) = -2;\ f'(x_0) = 12$

Note: $f(x_0)\ and\ f'(x_0)$ are derived by substituting $x_0 = 2$ into $f(x_i)\ and\ f'(x_i)$ where $i = 0$

Substituting the values into equation (2)

$x_1 = 2 - \frac{(-2)}{12} = 2.1667$

$x_1 = 2.1667$

Second Iteration:

Hence, substituting $i = 1$ into equation (1) to get the second approximation of the root

We have

$$\boldsymbol{x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}} \qquad \mathbf{(3)}$$

$x_1 = 2.1667;\ f(x_1) = 0.1718;\ f'(x_1) = 14.0838$

Note: $f(x_1)\ and\ f'(x_1)$ are derived by substituting $x_1 = 2.1667$ into $f(x_i)\ and\ f'(x_i)$ where $i = 1$

Substituting the values into equation (3)

$x_2 = 2.1667 - \frac{0.1718}{14.0838} = 2.1545$

$x_2 = 2.1545$

Third Iteration:

Hence, substituting $i = 2$ into equation (1) to get the third approximation of the root

We have

$$\boldsymbol{x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}} \qquad \mathbf{(4)}$$

$x_2 = 2.1545;\ f(x_2) = 0.0009;\ f'(x_2) = 13.9256$

Note: $f(x_1)\ and\ f'(x_1)$ are derived by substituting $x_2 = 2.1545$ into $f(x_i)\ and\ f'(x_i)$ where $i = 2$

Substituting the values into equation (4)

$x_3 = 2.1545 - \frac{0.0009}{13.9256} = 2.1544$

$x_3 = 2.1544$

Summary Table of Iterations

i (iteration)	X_i	$f(x_i)$	$f'(x_i)$	X_{i+1}	$\lvert X_{i+1} - X_i \rvert$
0	2	-2	12	2.1667	0.1667
1	2.1667	0.1718	14.0838	2.1545	0.0122
2	2.1545	0.0009	13.9256	2.1544	0.0001
3	2.1544	-0.0005	13.9243	2.1544	0

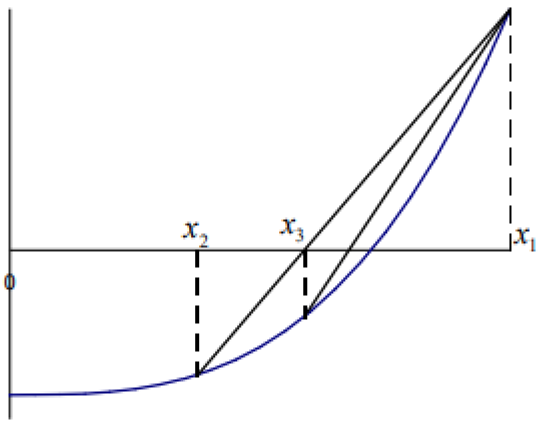
Thus, the root of the equation: $f(x) = x^3 - 10 = 0$ is <u>2.1544</u>

Regula-falsi method

- Also known as the false-position method, or linear interpolation method.

A regula-falsi or a method of false position assumes a test value for the solution of the equation.

- The *regula falsi* method start with two point, (a, f(a)) and (b,f(b)), satisfying the condition that f(a)f(b)<0.



Then, for an arbitrary x and the corresponding y,

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

gives the equation of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Setting $y = 0$, that is, where the chord crosses the x-axis,

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

Then, we evaluate $f(x_3)$. Just as in the case of root-bisection, if the sign is opposite that of $f(x_1)$, then a root lies in-between x_1 and x_3. Then, we replace x_2 by x_3 in equation

In just the same way, if the root lies between x_1 and x_3, we replace x_2 by x_1. We shall repeat this procedure until we are as close to the root as desired.

Question 5k-3

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the regula-falsi method.

$$f(1.4) = -0.192\,,\ \ f(1.7) = 0.756$$

A solution lies between x = 1.4 and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.4607595$

$$f(1.4607595) = -0.088983$$

The root lies between 1.46076 and 1.7. Let $x_1 = 1.46076$ and $x_2 = 1.7$.

$$x_4 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4607595 - (-0.088983)\frac{1.7 - 1.46076}{0.756 - (-0.088983)}$$

$= 1.485953$

Table for Regula-falsi method

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.495149	-0.011985
4	1.498346	-0.004118
5	1.499439	-0.001401
6	1.499810	-0.000475
7	1.499936	-0.000161
8	1.499978	-0.000055

Question 5L

- **Finding the Cube Root of 2 Using Regula Falsi**

- Since f(1)= -1, f(2)=6, we take as our starting bounds on the zero a=1 and b=2.
- Our first approximation to the zero is

$$x = b - \frac{b-a}{f(b) - f(a)}(f(b)) = 2 - \frac{2-1}{6+1}(6)$$

$$= 2 - 6/7 = 8/7 \approx 1.1429$$

- We then find the value of the function:

$$y = f(x) = (8/7)^3 - 2 \approx -0.5073$$

- Since f(a) and y are both negative, but y and f(b) have opposite signs

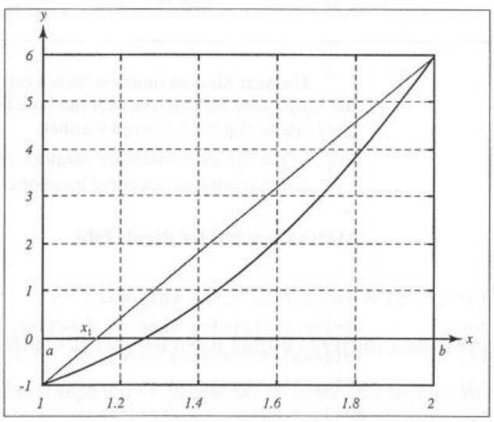


FIGURE 2.5 Graph of $y = x^3 - 2$ and approximation line on the interval [1, 2].

- Calculation of $\sqrt[3]{2}$ using *regula falsi.*

Step	a	b	x	y
1	1	2	1.1429	−0.50729
2	1.1429	2	1.2097	−0.22986
3	1.2097	2	1.2388	−0.098736
4	1.2388	2	1.2512	−0.041433
5	1.2512	2	1.2563	−0.017216
6	1.2563	2	1.2584	−0.0071239
7	1.2584	2	1.2593	−0.0029429
8	1.2593	2	1.2597	−0.0012148
9	1.2597	2	1.2598	−0.00050134
10	1.2598	2	1.2599	−0.00020687

Question 5L-2

Using the method of false position, find the real root of the equation $\boldsymbol{x^3 - 2x - 5 = 0}$. Where the real root lies between 2 and 2.1. (Up to 3 iterations and correct to 3 decimal places).

Let

$$f(x) = x^3 - 2x - 5$$

Given the roots as 2 and 2.1, therefore $a = 2\ and\ b = 2.1$

To find our approximation, we use the formula:

$$\boldsymbol{x_i = \frac{af(b) - bf(a)}{f(b) - f(a)}} \qquad (\mathbf{1})$$
$$\boldsymbol{where\ i = 1,2,3 \ldots}$$

First Iteration:

Where $i = 1; a = 2; b = 2.1; f(a) = -1;\ f(b) = 0.061$

Note: $f(a)\ and\ f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 1$

Therefore, substituting in the values into equation (1) we have

$$x_1 = \frac{2(0.061) - 2.1(-1)}{0.061 - (-1)}$$

$$x_1 = \frac{0.122 + 2.1}{1.061} = \frac{2.222}{1.061} = 2.0942$$

$$x_1 = 2.0942$$

Thus,

$$f(x_1) = f(2.0942) = (2.0942)^3 - 2(2.0942) - 5$$

$$f(x_1) = -0.0039$$

Since $f(x_1)$ is a negative value, therefore, the new root lies between (2.0942, 2.1) and $a = 2.0942$; $b = 2.1$

Second Iteration:

Where $i = 2; a = 2.0942; b = 2.1; f(a) = -0.0039;\ f(b) = 0.061$

Note: $f(a)$ *and* $f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 2$

Therefore, substituting in the values into equation (1) we have

$$x_2 = \frac{2.0942(0.061)-2.1(-0.0039)}{0.061-(-0.0039)}$$

$$x_2 = \frac{0.12775 + 0.00819}{0.0649} = \frac{0.13594}{0.0649} = 2.0946$$

$$x_2 = 2.0946$$

Thus,

$$f(x_2) = f(2.0946) = (2.0946)^3 - 2(2.0946) - 5$$

$$f(x_2) = 0.0005$$

Since $f(x_2)$ is a positive value, therefore, the new root lies between (2.0942, 2.0946) and $a = 2.0942$; $b = 2.0946$

Third Iteration:

Where $i = 3; a = 2.0942; b = 2.0946; f(a) = -0.0039;\ f(b) = 0.0005$

Note: $f(a)$ *and* $f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 3$

Therefore, substituting in the values into equation (1) we have

$$x_3 = \frac{2.0942(0.0005)-2.0946(-0.0039)}{0.0005-(-0.0039)}$$

$$x_3 = \frac{0.00105 + 0.00817}{0.0044} = \frac{0.00922}{0.0044} = 2.0952$$

$$x_3 = 2.0952$$

Summary Table of Iterations

i (iteration)	a	b	X_i	$f(x_i)$
1	2	2.1	2.0942	-0.0039
2	2.0942	2.1	2.0946	0.0005
3	2.0942	2.0946	2.0952	0.007

Therefore, after three iterations, the required approximate root correct to 3 decimal places is <u>2.095</u>

Secant Method

In the case of the secant method, it is not necessary that the root lie between the two initial points. As such, the condition $f(x_1)f(x_2) < 0$ is not needed. Following the same analysis with the case of the regula-falsi method,

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

Setting $y = 0$ gives

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

Thus, having found x_n, we can obtain x_{n+1} as,

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \; n = 2, 3, \ldots$$

By inspection, if $f(x_n) - f(x_{n-1}) = 0$, the sequence does not converge, because the formula fails to work for x_{n+1}. The regula-falsi scheme does not have this problem as the associated sequence always converges.

Question 5m

Find the roots of the equation f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the secant method.

$x_1 = 1.4, \; x_2 = 1.7$

$f(1.4) = -0.192, \; f(1.7) = 0.756$

A solution lies between $x = 1.4$ and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.460759$

$f(x_3) = -0.088983$

$$x_4 = x_3 - f(x_3)\frac{x_3 - x_2}{f(x_3) - f(x_2)} = 1.460759 - (-0.088983) \times \frac{1.460759 - 1.7}{-0.088983 - 0.756}$$

$= 1.485953$

If the scheme continues, the table for secant method will be

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.501487	0.003730
4	1.499949	-0.000129
5	1.500000	0.000000

Question 5n

Find the roots of the equation by the secant method:

$$f(x) = x^2 - 2x + 0.5$$

$$x_0 = 0$$

$$x_1 = 1$$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Question 5n-i

Find the roots of the equation by the secant method:

Use Secant method to find the root of:

$$f(x) = x^6 - x - 1$$

Two initial points $x_0 = 1$ *and* $x_1 = 1.5$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

k	x_k	$f(x_k)$
0	1.0000	-1.0000
1	1.5000	8.8906
2	1.0506	-0.7062
3	1.0836	-0.4645
4	1.1472	0.1321
5	1.1331	-0.0165
6	1.1347	-0.0005

Question 5n-ii

Given the equation: $f(x) = x^3 - 5x + 1$, where x_0 and x_1 are 2 and 2.5 respectively. Find the real root using the Secant method. (Up to 4 iterations and correct to 4 decimal places).

Let

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

Given $x_0 = 2 \ and \ x_1 = 2.5$

To find our approximation, we use the formula:

$$\boldsymbol{x_{i+1} = x_i - f(x_i)\left(\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})}\right)} \qquad (\mathbf{1})$$

where i=1,2,3 …

Let $\quad f(x_i) = f(x)$, where i=1,2,3 …

First Iteration:

Hence, substituting $i = 1$ into equation (1) to get the first approximation of the root

We have

$$\boldsymbol{x_2 = x_1 - f(x_1)\left(\frac{x_1 - x_0}{f(x_1) - f(x_0)}\right)} \qquad (\mathbf{2})$$

$x_0 = 2; x_1 = 2.5; \ f(x_0) = -1; \ f(x_1) = 4.125$

Note: $f(x_0) \ and \ f(x_1)$ are derived by substituting $x_0 = 2, \ x_1 = 2.5$ into $f(x_0) \ and \ f(x_1)$ respectively.

Substituting the values into equation (2)

$$x_2 = 2.5 - (4.125)\left(\frac{2.5 - 2}{4.125 - (-1)}\right)$$

$$x_2 = 2.5 - (4.125)\left(\frac{0.5}{5.125}\right)$$

$$x_2 = 2.5 - (4.125)(0.09756)$$

$$x_2 = 2.5 - 0.4024$$

$$x_2 = 2.0976$$

Thus,

$$f(x_2) = f(2.0976) = (2.0976)^3 - 5(2.0976) + 1$$

$$f(x_2) = -0.2587$$

Second Iteration:

Hence, substituting $i = 2$ into equation (1) to get the second approximation of the root

We have

$$\boldsymbol{x_3 = x_2 - f(x_2)\left(\frac{x_2 - x_1}{f(x_2) - f(x_1)}\right)} \qquad (\mathbf{3})$$

$x_1 = 2.5; x_2 = 2.0976; \ f(x_1) = 4.125; \ f(x_2) = -0.2587$

Note: $f(x_1) \ and \ f(x_2)$ are derived by substituting $x_1 = 2.5, \ x_2 = 2.0976$ into $f(x_1) \ and \ f(x_2)$ respectively.

Substituting the values into equation (3)

$$x_3 = 2.0976 - (-0.2587)\left(\frac{2.0976 - 2.5}{(-0.2587) - 4.125}\right)$$

$$x_3 = 2.0976 + (0.2587)\left(\frac{-0.4025}{-4.3837}\right)$$

$$x_3 = 2.0976 + (0.2587)(0.09182)$$

$$x_3 = 2.0976 + 0.0238$$

$$x_3 = 2.1214$$

Thus,

$$f(x_3) = f(2.1214) = (2.1214)^3 - 5(2.1214) + 1$$

$$f(x_3) = -0.0600$$

Third Iteration:

Hence, substituting $i = 3$ into equation (1) to get the third approximation of the root

We have

$$\boldsymbol{x_4 = x_3 - f(x_3)\left(\frac{x_3 - x_2}{f(x_3) - f(x_2)}\right)} \qquad (\mathbf{4})$$

$x_2 = 2.0976; x_3 = 2.1214;\ f(x_2) = -0.2587;\ f(x_3) = -0.0600$

Note: $f(x_2)$ *and* $f(x_3)$ are derived by substituting $x_2 = 2.0976,\ x_3 = 2.1214$ into $f(x_2)$ *and* $f(x_3)$ respectively.

Substituting the values into equation (4)

$$x_4 = 2.1214 - (-0.0600)\left(\frac{2.1214 - 2.0976}{(-0.0600) - (-0.2587)}\right)$$

$$x_4 = 2.1214 + (0.0600)\left(\frac{0.0238}{0.1987}\right)$$

$$x_4 = 2.1214 + (0.0600)(0.1198)$$

$$x_4 = 2.1214 + 0.0072$$

$$x_4 = 2.1286$$

Thus,

$$f(x_4) = f(2.1286) = (2.1286)^3 - 5(2.1286) + 1$$

$$f(x_4) = 0.0016$$

Fourth Iteration:

Hence, substituting $i = 4$ into equation (1) to get the fourth approximation of the root

We have

$$\boldsymbol{x_5 = x_4 - f(x_4)\left(\frac{x_4 - x_3}{f(x_4) - f(x_3)}\right)} \qquad (\mathbf{5})$$

$x_3 = 2.1214; x_4 = 2.1286;\ f(x_3) = -0.0600;\ f(x_4) = 0.0016$

Note: $f(x_3)$ *and* $f(x_4)$ are derived by substituting $x_3 = 2.1214,\ x_4 = 2.1286$ into $f(x_3)$ *and* $f(x_4)$ respectively.

Substituting the values into equation (5)

$$x_5 = 2.1286 - (0.0016)\left(\frac{2.1286 - 2.1214}{0.0016 - (-0.0600)}\right)$$

$$x_5 = 2.1286 - (0.0016)\left(\frac{0.0072}{0.0616}\right)$$

$$x_5 = 2.1286 - (0.0016)(0.1169)$$

$$x_5 = 2.1286 - 0.0002$$

$$x_5 = 2.1284$$

Thus,

$$f(x_5) = f(2.1284) = (2.1284)^3 - 5(2.1284) + 1$$

$$f(x_5) = -0.0002$$

Summary Table of Iterations

i (iteration)	**X_i**	**$f(x_i)$**	**X_{i+1}**	**$\lvert X_{i+1} - X_i \rvert$**
0	2	-1	2.5	0.5
1	2.5	4.125	2.0976	0.4022
2	2.0976	-0.2587	2.1214	0.0238
3	2.1214	-0.0600	2.1286	0.0072
4	2.1286	0.0016	2.1284	0.0002

Therefore, the root of the equation after 4 iterations correct to 4 decimal places is <u>2.1284</u>

Modified Secant Method

In this modified Secant method, only one initial guess is needed :

$$f'(x_i) \approx \frac{f(x_i + \delta\, x_i) - f(x_i)}{\delta\, x_i}$$

$$x_{i+1} = x_i - \frac{f(x_i)}{\dfrac{f(x_i + \delta\, x_i) - f(x_i)}{\delta\, x_i}} = \quad x_i - \frac{\delta\, x_i f(x_i)}{f(x_i + \delta\, x_i) - f(x_i)}$$

Question 5(o)

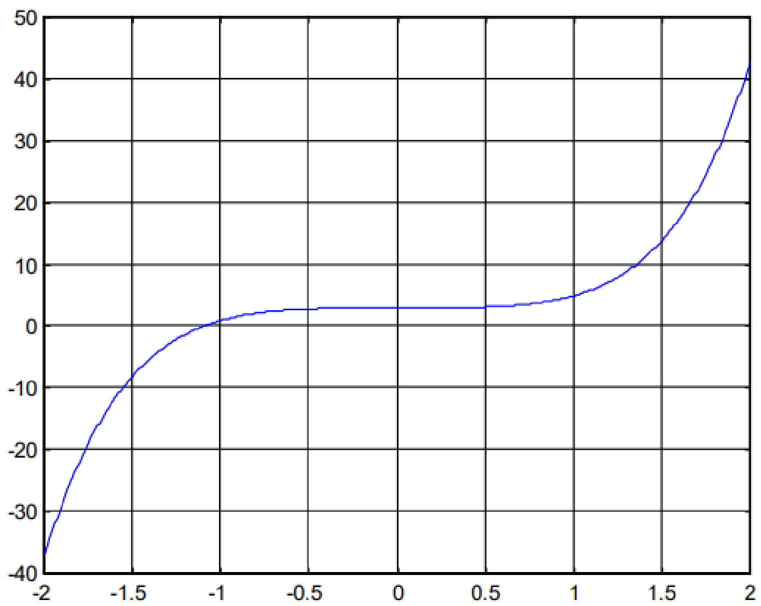
Find the roots of :

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

Initial points

$x_0 = -1 \; and \; x_1 = -1$

$with \;\; error < 0.001$



Fixed-point Iteration Method

- Also known as **one-point iteration** or **successive substitution**

- To find the root for $f(x) = 0$, we reformulate $f(x) = 0$ so that there is an x on one side of the equation.

$$f(x) = 0 \quad \Leftrightarrow \quad g(x) = x$$

- If we can solve $g(x) = x$, we solve $f(x) = 0$.
  - x is known as the fixed point of $g(x)$.
- We solve $g(x) = x$ by computing

$$x_{i+1} = g(x_i) \quad \text{with } x_0 \text{ given}$$

until x_{i+1} converges to x.

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

$$x^2 + 2x - 3 = 0 \quad \Rightarrow \quad 2x = 3 - x^2 \quad \Rightarrow \quad x = \frac{3 - x^2}{2}$$

$$\Rightarrow x_{i+1} = g(x_i) = \frac{3 - x_i^2}{2}$$

Reason: If x converges, i.e. $x_{i+1} \rightarrow x_i$

$$x_{i+1} = \frac{3 - x_i^2}{2} \rightarrow x_i = \frac{3 - x_i^2}{2}$$

$$\Rightarrow x_i^2 + 2x_i - 3 = 0$$

Question 5p

Use fixed point iteration to:

Find root of $f(x) = e^{-x} - x = 0$.
(Answer: $\alpha = 0.56714329$)

We put $x_{i+1} = e^{-x_i}$

i	x_i	ε_a (%)	ε_t (%)
0	0		100.0
1	1.000000	100.0	76.3
2	0.367879	171.8	35.1
3	0.692201	46.9	22.1
4	0.500473	38.3	11.8
5	0.606244	17.4	6.89
6	0.545396	11.2	3.83
7	0.579612	5.90	2.20
8	0.560115	3.48	1.24
9	0.571143	1.93	0.705
10	0.564879	1.11	0.399

- There are infinite ways to construct $g(x)$ from $f(x)$.

For example, $f(x) = x^2 - 2x - 3 = 0$ (ans: $x = 3$ or -1)

Case a:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x^2 = 2x + 3$$
$$\Rightarrow x = \sqrt{2x+3}$$
$$\Rightarrow g(x) = \sqrt{2x+3}$$

Case b:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x(x-2) - 3 = 0$$
$$\Rightarrow x = \frac{3}{x-2}$$
$$\Rightarrow g(x) = \frac{3}{x-2}$$

Case c:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow 2x = x^2 - 3$$
$$\Rightarrow x = \frac{x^2-3}{2}$$
$$\Rightarrow g(x) = \frac{x^2-3}{2}$$

So which one is better?

Case a

$$x_{i+1} = \sqrt{2x_i + 3}$$

1. $x_0 = 4$
2. $x_1 = 3.31662$
3. $x_2 = 3.10375$
4. $x_3 = 3.03439$
5. $x_4 = 3.01144$
6. $x_5 = 3.00381$

Converge!

Case b

$$x_{i+1} = \frac{3}{x_i - 2}$$

1. $x_0 = 4$
2. $x_1 = 1.5$
3. $x_2 = -6$
4. $x_3 = -0.375$
5. $x_4 = -1.263158$
6. $x_5 = -0.919355$
7. $x_6 = -1.02762$
8. $x_7 = -0.990876$
9. $x_8 = -1.00305$

Converge, but slower

Case c

$$x_{i+1} = \frac{x_i^2 - 3}{2}$$

1. $x_0 = 4$
2. $x_1 = 6.5$
3. $x_2 = 19.625$
4. $x_3 = 191.070$

Diverge!

s/n	**Root finding method**	**f(a)f(b)<0 assumption**	**2 initial point**	**1 initial point**	**Class**	**Derivatives needed**	**Formula**
1	**Bisection**	Yes	Yes		Bracket		x3 = (x1 + x2) / 2
2	**Regua Falsi**	Yes	Yes		Bracket		$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)}$
3	**Newton Raphson**			Yes	Open	Yes	$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$
4	**Secant**		Yes		Open		$x_3 = x_2 - f(x_2)\frac{x_2 - x_1}{f(x_2) - f(x_1)}$
5	**Modified Secant**			Yes	Open	Yes	x1 = x – [f(x) – f ’(x)]
6	**Fixedpoint Iteration**			Yes	Open		$x_{i+1} = g(x_i)$

COVENANT UNIVERSITY

CANAANLAND, KM 10, IDIROKO ROAD
P.M.B 1023, OTA, OGUN STATE, NIGERIA.

TITLE OF EXAMINATION: TEST 1
COLLEGE: College of Science and Technology
DEPARTMENT: Department of Computer and Information Sciences
SESSION: 2023/2024 **SEMESTER:** ALPHA
COURSE CODE: CSC431 **CREDIT UNIT:** 3
COURSE TITLE: Computational Science and Numerical Methods
INSTRUCTION: Answer ALL questions **TIME:** 1 HOUR

Question 1 – (24mks: each iteration of each method carries 2 marks)

Given a floating ball with specific gravity 0.6N, radius 5.5cm, and equation for depth x to which the ball is submerged under water: $x^3 - 0.165x^2 + 3.993 \times 10^{-4} = 0$

The ball will be submerged between x = 0 and x = 2R where R is the radius of the ball.

Hint: 1 Newton = 1 kg.m/s², so 5.5cm = 0.055m, $0 \leq x \leq 2R \Rightarrow 0 \leq x \leq 2(0.055) \Rightarrow 0 \leq x \leq 0.11$.

Calculate the depth to which the ball is submerged when floating in water by finding the root of the equation **in 2 iterations** using each of the methods below.

s/n	Root finding method	f(a)f(b)<0 assumption	2 initial points	1 initial point	Type of Method	Derivatives Needed	Formula
A	**Bisection**	Yes	Yes		Bracket		x3 = (x1 + x2) / 2
B	**Regua Falsi**	Yes	Yes		Bracket		$x_3 = x_1 - f(x_1)\dfrac{x_2 - x_1}{f(x_2) - f(x_1)}$
C	**Newton Raphson**			Yes	Open	Yes	$x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)}$
D	**Secant**		Yes		Open		$x_3 = x_2 - f(x_2)\dfrac{x_2 - x_1}{f(x_2) - f(x_1)}$ for [x0, x1], [x1, x2], [x2, x3], …
E	**Modified Secant**			Yes	Open	Yes	x1 = x – [f(x) – f '(x)]
F	**Fixedpoint Iteration**			Yes	Open		$x_{i+1} = g(x_i)$

Question 2 – (6mks)

A. Give at least two (2) pros and cons each of the following methods of solving non-linear equations: Bisection method, Newton-Raphson's method and Secant method. [**3 marks]**

B. Discuss the relevance of the course "computational science and numerical methods" (give at least 3 application areas). **[3 marks]**

Sample solution

Question 1

For a function $f(x) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$ with interval ($x_1 = 0$, $x_2 = 0.11$)

Bisection method **[4mks]**

$x_2 = \frac{x_0+x_1}{2}$ for $f(x_0) \times f(x_1) < 0$

1st iteration

$f(x0) = f(0) = 0^3 - 0.165 \times 0^2 + 3.993 \times 10^{-4} = 0.0003993 > 0$

$f(x1) = f(0.11) = 0.11^3 - 0.165 \times 0.11^2 + 3.993 \times 10^{-4}$

$f(x1) = f(0.11) = 0.001331 - 0.0019965 + 0.0003993 = -0.0002662 < 0$

$x_2 = \frac{x0+x1}{2} = \frac{0+0.11}{2} = 0.055$

2nd iteration

$f(x0) = f(0) = 0.0003993 > 0$

$f(x1) = f(0.11) = -0.0002662 < 0$

$f(x2) = f(0.055) = 0.055^3 - 0.165 \times 0.055^2 + 3.993 \times 10^{-4}$

$f(x2) = f(0.055) = 0.000166375 - 0.000499125 + 0.0003993 = -0.00006655 < 0$

$x_3 = \frac{x_2+x_0}{2}$ for $f(x_2) \times f(x_0) < 0$

$x_3 = \frac{x2+x0}{2} = \frac{0.055+0}{2} = 0.0275$

Regula falsi method **[4mks]**

$x_2 = x_0 - [f(x_0)\ \frac{x_1-x_0}{f(x_1)-f(x_0)}]$ for $f(x_0) \times f(x_1) < 0$

1st iteration

$f(x0) = f(0) = 0.0003993 > 0$

$f(x1) = f(0.11) = -0.0002662 < 0$

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

$x_2 = 0 - [0.0003993\ \frac{0.11-0}{-0.0002662 - 0.0003993}]$

$x_2 = 0 - [0.0003993\ \frac{0.11}{-0.0006655}]$

$x_2 = 0 - [0.0003993 \times -165.2892562$

$x_2 = 0 - -0.066$

$x_2 = 0.066$

2nd iteration

$f(x0) = f(0) = 0.0003993 > 0$

$f(x1) = f(0.11)$ = -0.0002662 < 0

$f(x2) = f(0.066) = 0.066^3 - 0.165 \times 0.066^2 + 3.993 \times 10^{-4}$

$f(x2) = f(0.066) = 0.000287496 - 0.00071874 + 0.0003993$ = -0.000031944 < 0

$x_3 = x_0 - [f(x_0)\ \frac{x_2 - x_0}{f(x_2) - f(x_0)}\ ]$ for $f(x_0) \times f(x_2) < 0$

$x_3 = 0 - [0.0003993\ \frac{0.066 - 0}{-0.000031944 - 0.0003993}\ ]$

$x_3 = 0 - [0.0003993\ \frac{0.066}{-0.000431244}\ ]$

$x_3 = 0 - [0.0003993 \times -153.0456076$

$x_3 = 0 - -0.061111111$

$x_3 = 0.061111111$

Newton Raphson method **[4mks]**

$x_1 = x_0 - [\ \frac{f(x_0)}{f\prime(x_0)}\ ]$, with radius 0.055m

1st iteration

$f(x_0) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$

$f'(x_0) = 3x^2 - 0.165(2x)$

$f(x_0) = f(0.055)$ = -0.00006655

$f'(x_0) = f'(0.055) = 3 \times 0.055^2 - 0.165 \times 2 \times 0.055$

$f'(x_0) = f'(0.055) = 0.009075 - 0.01815$ = -0.009075

$x_1 = x_0 - [\ \frac{f(x_0)}{f\prime(x_0)}\ ] = 0.055 - [\ \frac{-0.00006655}{-0.009075}\ ] = [\ 0.055 + 0.00733333\ ]$

$x_1 = 0.062333333$

2nd iteration

$f(x_1) = f(0.062333333) = 0.062333333^3 - 0.165 \times 0.062333333^2 + 3.993 \times 10^{-4}$

$f(x_1) = f(0.06267085) = 0.000242193 - 0.0006410983 + 0.0003993$ = -0.0000003947

$f'(x_1) = f'(0.06267085) = 3 \times 0.062333333^2 - 0.165 \times 2 \times 0.062333333$

$f'(x_1) = f'(0.06267085) = 0.011656333 - 0.020569999$ = -0.008913666

$x_2 = x_1 - [\frac{f(x_1)}{f\prime(x_1)}] = 0.06267085 - [\frac{-0.0000003947}{-0.008913666}] = [0.06267085 + 0.00004428]$

$x_2 = 0.06271513$

Secant method **[4mks]**

$x_2 = x_1 - [f(x_1)\ \frac{x_1 - x_0}{f(x_1) - f(x_0)}]$ for [x0, x1], [x1, x2], [x2, x3], …

1st iteration

$f(x0) = f(0) = 0.0003993$

$f(x1) = f(0.11)$ = -0.0002662

$x_2 = x_1 - [f(x_1)\ \frac{x_1 - x_0}{f(x_1) - f(x_0)}]$ for [x0, x1]

$x_2 = 0.11 - [-0.0002662 \times \frac{0.11 - 0}{-0.0002662 - 0.0003993}]$

$x_2 = 0.11$ – [-0.0002662 $\times\ -165.2892562$]

$x_2 = 0.11 - 0.044$

$x_2 = 0.066$

2nd iteration

$f(x0) = f(0) = 0.0003993$

$f(x1) = f(0.11)$ = -0.0002662

$f(x2) = f(0.066)$ = -0.000031944

$x_3 = x_2 - [f(x_2)\ \frac{x_2 - x_1}{f(x_2) - f(x_1)}]$ for [x1, x2]

$x_3 = 0.066 - [0.00031944\ \frac{0.066 - 0.11}{-0.000031944 - 0.0002662}]$

$x_3 = 0.066 - [0.0003993\ \frac{0.044}{-0.000298144}]$

$x_3 = 0.066 - [0.0003993 \times -147.579693]$

$x_3 = 0.066 - -0.058928571$

$x_3 = 0.124$. This solution has diverged. Regula-falsi scheme does not have this problem as the associated sequence always converges.

Modified Secant method **[4mks]**

$x_1 = x_0 - [\ f(x_0) - f'(x_0)\ ]$, with radius 0.055m

1st iteration

$f(x_0) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$

$f'(x_0) = 3x^2 - 0.165(2x)$

$f(x_0) = f(0.055)$ = -0.00006655

$f'(x_0) = f'(0.055)$ = -0.009075

$x_1 = x_0 - [f(x_0) - 'f(x_0)]$ = $0.055 - [-0.00006655 - -0.009075]$ = $0.055 + 0.00900845$

$x_1 = 0.06400845$

2nd iteration

$f(x_1) = f(0.06400845) = 0.06400845^3 - 0.165 \times 0.06400845^2 + 3.993 \times 10^{-4}$

$f(x_1) = f(0.06400845) = 0.00026225 - 0.00067602 + 0.0003993$ = -0.00001447

$f'(x_1) = f'(0.06267085) = 3 \times 0.06400845^2 - 0.165 \times 2 \times 0.06400845$

$f'(x_1) = f'(0.06267085) = 0.012291245 - 0.021122788$ = -0.008831543

$x_2 = x_1 - [f(x_1) - 'f(x_1)]$ = $0.06267085 - [-0.00001447 - -0.008831543]$

x_2= $0.06267085 - 0.008686843$

$x_2 = 0.053984007$

Fixed-point iteration method **[4mks]**

$for \quad f(x \quad); \; x \quad = g(x \quad)$

if $\quad f(x) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$ with radius 0.055m

A possible x = g(x) = $[0.165x^2 - 3.993 \times 10^{-4}]^{1/3}$ **at point 0.055**

1st Iteration

x1 = g(x0) = g(0.055) = $\left[0.165 \times 0.055^2 - 3.993 \times 10^{-4}\right]^{1/3}$

= $[0.000499 - 0.0003993]^{1/3}$

= $[0.000099825]^{1/3}$

= 0.046388796

2nd Iteration

x1 = g(x0) = g(0.055) = $[0.165 \times 0.046388796^2 - 3.993 \times 10^{-4}]^{1/3}$

= $[0.0003551 - 0.0003993]^{1/3}$

= $[0.000044233]^{1/3}$

= 0.0353657

Sample solution - Question 2

A. Give at least two (2) pros and cons each of the following methods of solving non-linear equations: Bisection method, Newton-Raphson's method and Secant method. [**3 marks]**

Method	**Pros**	**Cons**
Bisection	• Easy, Reliable, Convergent • One function evaluation per iteration • No knowledge of derivative is needed	• Slow • Needs an interval $[a, b]$ containing the root, i.e., $f(a), f(b) < 0$
Newton-Raphson	• Fast (if near the root) • Two function evaluations per iteration	• May diverge • Needs derivative and an initial guess x0 such that $f'(x_0)$ is nonzero
Secant	• Fast (slower than Newton) • One function evaluation per iteration • No knowledge of derivative is needed	• May diverge • Needs two initial points guess x_0, x_1 such that $f(x_0) - f(x_1)$ is nonzero

[0.5 mk × any 6 correct points = 3 mks]

B. Discuss the relevance of the course "computational science and numerical method" (give at least 3 application areas). **[3 marks]**

Definition

Computational science is a field in mathematics that uses advanced computing capabilities to understand and solve complex problems. It is an area of science that spans many disciplines, but at its core, it involves the development of models and simulations to understand natural systems. It is often thought of as an integration of three disciplines- mathematics, computer science, and science. It also involves the invention, implementation, testing, and application of algorithms and software used to solve large-scale scientific and engineering problems

A **numerical method** is an approximate computer method for solving a mathematical problem which often has no analytical solution. It entails making use of computers to solve problems by step-wise, repeated and iterative solution methods, which would otherwise be tedious or unsolvable by hand-calculations.

Relevance

This course is multi-disciplinary and cuts across various fields such as computational biology, computational chemistry, computational physics, computational finance, economics and engineering. It helps to solve complex problems and optimize various processes across these disciplines.

- Modern scientists increasingly rely on computational modelling and data analysis to explore and understand the natural world. Given the ubiquitous use in science and its critical importance to the future of science and engineering, computational science plays a central role in progress and scientific developments in the 21st Century.

- It aims at educating the next generation of cross-disciplinary science students with the knowledge, skills, and values needed to pose and solve current and new scientific, technological and societal challenges.
- Computational science focuses on the development of predictive computer models of the world around us. As studies of physical phenomena evolved to address increasingly complex systems, traditional experimentation is often infeasible. Thus, the discipline entails the development of new methods that make challenging problems tractable on modern computing platforms, providing scientists and engineers with new windows into the world around us.

Application Areas

1. **Engineering:** It cuts across various engineering disciplines such as Civil engineering, Mechanical engineering, Electrical engineering, etc. For example, numerical methods in Civil Engineering are now used routinely in structural analysis to determine the member forces and moments in structural systems, prior to design.
2. **Optimization:** useful for solving optimization problems which requires trial through numerous iterations e.g., scheduling tasks on processors in a heterogeneous multiprocessor computing network.
3. **Algorithm Trading:** it helps to automate process. Basically, it factors in time, cost and volume to aid decision making.
4. **Computation Biology:** studies biological systems. For example, airflow patterns in the respiratory tract, transport and disposition of chemicals through the body, etc.
5. **Computational Physics:** study and analysis of physical problems through computation and modelling. It covers areas such as fluid dynamics, astrophysics, thermodynamics, electromagnetics, etc.
6. **Computational Chemistry**: entails the study and prediction of chemical reactions. Also entails the understanding of molecular structures and properties. For example, transport and disposition of chemicals, predicting the evolution of crystals growing in an industrial crystallizer, etc.
7. **Computational Finance:** which has gained popularity in recent times. It helps in making the best investments and covers various aspects of finance. For example, calculation of insurance risks and price of insurance.
8. **Computation Economics:** this is similar to Computational Finance but different, as this branch focuses on predictive economic models.

[0.5 mk × any 6 correct points = 3 mks]

Question 3- Removed

Write a Python program that accepts these inputs; order of an equation, coefficients of each x in the equation, and the gradient of the equation **[2.5mk].** The program then imports ‘symbols’ function from ‘sympy’ library and generate the symbolic representation of the equation like this: f_of_(x) =+ “…” **[2.5mks].** Finally, the program accepts a value for x, computes f_of_(x), and outputs the value of f_of_(x) for any given function [**1mk].**

```
poly_power = int(input("To what polynomial power is the function? "))
```

```
coeffs = []
for i in range(poly_power):
    co_ef = float(input(f"Enter coefficient value for power {poly_power - i}: "))
    coeffs.append(co_ef)
poly_gradient = float(input("What is the gradient of the function? "))

from sympy import symbols
x = symbols('x')
f_of_x = 0
for i in range(poly_power):
    power = poly_power - i
    f_of_x += coeffs[i] * x**power
f_of_x += poly_gradient
print("f_of_x = ", f_of_x)

x_val = float(input("Please enter a value for x: "))
y = f_of_x.subs(x, x_val)
print(f"f_of_x = {y:.7f}")
```

7 Methods of Interpolation

LINEAR INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2) and x.

$$y = y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1}$$

QUADRATIC INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), and x.

$$y = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

CUBIC INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), and x.

$$y = y_1 \frac{(x - x_2)(x - x_3)(x - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)} + y_2 \frac{(x - x_1)(x - x_3)(x - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)} + y_3 \frac{(x - x_1)(x - x_2)(x - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)} + y_4 \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)}$$

Nth INTERPOLATION: GENERAL SYSTEM OF INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$
\begin{aligned}
y \;=\; & y_1 \frac{(x-x_2)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)\dots(x_1-x_{n-1})(x_1-x_n)} \\
& + y_2 \frac{(x-x_1)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)\dots(x_2-x_{n-1})(x_2-x_n)} \\
& + y_3 \frac{(x-x_1)(x-x_2)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)\dots(x_3-x_{n-1})(x_3-x_n)} \\
& + y_{n-1} \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_n)}{(x_{n-1}-x_1)(x_{n-1}-x_2)(x_{n-1}-x_3)\dots(x_{n-1}-x_{n-2})(x_{n-1}-x_n)} \\
& + y_n \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)(x_n-x_3)\dots(x_n-x_{n-2})(x_n-x_{n-1})}
\end{aligned}
$$

LAGRANGE METHOD

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
x = 16	y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$
\begin{aligned}
y \;=\; & y_1 \frac{(x-x_2)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)\dots(x_1-x_{n-1})(x_1-x_n)} \\
& + y_2 \frac{(x-x_1)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)\dots(x_2-x_{n-1})(x_2-x_n)} \\
& + y_3 \frac{(x-x_1)(x-x_2)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)\dots(x_3-x_{n-1})(x_3-x_n)} + \dots \\
& + y_{n-1} \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_n)}{(x_{n-1}-x_1)(x_{n-1}-x_2)(x_{n-1}-x_3)\dots(x_{n-1}-x_{n-2})(x_{n-1}-x_n)} \\
& + y_n \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)(x_n-x_3)\dots(x_n-x_{n-2})(x_n-x_{n-1})}
\end{aligned}
$$

$$
\begin{aligned}
y \;=\; & f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x) + f(x_4)L_4(x) \\
& + \dots + f(x_{n-1})L_{n-1}(x) + f(x_n)L_n(x)
\end{aligned}
$$

$$
y = \sum_{i>0}^{n} f(x_i)L_i(x)
$$

NEWTON'S DIVIDED DIFFERENCE INTERPOLATION

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
x = 16	y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$y = f(x_1) + f(x_1,x_2)(x - x_1) + f(x_1,x_2,x_3)(x - x_1)(x - x_2) +$$
$$f(x_1,x_2,x_3,x_4)(x - x_1)(x - x_2)(x - x_3) + \dots$$

Such that:

The coefficients $f[x_i, x_{i+1}, \dots, x_j]$ are the divided differences obtained from the divided difference table

$f(x_1) = y_1$

$f(x_1,x_2) = \frac{y_2 - y_1}{x_2 - x_1}$ where $y_2 - y_1$ is the linear contribution based on the first two data points

$f(x_1,x_2,x_3) = \frac{(y_3 - y_2)(y_2 - y_1)}{x_3 - x_1}$ where $(y_3 - y_2)(y_2 - y_1)$ is the quadratic contribution based on the first three data points

$f(x_1,x_2,x_3,x_4) = \frac{(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)}{x_4 - x_1}$ where $(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)$ is the cubic polynomial contribution based on the first four data points

Newton's divided difference table

xi	fi/yi	f(xi, xj) for FDD	f(xi, xj, xk) for SDD	f(xi, xj, xk, xL) for TDD
x1	f1			
		$f(x_2,x_1) = \frac{f_2 - f_1}{x_2 - x_1}$		
x2	f2		$f(x_3,x_2,x_1) = \frac{(y_3 - y_2)(y_2 - y_1)}{x_3 - x_1}$	
		$f(x_3,x_2) = \frac{f_3 - f_2}{x_3 - x_2}$		$f(x_4,x_3,x_2,x_1) = \frac{(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)}{x_4 - x_1}$
x3	f3		$f(x_4,x_3,x_2) = \frac{(y_4 - y_3)(y_3 - y_2)}{x_4 - x_2}$	
		$f(x_4,x_3) = \frac{f_4 - f_3}{x_4 - x_3}$		$f(x_5,x_4,x_3,x_2) = \frac{(y_5 - y_4)(y_4 - y_3)(y_3 - y_2)}{x_5 - x_2}$
x4	f4		$f(x_5,x_4,x_3) = \frac{(y_5 - y_4)(y_4 - y_3)}{x_5 - x_3}$	
		$f(x_5,x_4) = \frac{f_5 - f_4}{x_5 - x_4}$		

Question 7 - Assignment

Given a table of Velocity (m/s) as function of Time (s) below:

(s)	10	15	16	20	22.5	30
(m/s)	227.04	362.78	??	517.35	602.97	901.67

Find the velocity at t=16 seconds using

A. The Lagrangian method of

i. Linear interpolation;

$$y = f(x_1)L_1(x) + f(x_2)L_2(x)$$
$$= y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1}$$

ii. Quadratic interpolation;

$$y = f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x)$$
$$= y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

iii. Cubic interpolation;

$$y = f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x) + f(x_4)L_4(x)$$
$$= y_1 \frac{(x - x_2)(x - x_3)(x - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)} + y_2 \frac{(x - x_1)(x - x_3)(x - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)} + y_3 \frac{(x - x_1)(x - x_2)(x - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)} + y_4 \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)}$$

B. The values from Newton's divided difference table for the coefficients in

i. First divided difference;

$$y = f(x_1) + f(x_1, x_2)(x - x_1)$$

ii. Second divided difference;

$$y = f(x_1) + f(x_1, x_2)(x - x_1) + f(x_1, x_2, x_3)(x - x_1)(x - x_2)$$

iii. Third divided difference;

$$y = f(x_1) + f(x_1, x_2)(x - x_1) + f(x_1, x_2, x_3)(x - x_1)(x - x_2) + f(x_1, x_2, x_3, x_4)(x - x_1)(x - x_2)(x - x_3)$$

xi	fi/yi	f(xi, xj) for FDD	f(xi, xj, xk) for SDD	f(xi, xj, xk, xL) for TDD
x1	f1			
		$f(x_2, x_1) = \frac{f_2 - f_1}{x_2 - x_1}$		
x2	f2		$f(x_3, x_2, x_1) = \frac{(y_3 - y_2)(y_2 - y_1)}{x_3 - x_1}$	
		$f(x_3, x_2) = \frac{f_3 - f_2}{x_3 - x_2}$		$f(x_4, x_3, x_2, x_1) = \frac{(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)}{x_4 - x_1}$
x3	f3		$f(x_4, x_3, x_2) = \frac{(y_4 - y_3)(y_3 - y_2)}{x_4 - x_2}$	
		$f(x_4, x_3) = \frac{f_4 - f_3}{x_4 - x_3}$		$f(x_5, x_4, x_3, x_2) = \frac{(y_5 - y_4)(y_4 - y_3)(y_3 - y_2)}{x_5 - x_2}$
x4	f4		$f(x_5, x_4, x_3) = \frac{(y_5 - y_4)(y_4 - y_3)}{x_5 - x_3}$	
		$f(x_5, x_4) = \frac{f_5 - f_4}{x_5 - x_4}$		

8 Methods of solving first order Ordinary Differential Equations

An Ordinary Differential Equation (ODE) is a mathematical equation that involves an unknown function of one variable and its derivatives with respect to that variable. The "ordinary" in Ordinary Differential Equation distinguishes it from Partial Differential Equations (PDEs), which involve partial derivatives with respect to multiple variables.

The general form of a first-order PDE is: $F(x,y,u,\frac{\partial x}{\partial u},\frac{\partial y}{\partial u})=0$. $u(x,y)$ is the unknown function of the independent variables x and y, and F is a given function that describes the relationship between x, y, u, and its partial derivatives.

A first-order ordinary differential equation (ODE) is a mathematical equation that involves an unknown function and its first derivative with respect to an independent variable. The general form of a first-order ODE is: $\frac{dy}{dx} = f(x,y)$ or F $(x,y,\frac{dy}{dx})$ = 0. y is the unknown function of the independent variable x, and $f(x,y)$ is a given function that defines the relationship between x, y, and their derivatives. The solution to a first-order ODE is a function $y(x)$ that satisfies the given differential equation. Solving a first-order ODE involves finding this function $y(x)$ based on the provided equation and any initial conditions that specify the value of y at a particular point x_0. First Order ODEs are commonly encountered in various scientific and engineering applications. They describe rates of change or relationships involving a single variable and its derivative.

Higher-order ODEs are also common in physics and engineering, where they often arise from the modeling of physical systems with multiple degrees of freedom. Solutions to higher-order ODEs involve finding a function $y(x)$ that satisfies the given differential equation and any initial or boundary conditions. Methods for solving higher-order ODEs can be more involved than those for first-order ODEs. Common techniques include reduction to a system of first-order ODEs, substitution methods, and numerical methods. The behavior and solutions of higher-order ODEs can be complex and depend on the specific form of the equation and the given conditions. A higher-order ordinary differential equation (ODE) is an ODE that involves derivatives of a function up to some order greater than one. The general form of a higher-order ODE is: F $(x,y,\frac{dy}{dx},\frac{dy2}{dx2},\frac{dy3}{dx3},\dots,\frac{dyn}{dxn})$ = 0. E.g. a second-order ODE is one where the highest derivative involved is the second derivative: F $(x,y,\frac{dy}{dx},\frac{dy2}{dx2})$ = 0

Reducing a higher-order ordinary differential equation (ODE) to a system of first-order ODEs is a common technique that simplifies the problem and allows for the use of standard methods for solving first-order ODEs. The reduction is typically achieved by introducing new variables to represent the derivatives of the original function. This reduction is known as the "system of first-order ODEs" method or the "vectorization" of the higher-order ODE. It is a powerful technique that allows the application of well-established methods for solving first-order ODEs to higher-order problems.

To reduce a general n-th order ODE to a system of first-order ODEs, we introduce n new variables $y1,y2, \dots, yn$ to represent the first n derivatives of y;

F $(x,y,\frac{dy}{dx},\frac{dy2}{dx2},\frac{dy3}{dx3},\dots,\frac{dyn}{dxn})$ = 0 :

$$y_1 = y, \quad y_2 = \frac{dy}{dx}, \quad y_3 = \frac{d^2y}{dx^2}, \quad \ldots, \quad y_n = \frac{d^ny}{dx^n}$$

Now, we have a system of first-order ODEs:

$$\begin{aligned}
\frac{dy_1}{dx} &= y_2 \\
\frac{dy_2}{dx} &= y_3 \\
\frac{dy_3}{dx} &= y_4 \\
&\vdots \\
\frac{dy_{n-1}}{dx} &= y_n \\
\frac{dy_n}{dx} &= F\left(x, y_1, y_2, \ldots, y_n\right)
\end{aligned}$$

This system of first-order ODEs can then be solved using numerical or analytical methods, and the solutions can be used to reconstruct the solution to the original higher-order ODE by integrating the equations.

0 INTRO TO COMPUTATIONAL SCIENCE AND NUMERICAL METHODS

Numerical analysis ---> Numerical methods -----> Computational numerical methods ----> Computational numerical analysis

Numerical Analysis

Numerical Analysis is the study of numerical methods. Numerical analysis finds application in all fields of engineering and the physical sciences, and in the 21st century also the life and social sciences, medicine, business and even the arts. The GOAL of numerical analysis is the design and analysis of techniques/METHODS to give approximate but accurate solutions to hard problems.

Numerical methods

Numerical methods are mathematical attempts at finding approximate solutions of problems rather than the exact ones.

Computational numerical methods

Before modern computers, numerical methods often relied on hand formulas, using data from large printed tables. Since the mid20th century, computers calculate the required functions instead, but many of the same formulas continue to be used in software algorithms.

Computational numerical analysis

Current growth in computing power has enabled the use of more complex numerical analysis, providing detailed and realistic mathematical models in science and engineering. Numerical analysis continues this long tradition: rather than giving exact symbolic answers translated into digits and applicable only to real-world measurements, approximate solutions within specified error bounds are used.

1 Methods of Approximations

- Rounding off to significant figures
- Rounding off to decimal places
    - Working with arithmetic precision

2 Methods of Errors

    - Sources of errors
- Rounding errors
- Inherent errors
- Truncation errors
- True errors
- Relative true errors
- Absolute errors
- Relative absolute errors
- Approximate errors
- Relative approximate errors
- Absolute relative errors
- Percentage errors
    - Propagation of errors

3 Methods of Drawing the Lines of best fit

- Linearization
- Least squares curve fitting
- Group averages grouped averages curve fitting

4 Methods of solving Linear Systems of Equations

- Gaussian elimination
- Gauss-Jordan elimination
- LU decomposition
- Jacobi iteration
- Gauss-Seidal iteration

5 Methods of finding the roots of Algebraic and Transcendental Equations

- Bisection
- Newton-Raphson
- Regula-falsi
- Secant
- Modified Secant
- Fixed point Iteration

6 Methods of Finite Differences

- First forward difference
- First backward difference
- First central difference

7 Methods of Interpolation

- Lagrange Interpolation; linear, quadratic, cubic, …
- Newton’s divided difference interpolation; first, second, third, …
- Newton’s forward interpolation formula
- Newton’s backward interpolation formula

8 Methods of Numerical Integration

- Newton-coates Quadrature
- Trapezoidal rule
- Simpson’s one-third rule
- Simpson’s three-eight rule
- Romberg’s method

9 Methods of Solving First Order Ordinary Differential Equations

- Picard’s Method
- Euler Method
- Modified Euler Method
- Runge-Kutta first order method
- Runge-Kutta second order method
- Runge-Kutta third order method
- Runge-Kutta fourth order method

1 Methods of Approximations

Question 1

Without calling any in-built library or function, write a new function from scratch to

a. round-off any number to a stated precision of decimal place
b. return the absolute value of any number
c. find the natural log of any number
d. approximate any number to its nearest_integer
e. take the magnitude of any number
f. round off any number to a stated amount of significant figures
g. re-write 'f' using built-in libraries and functions.

Solution to Question 1f

ALGORITHM - To round off numbers to certain amount of significant figures

```
Given any 'number', with the 'num_sig_figs' to approximate the number to
If 'number' == 0,
        Return 0.0
Else
        Take the absolute value of 'number'
        Take the natural log of the absolute value
        Take 'number' to its nearest_integer
        Take the magnitude of the natural log
        Calculate rounding_factor = 10** (num_sig_figs – magnitude - 1)
        Find rounded_number = nearest_integer / rounding_factor
        If 'number' > 0
                Return rounded_number
        If 'number' < 0
                Return –rounded_number
```

PSEUDOCODE - To round off numbers to certain amount of significant figures

Without calling any in-built library or function, this is a new function from scratch to round off numbers to certain amount of significant figures

```
def round_to_significant_figures(number, num_sig_figs):
    if number == 0:
        return 0.0

    # Calculate the absolute value of 'number'
    def absolute_value(number):
        if number < 0:
            return -number
        else:
            return number
    abs_number = absolute_value(number)

    # Calculate the natural logarithm of the absolute value
    def custom_ln(number, num_terms=100):
        if number == 1:
            return 0.0
        elif number < 1:
            number = 1 / number
            num_terms = -num_terms
        nat_log = 0.0
        for n in range(1, num_terms + 1):
            term = ((number - 1) ** n) / n
            if n % 2 == 0:
                nat_log -= term
            else:
                nat_log += term
        return nat_log

    # Calculate the magnitude of the natural log
```

```
    def custom_floor(nat_log):
        if nat_log >= 0:
            return int(nat_log)
        else:
            integer_part = int(nat_log)
            if integer_part == nat_log:
                return integer_part
            else:
                return integer_part - 1
    magnitude = custom_floor(custom_ln(abs_number))

    # Calculate rounding_factor
    rounding_factor = 10 ** (num_sig_figs - magnitude - 1)

    # Use rounding factor to round number to the specified significant figures
    def custom_roundoff(number):
        decimal_part = number - int(number)
        if decimal_part < 0.5:
            return int(number)
        else:
            return int(number) + 1
    rounded_number = custom_roundoff(abs_number * rounding_factor) / rounding_factor

    # Restore the sign
    if number > 0:
        return rounded_number
    else:
        return -rounded_number

#round_to_significant_figures(number, num_sig_figs)
```

Arithmetic precision

It might be necessary to round off numbers to make them useful for numerical computation, moreso as it would require an infinite computer memory to store an unending number.

The precision of a number is an indication of the number of digits that have been used to express it. In scientific computing, it is the number of significant digits or numbers, while in management and financial systems, it is the number of decimal places. We are quite aware that most currencies in the world are quoted to two decimal places.

Arithmetic precision (often referred to simply as precision) is the specified number of significant figures or digits to which the number of interest is to be rounded.

2 Methods of errors

Rounding Errors

These are errors incurred by truncating a sequence of digits representing a number, as we saw in the case of representing the rational number 3/7 by 2.3333, instead of 2.3333….., which is an unending number. Apart from being unable to write this number in an exact form by hand, our instruments of calculation, be it the calculator or the computer, can only handle a finite string of digits. Rounding errors can be reduced if we change the calculation procedure in such a way as to avoid the subtraction of nearly equal numbers or division by a small number. It can also be reduced by retaining at least one more significant figure at each step than the one given in the data, and then rounding off at the last step.

Inherent Errors

As the name implies, these are errors that are inherent in the statement of the problem itself. This could be due to the limitations of the means of calculation, for instance, the calculator or the computer. This error could be reduced by using a higher precision of calculation.

Truncation Errors

If we truncate Taylor's series, which should be an infinite series, then some error is incurred. This is the error associated with truncating a sequence or by terminating an iterative process. This kind of error also results when, for instance, we carry out numerical differentiation or integration, because we are replacing an infinitesimal process with a finite one. In either case, we would have required that the elemental value of the independent variable tend to zero in order to get the exact value.

Absolute Error, Relative True Error, Relative Approximate Error and Percentage Error etc.

Question 2

a. A student measured the length of a string of actual length 72.5 cm as 72.4 cm.
   i. Calculate the absolute error and the percentage error
   ii. Write a function that accepts measured length and actual length to output absolute error and the percentage error.

b. The derivative of a function f(x) can be approximated by the equation

   $$f'(x) \approx \frac{f(x+h) - f(x)}{h}, \qquad \text{if } f(x) = 7e^{0.5x}, \text{ and } h = 0.3,$$

   i. Find the true value, the approximate value, true error, and relative error of f'(2)
   ii. If true values are not known or are very difficult to obtain,
   then Approximate error (E_a) = Present Approximation – Previous approximation.
   For f(x) = $7e^{0.5x}$ at x = 2, find
      - f'(2) using h = 0.3
      - f'(2) using h = 0.15

- approximate error, and relative approximate error of f'(2)

iii. Write a function that takes in any value of x for the derivative of a function f(x) approximated by the equation f'(x) = [f(x + h) - f(x)] / h for f(x) = 7e^0.5x, h1 = 0.3, h2 = 0.15, and returns true value, approximate value, true error, relative error, approximate error, and relative approximate error of f'(x)

Solution to Question 2a

(2ai)

Absolute error = | actual value – measured value |
Relative absolute error = | actual value – measured value | / actual value
The percentage error = Relative absolute error x 100
Absolute error = | 72.5 – 72.4 |= 0.01.
The percentage error = (0.1 / 72.5) x 100 = 0.1379

(2aii)

```
def calculate_errors(actual_length, measured_length):
    absolute_error = abs(actual_length - measured_length)
    relative_error = absolute_error / actual_length
    percentage_error = relative_error * 100

    return {
        "Absolute Error": absolute_error,
        "Relative Absolute Error": relative_error,
        "Percentage Error": percentage_error
    }

calculate_errors(actual_length, measured_length)
```

Solution to Question 2b

(2bi)
Approximate value of f'(x), for x = 2, and h = 0.3

$$f'(2) \approx \frac{f(2+0.3) - f(2)}{0.3}$$

$$= \frac{f(2.3) - f(2)}{0.3}$$

$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$

$$= \frac{22.107 - 19.028}{0.3} \quad = 10.263$$

(2bi cont'd)
The exact or true value of f'(2) can be found by using our knowledge of differential calculus

$$f(x) = 7e^{0.5x}$$

$$f'(x) = 7 \times 0.5 \times e^{0.5x}$$

$$= 3.5e^{0.5x}$$

$$f'(2) = 3.5e^{0.5(2)}$$

$$= 9.5140$$

(2bi cont'd)

True Error = True Value – Approximate Value
E_t = 9.5140 – 10263 = -0.722

Relative true error = (True value – Approximate value) / True value

= (9.5140 – 10263) / 9.5140= -0.722 / 9.5140

(2bii)

For x = 2, and h = 0.3

- Approximate value of f'(x) = 10.263

For x = 2, and h = 0.15

- Approximate value of f'(x) =

$$f^{'}(2) \approx \frac{f(2+0.15)-f(2)}{0.15}$$

$$= \frac{f(2.15)-f(2)}{0.15}$$

$$= \frac{7e^{0.5(2.15)} - 7e^{0.5(2)}}{0.15}$$

$$= \frac{20.50-19.028}{0.15} = 9.8800$$

Approximate error (E_a) = Present Approximation – Previous approximation

= 9.8800 – 10.263

= -0.38300

Relative approximate error = Approximate error / Previous approximation

$$= \frac{-0.38300}{9.8800} = -0.038765$$

(Question 2biii)

```
import math
def f(x):
    return 7 * math.exp(0.5 * x)
def derivative_of_f(x):
    return 3.5 * math.exp(0.5 * x)

def calculate_derivative_error(x, first_h, f, second_h=None):
    true_value = derivative_of_f(x)
    first_approximate_value = (f(x + first_h) - f(x)) / first_h
    true_error = abs(true_value - first_approximate_value)
    relative_true_error = true_error / true_value
    if second_h is not None:
        second_approximate_value = (f(x + second_h) - f(x)) / second_h
        approximate_error = abs(second_approximate_value - first_approximate_value)
        relative_approximate_error = approximate_error / second_approximate_value
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "Second Approximate Value": second_approximate_value,
            "Approximate Error": approximate_error,
            "Relative Approximate Error": relative_approximate_error
        }
    else:
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "True Error": true_error,
            "Relative True Error": relative_true_error
        }
#calculate_derivative_error(x, first_h, f)
#calculate_derivative_error(x, first_h, f, second_h)
```

Propagation of errors

In numerical methods, the calculations are not made with exact numbers. How do these inaccuracies propagate through the calculations?

Question 2c

Find the bounds for the propagation in adding two numbers. For example if one is calculating $X+Y$ where

$$X = 1.5 \pm 0.05$$

$$Y = 3.4 \pm 0.04$$

Solution

Maximum possible value of $X = 1.55$

Maximum possible value of $Y = 3.44$

Maximum possible value of $X + Y = 1.55 + 3.44 = 4.99$

Minimum possible value of $X = 1.45$.

Minimum possible value of $Y = 3.36$.

Minimum possible value of $X + Y = 1.45 + 3.36 = 4.81$

Hence

$$4.81 \leq X + Y \leq 4.99.$$

Propagation of Errors In Formula

If f is a function of several variables $X_1, X_2, X_3, \ldots\ldots, X_{n-1}, X_n$

then the maximum possible value of the error in f is

$$\Delta f \approx \left|\frac{\partial f}{\partial X_1}\Delta X_1\right| + \left|\frac{\partial f}{\partial X_2}\Delta X_2\right| + \ldots\ldots + \left|\frac{\partial f}{\partial X_{n-1}}\Delta X_{n-1}\right| + \left|\frac{\partial f}{\partial X_n}\Delta X_n\right|$$

Question 2d

The strain in an axial member of a square cross-section is given by

$$\in = \frac{F}{h^2 E}$$

Given F = 72

h = 4 x 10^{-3}

E = 70 x 10^{9}

Find the maximum possible error in the measured strain.

Solution

$$\in = \frac{72}{(4\times10^{-3})^2(70\times10^9)}$$
$$= 64.286\times10^{-6}$$
$$= 64.286\mu$$

$$\Delta\in = \left|\frac{\partial\in}{\partial F}\Delta F\right| + \left|\frac{\partial\in}{\partial h}\Delta h\right| + \left|\frac{\partial\in}{\partial E}\Delta E\right|$$

$$\frac{\partial\in}{\partial F} = \frac{1}{h^2E} \qquad \frac{\partial\in}{\partial h} = -\frac{2F}{h^3E} \qquad \frac{\partial\in}{\partial E} = -\frac{F}{h^2E^2}$$

Thus

$$\Delta E = \left|\frac{1}{h^2E}\Delta F\right| + \left|\frac{2F}{h^3E}\Delta h\right| + \left|\frac{F}{h^2E^2}\Delta E\right|$$
$$= \left|\frac{1}{(4\times10^{-3})^2(70\times10^9)}\times0.9\right| + \left|\frac{2\times72}{(4\times10^{-3})^3(70\times10^9)}\times0.0001\right|$$
$$+ \left|\frac{72}{(4\times10^{-3})^2(70\times10^9)^2}\times1.5\times10^9\right|$$

Hence

$\in = (64.286\mu \pm 5.3955\mu)$

Question 2e

Subtraction of numbers that are nearly equal can create unwanted inaccuracies. Using the formula for error propagation, show that this is true.

Solution

Let z = x – y, Then

$$|\Delta z| = \left|\frac{\partial z}{\partial x}\Delta x\right| + \left|\frac{\partial z}{\partial y}\Delta y\right|$$
$$= |(1)\Delta x| + |(-1)\Delta y|$$
$$= |\Delta x| + |\Delta y|$$

So the relative error or relative change is

$$\left|\frac{\Delta z}{z}\right| = \frac{|\Delta x| + |\Delta y|}{|x - y|}$$

Check: $x = 2 \pm 0.001$

$$y = 2.003 \pm 0.001$$
$$\left|\frac{\Delta z}{z}\right| = \frac{|0.001| + |0.001|}{|2 - 2.003|}$$
$$= 0.667$$

Percentage error = 66.67%

Taylor series

Some examples of common Taylor series

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

The general form of Taylor series is given as

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots$$

provided that all derivatives of f(x) are continuous and exist in the interval [x,x+h]

What does this mean in plain English?

As Archimedes would have said, "*Give me the value of the function at a single point, and the (first, second, and so on) values of all its derivatives at that single point, and I can give you the value of the function at any other point*"

Question 2f

Find the value of f(6) given that f(4) =125, f'(4) = 74, f''(4) = 30, f'''(4) = 6 and all other higher order derivatives of f(x) at x=4 are zero.

Solution

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + \cdots$$

x = 4

h = 6 – 4 = 2

Since the higher order derivatives are zero,

$$f(4+2) = f(4) + f'(4)2 + f''(4)\frac{2^2}{2!} + f'''(4)\frac{2^3}{3!}$$

$$f(6) = 125 + 74(2) + 30\left(\frac{2^2}{2!}\right) + 6\left(\frac{2^3}{3!}\right)$$

= 125 + 148 + 60 + 8

= 341

Note that to find f(6) exactly, we only need the value of the function and all its derivatives at some other point, in this case x = 4.

Error in Taylor series

The Taylor polynomial of order n of a function f(x) with (n+1) continuous derivatives in the domain [x,x+h] is given by

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2!}+\cdots+f^{(n)}(x)\frac{h^n}{n!}+R_n(x)$$

where the remainder is given by

$$R_n(x)=\frac{(x-h)^{n+1}}{(n+1)!}f^{(n+1)}(c)$$

Such that

$$x<c<x+h$$

that is, c is some point in the domain [x,x+h]

Derivation for Maclaurin Series for e^x

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots$$

The Maclaurin series is simply the Taylor series about the point x=0

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2!}+f'''(x)\frac{h^3}{3!}+f''''(x)\frac{h^4}{4}+f'''''(x)\frac{h^5}{5}+\cdots$$

$$f(0+h)=f(0)+f'(0)h+f''(0)\frac{h^2}{2!}+f'''(0)\frac{h^3}{3!}+f''''(0)\frac{h^4}{4}+f'''''(0)\frac{h^5}{5}+\cdots$$

Since f(x) = e^x, f'(x) = e^x, f''(x) = e^x, $f^n(x)$ = e^x, and $f^n(0)$ = e^0 = 1;

The Maclaurin series is then:

$$f(h)=(e^0)+(e^0)h+\frac{(e^0)}{2!}h^2+\frac{(e^0)}{3!}h^3\ldots$$

$$=1+h+\frac{1}{2!}h^2+\frac{1}{3!}h^3\ldots$$

Therefore

$$f(x)=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots$$

It can be seen that as the number of terms used increases, the error bound decreases and hence a better estimate of the function can be found.

Question 2g

How many terms would it require to get an approximation of e^1 within a magnitude of true error of less than $10^{-6.}$

Solution

Using (n + 1) terms of Taylor seriesgives error bound of

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c) \qquad x = 0, h = 1, f(x) = e^x$$

$$R_n(0) = \frac{(0-1)^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

$$= \frac{(-1)^{n+1}}{(n+1)!} e^c$$

Since

x < c < x+h

0 < c < 0+1

0 < c < 1

$$\frac{1}{(n+1)!} < |R_n(0)| < \frac{e}{(n+1)!}$$

So if we want to find out how many terms it would require to get an approximation of e^1 magnitude of true error of less than 10^6,

$$\frac{e}{(n+1)!} < 10^{-6}$$

$$(n+1)! > 10^6 e$$

$$(n+1)! > 10^6 \times 3$$

$$n \geq 9$$

So 9 terms or more are needed to get a true error less than 10^{-6}.

3 Drawing line of best fit

The process of fitting a curve to a set of data is called curve-fitting.

Linearisation

A nonlinear relationship can be linearised and the resulting graph analysed to bring out the relationship between variables.

y = ix + j ----------> linear or straight line graph, i=slope, j=intercept

y = ix^2 + jx + k --------> quadratic graph or curve

y = ix^n + jx + k: n >= 3 --------> polynomial or sinusoidal wave form graph

y = ie^x ---------------> ?? non-linear graph

y = $2\log_x i3$ --------> ?? non-linear graph

Remember:

ln(x) is the natural logarithm to the base 'e' ≈ 2.71828, often referred to simply as "log."

$\log_{10}(x)$ is the common logarithm to the base 10, often referred to simply as "log."

In mathematical notation, the distinction is clear:

ln(x) = $\log_e(x)$, where 'e' is the base of the natural logarithm.

log(x) = $\log_{10}(x)$ where 10 is the base of common logarithm.

Case 1: $y = ae^x$.

(i) We could take the logarithm of both sides to base e:

$$\ln y = \ln(ae^x) = \ln a + \ln e^x = x + \ln a,$$

since $\ln e^x = x$. Thus, a plot of $\ln y$ against x gives a linear graph with slope unity and a y-intercept of ln a.

(ii) We could also have plotted y against e^x. The result is a linear graph through the origin, with slope equal to a.

Case 2: $T = 2\pi\sqrt{\frac{l}{g}}$

We can write this expression in three different ways:

(i) $\ln T = \ln(2\pi) + \frac{1}{2}\ln\left(\frac{l}{g}\right) = \ln(2\pi) + \frac{1}{2}(\ln l - \ln g)$.

Rearranging, we obtain,

$$\ln T = \frac{1}{2}\ln l + \left(\ln(2\pi) - \frac{1}{2}\ln g\right)$$

writing this in the form $y = mx + c$, we see that a plot of $\ln T$ against $\ln l$ gives a slope of 0.5 and a $\ln T$ intercept of $\left(\ln(2\pi) - \frac{1}{2}\ln g\right)$. Once the intercept is read of the graph, you can then calculate the value of g.

(ii) $T = \frac{2\pi}{\sqrt{g}}\sqrt{l}$

A plot of T versus $\sqrt{l}$ gives a linear graph through the origin (as the intercept is zero). The slope of the graph is $\frac{2\pi}{\sqrt{g}}$, from which the value of g can be recovered.

Case 3: $N = N_0 e^{-\lambda t}$

The student can show that a plot of $\ln N$ versus t will give a linear graph with slope $-\lambda$, and $\ln N$ intercept is $\ln N_0$.

What other functions of N and t could you plot in order to get λ and N_0?

Case 4: $\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$

We rearrange the equation:

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$$

A plot of v^{-1} (y-axis) versus u^{-1} (x-axis) gives a slope of -1 and a vertical intercept of $\frac{1}{f}$.

Question 3a

A student obtained the following reading with a mirror in the laboratory.

u	10	20	30	40	50
v	-7	-10	-14	-15	-17

Linearise the relationship $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$. Plot the graph of v^{-1} versus u^{-1} and draw the line of best fit. Hence, find the focal length of the mirror. All distances are in cm.

Solution

u	v	1/u	1/v
10	-7	0.1	-0.14286
20	-10	0.05	-0.1
30	-14	0.033333	-0.07143
40	-15	0.025	-0.06667
50	-17	0.02	-0.05882

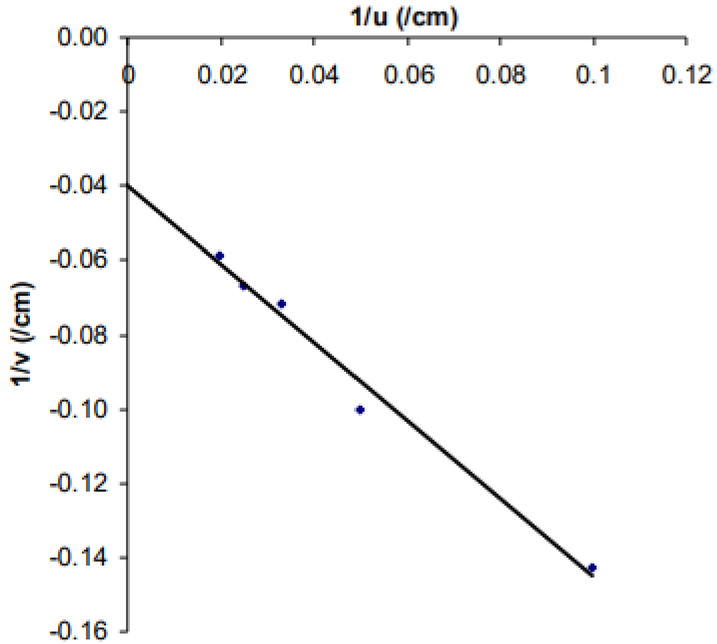


Fig. 1.1: Linear graph of the function $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$

The slope is -1.05 and the intercept -0.04. From $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$, we see that the intercept is $\frac{1}{f} = -0.04$, or $f = -\frac{1}{0.04} = -25$ cm.

Method of least squares curve fitting

The least square method entails minimizing the sum of the squares of the difference between the measured value and the one predicted by the assumed equation.

$$m = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2}$$

$$c = \bar{y} - m\bar{x}$$

Question 3b

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

Thus, for the following set of readings:

t	5	12	19	26	33
x	23	28	32	38	41

The table can be extended to give

t	5	12	19	26	33	Σ=95	$\bar{t}$ =19
x	23	28	32	38	41	Σ=162	$\bar{x}$ =32.4
tx	115	336	608	988	1353	Σ=3400	$\overline{tx}$=680
t^2	25	144	361	676	1089	Σ=2295	$\overline{x^2}$ =459

$$m = \frac{\overline{tx} - \bar{t}\,\bar{x}}{\overline{t^2} - \bar{t}^2} = \frac{680 - 19 \times 32.4}{459 - 19^2} = 0.6571$$
$$c = \bar{x} - m\bar{t} = 32.4 - 0.6571 \times 19 = 19.9151$$

Hence, the relationship between x and t is,

$$x = 0.6571t + 19.9151$$

Method of group averages curve fitting

As the name implies, a set of data is divided into two groups, each of which is assumed to have a zero sum of residuals.

$$\bar{y}_1 = m\bar{x}_1 + c$$
$$\bar{y}_2 = m\bar{x}_2 + c$$

Subtracting,

$$\bar{y}_1 - \bar{y}_2 = m(\bar{x}_1 - \bar{x}_2)$$
$$m = \frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1 - \bar{x}_2}$$

and

$$c = \bar{y}_1 - m\bar{x}_1$$

Question 3c

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

First divide the data into 2 groups

t	5	12	19
x	23	28	32

and

t	26	33
x	38	41

The tables can be extended to give, for Table 3:

t	5	12	19	Σ=36	$\bar{t}_1$=12
x	23	28	32	Σ=83	$\bar{x}_1$=27.666667

and for Table 4:

t	26	33	Σ=59	$\bar{t}_2$=29.5
x	38	41	Σ=79	$\bar{x}_2$=39.5

$$m = \frac{\bar{x}_1 - \bar{x}_2}{\bar{t}_1 - \bar{t}_2} = \frac{27.666667 - 39.5}{12 - 29.5} = 0.67619$$

and

$$c = \bar{x}_1 - m\bar{t}_1 = 27.666667 - (0.67619 \times 12)$$
$$= 19.552387$$

Thus, the equation of best fit is,

$$x = 0.67619t + 19.552387$$

4 Methods of Linear Systems of Equation

Let us consider a linear system of equations

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

This can be written in the form

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

Gaussian Elimination involving 2 variables

Question 4a: find x and y given that

2x + 3y = 13

x – y = -1

The augmented matrix representing our system of two equations is

$$\left[\begin{array}{cc|c} 2 & 3 & 13 \\ 1 & -1 & -1 \end{array}\right]$$

By Gaussian elimination, we seek to make every entry below the main diagonal zero. This we achieve by reducing 1 to zero, making use of the first row.

Thus,

$$5y = 15 \Rightarrow y = 3$$

Substituting this in the first row gives

$$2x + 3(3) = 13$$

from which we obtain $x = 2$.

The process of reducing every element below the main diagonal to zero (row echelon form) is called Gaussian Elimination. That of substituting obtained values to calculate other variables is called Back Substitution.

Gaussian Elimination involving 2 variables

The same process can be carried over to the case of a system of three equations.

Question 4b:

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

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

This yields (by Gaussian elimination)

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right] \begin{array}{c} \xrightarrow{(ii)'=(i)-2(ii)} \\ \xrightarrow{(iii)=(i)-(2/3)(iii)} \end{array} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 7/3 & 5/3 & 3 \end{array}\right]$$

$$\xrightarrow{(iii)''=(ii)'+(15/7)(iii)'} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

Upon back substitution,

$$-10z = 10 \text{ or } z = -1$$

$$z = -1;\ y + z = 1 \Rightarrow y = 2;\ 2x + y - z = 5 \Rightarrow x = 1$$

Traditionally, in Mathematics, it is usual to use indices such as x_1, x_2, etc. instead of x, y, z. Do you have any idea why this is so? It is because if we stay with the alphabets, we shall soon run out of symbols. Bear in mind that not all the alphabets can be employed as variables; as an example, a, b, c are commonly used as constants. In addition, it makes it easy to associate the coefficients a_{11}, a_{12}, etc. with x_1, x_2, etc. respectively. More importantly in numerical work, it makes programming easier. For instance for our system of three equations, we could use the more general notation:

The general Gaussian elimination for linear system of 3 variables is thus given as:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] \xrightarrow[(iii)'=a_{11}(i)-a_{11}(iii)]{(ii)'=a_{12}(i)-a_{11}(ii)} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & a_{32}' & a_{33}' & a_{34}' \end{array}\right]$$

$$\xrightarrow{(iii)''=a_{32}(ii)'+a_{22}(iii)'} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & 0 & a_{33}'' & a_{34}'' \end{array}\right]$$

Question 4b-2:

$$\begin{aligned} -3x1 + 2x2 - x3 &= -1, \\ 6x1 - 6x2 + 7x3 &= -7, \\ 3x1 - 4x2 + 4x3 &= -6. \end{aligned}$$

First write out the augmented matrix

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 6 & -6 & 7 & -7 \\ 3 & -4 & 4 & -6 \end{pmatrix}$$

Perform row reduction by multiplying the first row by 2 (the lcm of all x1's), then add first row to both second and third row

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & -2 & 3 & -7 \end{pmatrix}$$

Perform row reduction by multiplying the second row by -1 (the lcm of x2's in rows 2 and 3), then add second row to third row

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

$$\begin{aligned} &-2x3 = 2 \rightarrow x3 = -1 \\ &-2x2 = -9 - 5x3 = -4 \rightarrow x2 = 2, \\ &-3x1 = -1 - 2x2 + x3 = -6 \rightarrow x1 = 2. \end{aligned}$$

Therefore, we have:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_n \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$$

Gauss-Jordan Elimination

This entails eliminating in addition to the entries below the major diagonal, the entries above it, so that the main matrix is a diagonal matrix. In that case, the solution to the system is given by dividing the element in the augmented part of the matrix by the diagonal element for that row.

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

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

In solving the same problem using Gauss-Jordan elimination, we continue from completion of the Gaussian elimination part.

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right] \xrightarrow[(ii)'=(iii)-2(ii)]{(i)'=(iii)-10(i)} \left[\begin{array}{ccc|c} -20 & 10 & 0 & -40 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

$$\xrightarrow{(i)''=(ii)'+(i)'} \left[\begin{array}{ccc|c} -20 & 0 & 0 & -20 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

It follows that,

$-20x = -20$ or $x = 1$; $-10y = 20$ or $y = 2$; and $-10z = 10$ or $z = -1$

(lower and upper echelon) - LU decomposition

Suppose we could write the matrix

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

This implies that

$$l_{11}u_{11} = a_{11},\ l_{11}u_{12} = a_{12},\ l_{11}u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11},\ a_{22} = l_{21}u_{12} + l_{22}u_{22},\ a_{23} = l_{21}u_{13} + l_{22}u_{23}$$
$$a_{31} = l_{31}u_{11},\ a_{32} = l_{31}u_{12} + l_{32}u_{22},\ a_{33} = l_{31}u_{13} + l_{32}u_{23} + l_{33}u_{33}$$

Without loss of generality, we could set the diagonal elements of the L matrix equal to 1.

Then,

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

Multiplying out the right side of equation 3.19,

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

From the equality of matrices, this requires that,

$$u_{11} = a_{11}$$

$$u_{12} = a_{12}$$
$$u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11} \qquad \Rightarrow l_{21} = a_{21}/u_{11} = a_{21}/a_{11}$$
$$a_{31} = l_{31}u_{11} \qquad \Rightarrow l_{31} = a_{31}/u_{11} = a_{31}/a_{11}$$
$$a_{22} = l_{21}u_{12} + u_{22} \text{ , or } u_{22} = a_{22} - l_{21}u_{12} = a_{22} - \frac{a_{21}}{u_{11}}u_{12}$$
$$\Rightarrow u_{22} = a_{22} - \frac{a_{21}}{a_{11}}a_{12}$$
$$a_{23} = l_{21}u_{13} + u_{23}\text{, or } u_{23} = a_{23} - l_{21}u_{13} = a_{23} - \frac{a_{21}}{u_{11}}u_{13}$$
$$\Rightarrow u_{23} = a_{23} - \frac{a_{21}}{a_{11}}a_{13}$$
$$l_{32} = \frac{a_{32} - l_{31}u_{12}}{u_{22}} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{u_{11}}u_{12}\right]$$
$$\Rightarrow l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}}a_{12}\right]$$
$$a_{32} = l_{31}u_{12} + l_{32}u_{22}$$
$$a_{33} = l_{31}u_{13} + l_{32}u_{23} + u_{33}$$
$$\Rightarrow u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23}$$

You can see that we have determined all the nine elements of the two matrices in terms of the elements of the original matrix.

Once we have obtained L and U, then we can write the original equation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

as

$$LU\mathbf{x} = \mathbf{y}$$

where $\mathbf{x}$ and $\mathbf{y}$ are column vectors.

We shall write $\mathbf{w} = U\mathbf{x}$

Then,

$$L\mathbf{w} = \mathbf{y}$$

Now we continue to solving **Question5b** again using LU decomposition

The corresponding matrix is

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

$$u_{11} = a_{11} = 2$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = -1$$
$$l_{21} = a_{21} / a_{11} = 1/2$$
$$l_{31} = a_{31} / a_{11} = 3/2$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 3 - \frac{1}{2}(1) = 5/2$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - \frac{1}{2}(-1) = 2 + \frac{1}{2} = 5/2$$
$$l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{5/2}\left[-2 - \frac{3}{2}(1) \right] = -7/5$$
$$u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = -4 - (3/2)(-1) - (-7/5)(5/2) = 1$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 5/2 & 5/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 2 \\ 3 & -2 & -4 \end{bmatrix}$$

The above decomposition is correct as the multiplication of L and U gives the original matrix.

The original equation is equivalent to

$$LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$$

$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 3 \end{bmatrix}$$

Solving,

$$w_1 = \quad 5$$

$$\frac{1}{2}w_1 + w_2 = 5 \text{ or } w_2 = 5 - \frac{1}{2}w_1 = 5 - \frac{1}{2}(5) = \frac{5}{2}$$

$$\frac{3}{2}w_1 - \frac{7}{5}w_2 + w_3 = 3\text{, or } w_3 = 3 + \frac{7}{5}w_2 - \frac{3}{2}w_1 = 3 + \frac{7}{5}\left(\frac{5}{2}\right) - \frac{3}{2}(5) = -1$$

$U\mathbf{x} = \mathbf{w}$ implies:

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

By back substitution,

$$x_3 = -1$$

$$\frac{5}{2}x_2 + \frac{5}{2}x_3 = \frac{5}{2} \Rightarrow \frac{5}{2}x_2 = \frac{5}{2} - \frac{5}{2}x_3 = \frac{5}{2} - \frac{5}{2}(-1) = 5$$

$$x_2 = 2$$

$$2x_1 + x_2 - x_3 = 5$$

$$x_1 = \frac{5 - x_2 + x_3}{2} = \frac{5 - 2 + (-1)}{2} = 1$$

The solution set is therefore,

$$x_1 = 1,\ y = 2,\ z = -1.$$

Question 4c

Solve the system of linear equations $x + y + z = -1$, $x + 2y + 2z = -4$, $9x + 6y + z = 7$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

(i) Gaussian elimination

Initial augmented matrix

1	1	1	-1
1	2	2	-4
9	6	1	7

First round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	-3	-8	16

Second round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

Last matrix for Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

First round of Jordan elimination

5	5	0	2
0	5	0	-8
0	0	-5	7

Second round of Jordan elimination

-25	0	0	-50
0	5	0	-8
0	0	-5	7

(iii) LU decomposition

$x + y + z = -1$

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

$9x + 6y + z = 7$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7 \text{, or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$$x_3 = -7/5 = -1.4$$

$$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$$

$$x_2 = -8/5 = -1.6$$

$$x_1 + x_2 + x_3 = -1$$

$$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \frac{10}{5} = 2$$

The solution set is therefore,

$$x_1 = 2,\ y = -1.6,\ z = -1.4.$$

Question 4d

Solve the system of linear equations $x + 2y + 2z = -2$, $2x + 2y + z = -4$, $9x + 6y + 2z = -14$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

Gaussian elimination

Initial augmented matrix

1	2	2	-2
2	2	1	-4
9	6	2	-14

First round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	-12	-16	4

Second round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

Answers

x	0
y	-3
z	2

Gauss-Jordan elimination

Last matrix for Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

First round of Jordan elimination

4	8	0	-24
0	-8	0	24
0	0	-4	-8

Second round of Jordan elimination

32	0	0	0
0	-8	0	24
0	0	-4	-8

LU decomposition

$$x + 2y + 2z = -4$$
$$9x + 6y + z = 7$$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

W got the decomposition right, as the multiplication of the L and U gives the original matrix.

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7\text{, or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$$x_3 = -7/5 = -1.4$$
$$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$$
$$x_2 = -8/5 = -1.6$$
$$x_1 + x_2 + x_3 = -1$$
$$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \frac{10}{5} = 2$$

The solution set is therefore,

$$x_1 = 2\text{, } y = -1.6\text{, } z = -1.4\text{.}$$

It is usually a good practice to revert to fractions to avoid incurring rounding errors.

5 Methods for finding Roots of Algebraic and Transcendental equations

In all scientific fields, there's always the need to find the root of an equation, equivalently the zero of a function. Numerical methods allow for more complicated cases of handling roots of quadratic and polynomial equations.

Bisection method

As the name implies, we obtain the points x_1 and x_2, such that $f(x_2)\, f(x_1) < 0$, meaning that the value of f has opposite signs at the two points, which points to the fact that a root exists between x_1 and x_2. We approximate this root by the average of the two, i.e., $(x_1 + x_2) / 2$. Let this be x_3. Then we evaluate $f(x_3)$. x_3 is then combined with x_1 or x_2, depending on the one at which the sign of the function is opposite $f(x_3)$. This gives x_4. This process is repeated until f(x) attains the prescribed tolerance. The convergence of the Bisection method is slow and steady.

Bisection Method

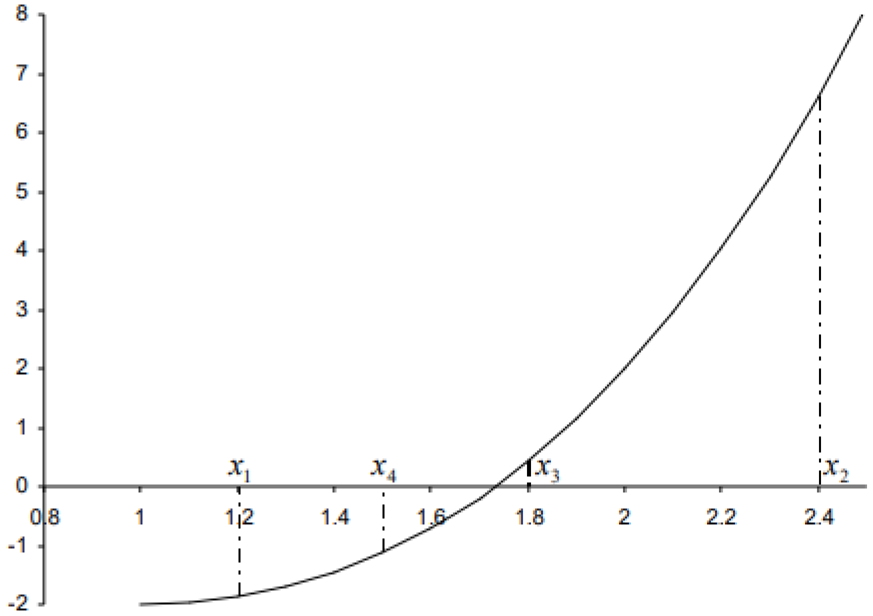


Fig.

- The **Bisection method** is one of the simplest methods to find a zero of a nonlinear function.
- It is also called **interval halving** method.
- To use the Bisection method, one needs an initial interval that is known to contain a zero of the function.
- The method systematically reduces the interval. It does this by dividing the interval into two equal parts, performs a simple test and based on the result of the test, half of the interval is thrown away.
- The procedure is repeated until the desired interval size is obtained.

Bisection Algorithm

Assumptions:

- $f(x)$ is continuous on $[a,b]$
- $f(a)\, f(b) < 0$

Loop

1. Compute the mid point $c=(a+b)/2$
2. Evaluate $f(c)$
3. If $f(a)\, f(c) < 0$ then new interval $[a, c]$
   If $f(a)\, f(c) > 0$ then new interval $[c, b]$

End loop

Question 5a

Can you use Bisection method to find a zero of:

$f(x) = x^3 - 3x + 1$ in the interval [0,2]?

Answer:

$f(x)$ is continuous on $[0,2]$

and $\mathrm{f}(0) * \mathrm{f}(2) = (1)(3) = 3 > 0$

$\Rightarrow$ Assumptions are not satisfied

$\Rightarrow$ Bisection method can not be used

Question 5b

Can you use Bisection method to find a zero of:

$f(x) = x^3 - 3x + 1$ in the interval [0,1]?

Answer:

$f(x)$ is continuous on $[0,1]$

and $\mathrm{f}(0) * \mathrm{f}(1) = (1)(-1) = -1 < 0$

$\Rightarrow$ Assumptions are satisfied

$\Rightarrow$ Bisection method can be used

Question 5c

Find the root of:

$$f(x) = x^3 - 3x + 1 \text{ in the interval } : [0,1]$$

* $f(x)$ is continuous

* $f(0) = 1,\ f(1) = -1 \Rightarrow f(a)\, f(b) < 0$

$\Rightarrow$ Bisection method can be used to find the root

Iteration	a	b	c= $\frac{(a+b)}{2}$	f(c)	$\frac{(b-a)}{2}$
1	0	1	0.5	-0.375	0.5
2	0	0.5	0.25	0.266	0.25
3	0.25	0.5	.375	-7.23E-3	0.125
4	0.25	0.375	0.3125	9.30E-2	0.0625
5	0.3125	0.375	0.34375	9.37E-3	0.03125

Question 5d

Find a zero of the function $f(x) = 2x^3 - 3x^2 - 2x + 3$ between the points 1.4 and 1.7, using the bisection method. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(1.4) = -0.192$$
$$f(1.7) = 0.756$$
$$x_3 = \frac{1.4 + 1.7}{2} = 1.55$$

$$f(1.55) = 1.4025 \times 10^{-1}$$

$$x_4 = \frac{1.55 + 1.4}{2} = 1.475$$

$$f(1.475) = -0.0588$$

$$x_5 = \frac{1.55 + 1.475}{2} = 1.5125$$

This confirm that the Table for Bisection method is indeed true

n	x	$f(x)$
1	1.55	0.14025
2	1.475	-5.88E-02
3	1.5125	3.22E-02
4	1.49375	-1.54E-02
5	1.503125	7.87E-03
6	1.498437	-3.89E-03
7	1.500781	1.96E-03
8	1.499609	-9.76E-04
9	1.500195	4.89E-04
10	1.499902	-2.44E-04
11	1.500049	1.22E-04
12	1.499976	-6.10E-05

CONVERGENCE ANALYSIS OF BISECTION METHOD

Bisection method obtains an interval that is guaranteed to contain a zero of the function.

Questions:

- What is the best estimate of the zero of $f(x)$?
- What is the error level in the obtained estimate?

The <u>best estimate</u> of the zero of the function $f(x)$ after the first iteration of the Bisection method is the mid point of the initial interval:

$$\textit{Estimate of the zero}: \ r = \frac{b+a}{2}$$

$$\textit{Error} \leq \frac{b-a}{2}$$

$$\textit{Given} \quad f(x),\ a,\ b,\ \textit{and}\ \ \varepsilon$$

How many iterations are needed such that: $|x - r| \leq \varepsilon$

where r is the zero of $f(x)$ and x is the bisection estimate (i.e., $x = c_k$)?

$$\boxed{n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)}}$$

$$\boxed{n \geq \log_2\left(\frac{\text{width of initial interval}}{\text{desired error}}\right) = \log_2\left(\frac{b-a}{\varepsilon}\right)}$$

Two common stopping criteria

1. Stop after a fixed number of iterations
2. Stop when the absolute error is less than a specified value

After n iterations :

$$\left| error \right| = \left| r - c_n \right| \leq E_a^n = \frac{b-a}{2^n} = \frac{\Delta x^0}{2^n}$$

Question 5e

$a = 6,\ b = 7,\ \varepsilon = 0.0005$

How many iterations are needed such that: $\left| x - r \right| \leq \varepsilon$?

$$n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1) - \log(0.0005)}{\log(2)} = 10.9658$$

$$\Rightarrow n \geq 11$$

Question 5f

- Use Bisection method to find a root of the equation x = cos(x) with absolute error <0.02 (assume the initial interval [0.5, 0.9])

What is *f (x)* ?

Are the assumptions satisfied ?

How many iterations are needed ?

How to compute the new estimate ?

Question 5f (i) – what is f(x)?

x = cos(x)

f(x) = x – cos(x)

Question 5f (ii) – Are the assumptions satisfied?

Assuming interval [0.5, 0.9]

f(0.5) = 0.5 – cos(0.5) = -0.3776; This is a negative value

f(0.9) = 0.9 – cos(0.9) = 0.2784; This is a positive value

f(0.5)*f(0.9) = -0.3776 * 0.2764 < 0; Assumption is therefore satisfied.

Bisection method can be used.

Question 5f (iii) – How many iterations are needed?

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

a = 0.5, b = 0.9, ε = 0.02

n $\geq$ [log(0.9 – 0.5) – log(0.02)] / log(2)

n $\geq$ [-0.3979 - -1.6990] / 0.3010

n $\geq$ 1.3011 / 0.3010

n $\geq$ 4.3226

n $\geq$ 5

Question 5f (iii) – How to compute the new estimate?

$$Estimate\ of\ the\ zero:\ r=\frac{b+a}{2}\ ,\quad Error\leq\frac{b-a}{2}$$

r1 = (0.9 + 0.5) / 2 = 0.7; **Error < (0.9 - 0.5)/2 $\leq$ 0.2;**

f(0.7) = 0.7 – cos(0.7) = 0.7 – 0.9999 = -0.2999

f(0.5) = -0.3776; f(0.9) = 0.2784; f(0.7) = -0.2999

r2 = (0.7 + 0.9) / 2 = 0.8, **Error < (0.9 - 0.7) / 2 $\leq$ 0.1**

f(0.8) = 0.8 – cos(0.8) = 0.8 – 0.9999 = -0.1999

f(0.7) = -0.2999; f(0.9) = 0.2784; f(0.8) = -0.1999

r3 = (0.8 + 0.9) / 2 = 0.85, **Error < (0.9 - 0.8) / 2 $\leq$ 0.5**

f(0.85) = 0.85 – cos(0.85) = 0.85 – 0.9999 = -0.1499

f(0.8) = -0.1999; f(0.9) = 0.2784; f(0.85) = -0.1499

r4 = (0.85 + 0.9) / 2 = 0.875, **Error < (0.9 - 0.85) / 2 $\leq$ 0.025**

f(0.875) = 0.875 – cos(0.875) = 0.875 – 0.9999 = -0.1249

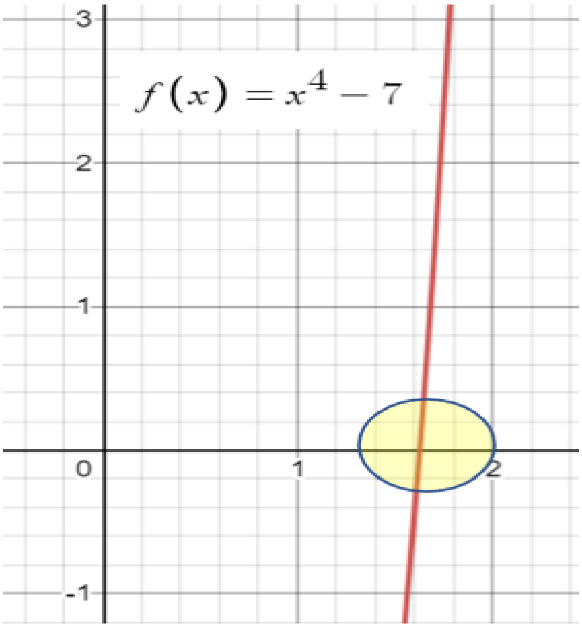
f(0.85) = -0.1499; f(0.9) = 0.2784; f(0.875) = -0.1249

r5 = (0.875 + 0.9) / 2 = 0.8875, **Error < (0.9 - 0.875) / 2 $\leq$ 0.02**

f(0.8875) = 0.8875 – cos(0.8875) = 0.8875 – 0.9999 = -0.1124

f(0.875) = -0.1249; f(0.9) = 0.2784; f(0.8875) = -0.1124

Question 5g



Find the 3rd approximation of the root of f(x) = x^4–7 using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(left)	f(mid)	f(right)	New Interval	Midpoint
f(1) = -6	f(1.5) = -2	f(2) = 9	(1.5, 2)	1.75
f(1.5) = -2	f(1.75) = 2.4	f(2) = 9	(1.5, 1.75)	1.625
f(1.5) = -2	f(1.625) = -.03	f(1.75) = 2.4	(1.625, 1.75)	1.6875

$f(x) = x^4 – 7$

$f(2) = (2)^4 – 7 = 9$; this is positive

$f(1) = (1)^4 – 7 = -6$; this is negative

$f(2)*f(1) = 9 * -6 < 0$; Assumption is therefore satisfied. Bisection method can be used.

for;

Starting interval (1, 2)

mid x = [2+1] / 2 = 1.5; Initial estimate

$f(mid) = f(1.5) = (1.5)^4 – 7 = 5.0625 – 7 = -1.9375$.

for;

f(2) = 9, f(1) = -6, **f(1.5) = -1.9375**

Next interval (2, 1.5)

mid x = [2+1.5]/2 = 1.75; first approximation

$f(mid) = f(1.75) = (1.75)^4 – 7 = 9.3789 – 7 = 2.3789$.

for;

f(2) = 9**,** **f(1.5) = -1.9375,** **f(1.75) = 2.3789**

Next interval (1.75, 1.5)

mid x = [1.75+1.5]/2 = 1.625; second approximation

$f(mid) = f(1.625) = (1.625)^4 – 7 = 6.9729 – 7 = -0.0271$.

for;

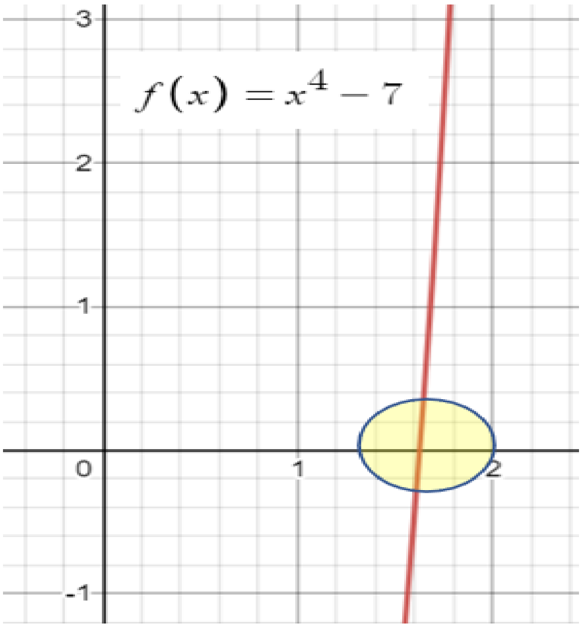
f(1.75) = 2.3789, f(1.5) = -1.9375, **f(1.625) = -0.0271**

Next interval (1.75, 1.625)

mid x = [1.75+1.625]/2 = 1.6875; third approximation

$f(mid) = f(1.6875) = (1.6875)^4 – 7 = 8.1091 – 7 = 1.1091$. **Stop.**

Question 5h



Find the 3rd approximation of the root of f(x) = 10 – x^2 using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(x) = 10 – x^2

> **f(2) = 10 - $(2)^2$ = 6; this is positive**
>
> **f(1) = 10 - $(1)^2$ = 9; this is also positive**
>
> **f(2)*f(1) = 6 * 9 !< 0;** Assumption is NOT satisfied. Bisection method cannot be used.

Question 5h

Given a floating ball with a specific gravity of 0.6 and has a radius of 5.5 cm. You are asked to find the depth to which the ball is submerged when floating in water.

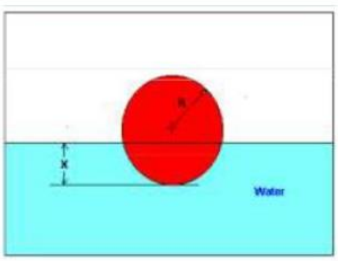


Diagram of the floating ball

The equation that gives the depth x to which the ball is submerged under water is given by:

$$x^3 - 0.165x^2 + 3.993x10^{-4} = 0$$

1. Use the Bisection method of finding roots of equations to find the depth x to which the ball is submerged under water. Conduct three iterations to estimate the root of the above equation.
2. Find the absolute relative approximate error at the end of each iteration.
3. Use both false position and newton methods to solve the roots of the equations.

Hint: From the Physics point of view, the ball would be submerged between $x = 0$ and $x = 2R$, where R = radius of the ball.

That is, $0 \leq x \leq 2R$ ==> $0 \leq x \leq 2(0.055)$ ==> $0 \leq x \leq 0.11$

Newton-Raphson Method

(Also known as Newton's Method)

Given an initial guess of the root x_0, Newton-Raphson method uses information about the function and its derivative at that point to find a better guess of the root.

Assumptions:

- $f(x)$ is continuous and the first derivative is known
- An initial guess x_0 such that $f'(x_0) \neq 0$ is given

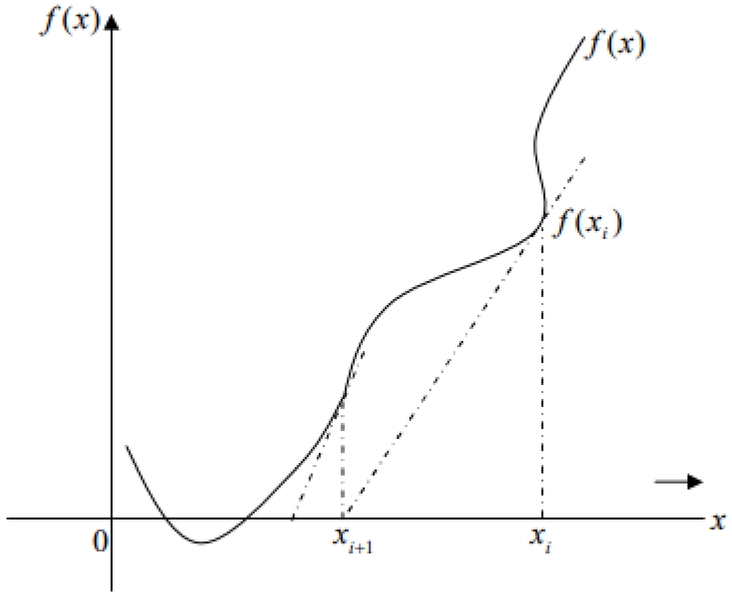
It is quite clear that the function $f(x)$ must be differentiable for you to be able apply the Newton-Raphson method.

More generally,

$$x_{i+1} = x_i + \Delta x = x_i - \frac{f(x_i)}{f'(x_i)}$$

With an initial guess of x_0, we can then get a sequence x_1, x_2, …, which we expect to converge to the root of the equation.

Newton-Raphson method is equivalent to taking the slope of the function f(x) at the i^{th} iterative point, and the next approximation is the point where the slope intersects the x axis.



Question 5i

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ starting with the point 1.4, using the **Newton-Raphson Method**. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(x) = 2x^3 - 3x^2 - 2x + 3$$

$$f'(x) = 6x^2 - 6x - 2$$

$$x_0 = 1.4$$

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

$$= \frac{6x_0^{\ 3} - 6x_0^{\ 2} - 2x_0 - 2x_0^{\ 3} + 3x_0^{\ 2} + 2x_0 - 3}{6x_0^{\ 2} - 6x_0 - 2}$$

$$= \frac{4x_0^{\ 3} - 3x_0^{\ 2} - 3}{6x_0^{\ 2} - 6x_0 - 2}$$

$$= \frac{4(1.4)^3 - 3(1.4)^2 - 3}{6(1.4)^2 - 6(1.4) - 2}$$

$$= 1.5412$$

$x_1 = 1.5412$, $|x_1 - x_0| = 0.1412$

$x_2 = 1.5035$, $|x_2 - x_1| = 0.0377$

$x_3 = 1.5$, $|x_3 - x_2| = 0.0035$

$x_4 = 1.5$, $|x_4 - x_3| = 0$

Question 5j-2

Find a zero of the function $f(x) = x^3 - 2x^2 + x - 3$, $x_0 = 4$

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

Iteration 1: $\quad x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)} = 4 - \dfrac{33}{33} = 3$

Iteration 2: $\quad x_2 = x_1 - \dfrac{f(x_1)}{f'(x_1)} = 3 - \dfrac{9}{16} = 2.4375$

Iteration 3: $\quad x_3 = x_2 - \dfrac{f(x_2)}{f'(x_2)} = 2.4375 - \dfrac{2.0369}{9.0742} = 2.2130$

k (Iteration)	$\mathbf{x_k}$	$\mathbf{f(x_k)}$	$\mathbf{f'(x_k)}$	$\mathbf{x_{k+1}}$	$\mathbf{\lvert x_{k+1} - x_k \rvert}$
0	**4**	**33**	**33**	**3**	1
1	**3**	**9**	**16**	**2.4375**	0.5625
2	**2.4375**	**2.0369**	**9.0742**	2.2130	0.2245
3	2.2130	0.2564	6.8404	2.1756	0.0384
4	2.1756	0.0065	6.4969	2.1746	0.0010

Question 5j-2

Use Newton's Method to find a root of:

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

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

Five Iterations of the Solution

k	x_k	$f(x_k)$	$f'(x_k)$	ERROR
0	1.0000	-1.0000	2.0000	
1	1.5000	0.8750	5.7500	0.1522
2	1.3478	0.1007	4.4499	0.0226
3	1.3252	0.0021	4.2685	0.0005
4	1.3247	0.0000	4.2646	0.0000
5	1.3247	0.0000	4.2646	0.0000

Question 5j-3

Use Newton's Method to find a root of:

$f(x) = e^{-x} - x$

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

x_k	$f(x_k)$	$f'(x_k)$	$\frac{f(x_k)}{f'(x_k)}$
1.0000	-0.6321	-1.3679	0.4621
0.5379	0.0461	-1.5840	-0.0291
0.5670	0.0002	-1.5672	-0.0002
0.5671	0.0000	-1.5671	-0.0000

Question 5k

Estimates of the root of: x-cos(x)=0.

0.60000000000000		**Initial guess**
0.74401731944598	1	correct digit
0.73909047688624	4	correct digits
0.73908513322147	10	correct digits
0.73908513321516	14	correct digits

Question 5k-2

Given the equation: $\boldsymbol{f(x) = x^3 - 10 = 0}$ which root lies between 2 and 3. Find the real root using the Newton Raphson method (Up to 3 iterations and correct to 4 decimal places).

Taking $x_0 = 2$.

Let

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

and

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

Given $x_0 = 2$

$$\boldsymbol{x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}} \qquad (1)$$

Where i=0,1,2,3…

Let $f(x_i) = f(x)$ and $f'(x_i) = f'(x)$

First Iteration:

Hence, substituting $i = 0$ into equation (1) to get the first approximation of the root

We have

$$\boldsymbol{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}} \qquad \textbf{(2)}$$

$$x_0 = 2;\ f(x_0) = -2;\ f'(x_0) = 12$$

Note: $f(x_0)\ and\ f'(x_0)$ are derived by substituting $x_0 = 2$ into $f(x_i)\ and\ f'(x_i)$ where $i = 0$

Substituting the values into equation (2)

$$x_1 = 2 - \frac{(-2)}{12} = 2.1667$$

$$x_1 = 2.1667$$

Second Iteration:

Hence, substituting $i = 1$ into equation (1) to get the second approximation of the root

We have

$$\boldsymbol{x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}} \qquad \textbf{(3)}$$

$$x_1 = 2.1667;\ f(x_1) = 0.1718;\ f'(x_1) = 14.0838$$

Note: $f(x_1)\ and\ f'(x_1)$ are derived by substituting $x_1 = 2.1667$ into $f(x_i)\ and\ f'(x_i)$ where $i = 1$

Substituting the values into equation (3)

$$x_2 = 2.1667 - \frac{0.1718}{14.0838} = 2.1545$$

$$x_2 = 2.1545$$

Third Iteration:

Hence, substituting $i = 2$ into equation (1) to get the third approximation of the root

We have

$$\boldsymbol{x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}} \qquad \textbf{(4)}$$

$$x_2 = 2.1545;\ f(x_2) = 0.0009;\ f'(x_2) = 13.9256$$

Note: $f(x_1)\ and\ f'(x_1)$ are derived by substituting $x_2 = 2.1545$ into $f(x_i)\ and\ f'(x_i)$ where $i = 2$

Substituting the values into equation (4)

$$x_3 = 2.1545 - \frac{0.0009}{13.9256} = 2.1544$$

$$x_3 = 2.1544$$

Summary Table of Iterations

i (iteration)	X_i	$f(x_i)$	$f'(x_i)$	X_{i+1}	$\lvert X_{i+1} - X_i \rvert$
0	2	-2	12	2.1667	0.1667
1	2.1667	0.1718	14.0838	2.1545	0.0122
2	2.1545	0.0009	13.9256	2.1544	0.0001
3	2.1544	-0.0005	13.9243	2.1544	0

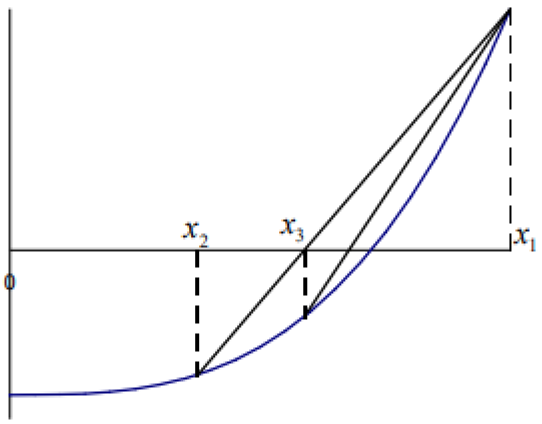
Thus, the root of the equation: $f(x) = x^3 - 10 = 0$ is <u>2.1544</u>

Regula-falsi method

- Also known as the false-position method, or linear interpolation method.

A regula-falsi or a method of false position assumes a test value for the solution of the equation.

- The *regula falsi* method start with two point, (a, f(a)) and (b,f(b)), satisfying the condition that f(a)f(b)<0.



Then, for an arbitrary x and the corresponding y,

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

gives the equation of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Setting $y = 0$, that is, where the chord crosses the x-axis,

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

Then, we evaluate $f(x_3)$. Just as in the case of root-bisection, if the sign is opposite that of $f(x_1)$, then a root lies in-between x_1 and x_3. Then, we replace x_2 by x_3 in equation

In just the same way, if the root lies between x_1 and x_3, we replace x_2 by x_1. We shall repeat this procedure until we are as close to the root as desired.

Question 5k-3

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the regula-falsi method.

$$f(1.4) = -0.192\,,\ f(1.7) = 0.756$$

A solution lies between x = 1.4 and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.4607595$

$$f(1.4607595) = -0.088983$$

The root lies between 1.46076 and 1.7. Let $x_1 = 1.46076$ and $x_2 = 1.7$.

$$x_4 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4607595 - (-0.088983)\frac{1.7 - 1.46076}{0.756 - (-0.088983)}$$

$= 1.485953$

Table for Regula-falsi method

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.495149	-0.011985
4	1.498346	-0.004118
5	1.499439	-0.001401
6	1.499810	-0.000475
7	1.499936	-0.000161
8	1.499978	-0.000055

Question 5L

- **Finding the Cube Root of 2 Using Regula Falsi**

- Since f(1)= -1, f(2)=6, we take as our starting bounds on the zero a=1 and b=2.
- Our first approximation to the zero is

$$x = b - \frac{b-a}{f(b) - f(a)}(f(b)) = 2 - \frac{2-1}{6+1}(6)$$
$$= 2 - 6/7 = 8/7 \approx 1.1429$$

- We then find the value of the function:

$$y = f(x) = (8/7)^3 - 2 \approx -0.5073$$

- Since f(a) and y are both negative, but y and f(b) have opposite signs

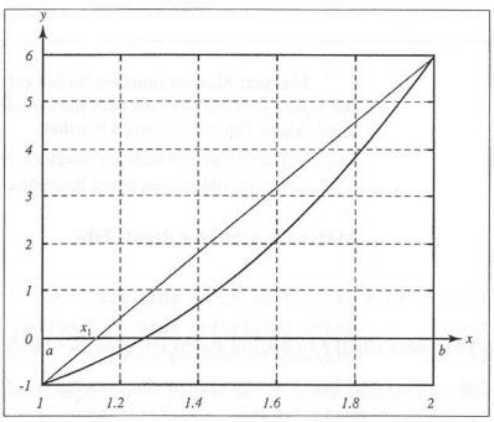


FIGURE 2.5 Graph of $y = x^3 - 2$ and approximation line on the interval [1, 2].

• Calculation of $\sqrt[3]{2}$ using *regula falsi*.

Step	a	b	x	y
1	1	2	1.1429	−0.50729
2	1.1429	2	1.2097	−0.22986
3	1.2097	2	1.2388	−0.098736
4	1.2388	2	1.2512	−0.041433
5	1.2512	2	1.2563	−0.017216
6	1.2563	2	1.2584	−0.0071239
7	1.2584	2	1.2593	−0.0029429
8	1.2593	2	1.2597	−0.0012148
9	1.2597	2	1.2598	−0.00050134
10	1.2598	2	1.2599	−0.00020687

Question 5L-2

Using the method of false position, find the real root of the equation $\boldsymbol{x^3 - 2x - 5 = 0}$. Where the real root lies between 2 and 2.1. (Up to 3 iterations and correct to 3 decimal places).

Let

$$f(x) = x^3 - 2x - 5$$

Given the roots as 2 and 2.1, therefore $a = 2 \text{ and } b = 2.1$

To find our approximation, we use the formula:

$$\boldsymbol{x_i = \frac{af(b) - bf(a)}{f(b) - f(a)}} \qquad (\mathbf{1})$$
$$\boldsymbol{where\ i = 1, 2, 3 \ldots}$$

First Iteration:

Where $i = 1; a = 2; b = 2.1; f(a) = -1;\ f(b) = 0.061$

Note: $f(a) \text{ and } f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 1$

Therefore, substituting in the values into equation (1) we have

$$x_1 = \frac{2(0.061) - 2.1(-1)}{0.061 - (-1)}$$

$$x_1 = \frac{0.122 + 2.1}{1.061} = \frac{2.222}{1.061} = 2.0942$$

$$x_1 = 2.0942$$

Thus,

$$f(x_1) = f(2.0942) = (2.0942)^3 - 2(2.0942) - 5$$

$$f(x_1) = -0.0039$$

Since $f(x_1)$ is a negative value, therefore, the new root lies between (2.0942, 2.1) and $a = 2.0942$; $b = 2.1$

Second Iteration:

Where $i = 2; a = 2.0942; b = 2.1; f(a) = -0.0039; f(b) = 0.061$

Note: $f(a)$ *and* $f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 2$

Therefore, substituting in the values into equation (1) we have

$$x_2 = \frac{2.0942(0.061) - 2.1(-0.0039)}{0.061 - (-0.0039)}$$

$$x_2 = \frac{0.12775 + 0.00819}{0.0649} = \frac{0.13594}{0.0649} = 2.0946$$

$$x_2 = 2.0946$$

Thus,

$$f(x_2) = f(2.0946) = (2.0946)^3 - 2(2.0946) - 5$$

$$f(x_2) = 0.0005$$

Since $f(x_2)$ is a positive value, therefore, the new root lies between (2.0942, 2.0946) and $a = 2.0942$; $b = 2.0946$

Third Iteration:

Where $i = 3; a = 2.0942; b = 2.0946; f(a) = -0.0039; f(b) = 0.0005$

Note: $f(a)$ *and* $f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 3$

Therefore, substituting in the values into equation (1) we have

$$x_3 = \frac{2.0942(0.0005) - 2.0946(-0.0039)}{0.0005 - (-0.0039)}$$

$$x_3 = \frac{0.00105 + 0.00817}{0.0044} = \frac{0.00922}{0.0044} = 2.0952$$

$$x_3 = 2.0952$$

Summary Table of Iterations

i (iteration)	**a**	**b**	**X_i**	**$f(x_i)$**
1	2	2.1	2.0942	-0.0039
2	2.0942	2.1	2.0946	0.0005
3	2.0942	2.0946	2.0952	0.007

Therefore, after three iterations, the required approximate root correct to 3 decimal places is <u>2.095</u>

Secant Method

In the case of the secant method, it is not necessary that the root lie between the two initial points. As such, the condition $f(x_1)f(x_2) < 0$ is not needed. Following the same analysis with the case of the regula-falsi method,

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

Setting $y = 0$ gives

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

Thus, having found x_n, we can obtain x_{n+1} as,

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \; n = 2, 3, \ldots$$

By inspection, if $f(x_n) - f(x_{n-1}) = 0$, the sequence does not converge, because the formula fails to work for x_{n+1}. The regula-falsi scheme does not have this problem as the associated sequence always converges.

Question 5m

Find the roots of the equation f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the secant method.

$x_1 = 1.4, \; x_2 = 1.7$

$f(1.4) = -0.192, \; f(1.7) = 0.756$

A solution lies between $x = 1.4$ and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$$= 1.460759$$

$f(x_3) = -0.088983$

$$x_4 = x_3 - f(x_3)\frac{x_3 - x_2}{f(x_3) - f(x_2)} = 1.460759 - (-0.088983) \times \frac{1.460759 - 1.7}{-0.088983 - 0.756}$$

$$= 1.485953$$

If the scheme continues, the table for secant method will be

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.501487	0.003730
4	1.499949	-0.000129
5	1.500000	0.000000

Question 5n

Find the roots of the equation by the secant method:

$$f(x) = x^2 - 2x + 0.5$$

$$x_0 = 0$$

$$x_1 = 1$$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Question 5n-i

Find the roots of the equation by the secant method:

Use Secant method to find the root of :

$f(x) = x^6 - x - 1$

Two initial points $x_0 = 1$ *and* $x_1 = 1.5$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

k	x_k	$f(x_k)$
0	1.0000	-1.0000
1	1.5000	8.8906
2	1.0506	-0.7062
3	1.0836	-0.4645
4	1.1472	0.1321
5	1.1331	-0.0165
6	1.1347	-0.0005

Question 5n-ii

Given the equation: $f(x) = x^3 - 5x + 1$, where x_0 and x_1 are 2 and 2.5 respectively. Find the real root using the Secant method. (Up to 4 iterations and correct to 4 decimal places).

Let

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

Given $x_0 = 2 \; and \; x_1 = 2.5$

To find our approximation, we use the formula:

$$\boldsymbol{x_{i+1} = x_i - f(x_i)\left(\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})}\right)} \qquad (\mathbf{1})$$

where i=1,2,3 …

Let $f(x_i) = f(x)$, where i=1,2,3 …

First Iteration:

Hence, substituting $i = 1$ into equation (1) to get the first approximation of the root

We have

$$\boldsymbol{x_2 = x_1 - f(x_1)\left(\frac{x_1 - x_0}{f(x_1) - f(x_0)}\right)} \qquad (\mathbf{2})$$

$x_0 = 2; x_1 = 2.5; \; f(x_0) = -1; \; f(x_1) = 4.125$

Note: $f(x_0) \; and \; f(x_1)$ are derived by substituting $x_0 = 2, \; x_1 = 2.5$ into $f(x_0) \; and \; f(x_1)$ respectively.

Substituting the values into equation (2)

$$x_2 = 2.5 - (4.125)\left(\frac{2.5 - 2}{4.125 - (-1)}\right)$$

$$x_2 = 2.5 - (4.125)\left(\frac{0.5}{5.125}\right)$$

$$x_2 = 2.5 - (4.125)(0.09756)$$

$$x_2 = 2.5 - 0.4024$$

$$x_2 = 2.0976$$

Thus,

$$f(x_2) = f(2.0976) = (2.0976)^3 - 5(2.0976) + 1$$

$$f(x_2) = -0.2587$$

Second Iteration:

Hence, substituting $i = 2$ into equation (1) to get the second approximation of the root

We have

$$\boldsymbol{x_3 = x_2 - f(x_2)\left(\frac{x_2 - x_1}{f(x_2) - f(x_1)}\right)} \qquad (\mathbf{3})$$

$x_1 = 2.5; x_2 = 2.0976; \; f(x_1) = 4.125; \; f(x_2) = -0.2587$

Note: $f(x_1) \; and \; f(x_2)$ are derived by substituting $x_1 = 2.5, \; x_2 = 2.0976$ into $f(x_1) \; and \; f(x_2)$ respectively.

Substituting the values into equation (3)

$$x_3 = 2.0976 - (-0.2587)\left(\frac{2.0976 - 2.5}{(-0.2587) - 4.125}\right)$$

$$x_3 = 2.0976 + (0.2587)\left(\frac{-0.4025}{-4.3837}\right)$$

$$x_3 = 2.0976 + (0.2587)(0.09182)$$

$$x_3 = 2.0976 + 0.0238$$

$$x_3 = 2.1214$$

Thus,

$$f(x_3) = f(2.1214) = (2.1214)^3 - 5(2.1214) + 1$$

$$f(x_3) = -0.0600$$

Third Iteration:

Hence, substituting $i = 3$ into equation (1) to get the third approximation of the root

We have

$$\boldsymbol{x_4 = x_3 - f(x_3)\left(\frac{x_3 - x_2}{f(x_3) - f(x_2)}\right)} \qquad (\mathbf{4})$$

$x_2 = 2.0976; x_3 = 2.1214;\ f(x_2) = -0.2587;\ f(x_3) = -0.0600$

Note: $f(x_2)$ *and* $f(x_3)$ are derived by substituting $x_2 = 2.0976,\ x_3 = 2.1214$ into $f(x_2)$ *and* $f(x_3)$ respectively.

Substituting the values into equation (4)

$$x_4 = 2.1214 - (-0.0600)\left(\frac{2.1214 - 2.0976}{(-0.0600) - (-0.2587)}\right)$$

$$x_4 = 2.1214 + (0.0600)\left(\frac{0.0238}{0.1987}\right)$$

$$x_4 = 2.1214 + (0.0600)(0.1198)$$

$$x_4 = 2.1214 + 0.0072$$

$$x_4 = 2.1286$$

Thus,

$$f(x_4) = f(2.1286) = (2.1286)^3 - 5(2.1286) + 1$$

$$f(x_4) = 0.0016$$

Fourth Iteration:

Hence, substituting $i = 4$ into equation (1) to get the fourth approximation of the root

We have

$$\boldsymbol{x_5 = x_4 - f(x_4)\left(\frac{x_4 - x_3}{f(x_4) - f(x_3)}\right)} \qquad (\mathbf{5})$$

$x_3 = 2.1214; x_4 = 2.1286;\ f(x_3) = -0.0600;\ f(x_4) = 0.0016$

Note: $f(x_3)$ *and* $f(x_4)$ are derived by substituting $x_3 = 2.1214,\ x_4 = 2.1286$ into $f(x_3)$ *and* $f(x_4)$ respectively.

Substituting the values into equation (5)

$$x_5 = 2.1286 - (0.0016)\left(\frac{2.1286 - 2.1214}{0.0016 - (-0.0600)}\right)$$

$$x_5 = 2.1286 - (0.0016)\left(\frac{0.0072}{0.0616}\right)$$

$$x_5 = 2.1286 - (0.0016)(0.1169)$$

$$x_5 = 2.1286 - 0.0002$$

$$x_5 = 2.1284$$

Thus,

$$f(x_5) = f(2.1284) = (2.1284)^3 - 5(2.1284) + 1$$

$$f(x_5) = -0.0002$$

Summary Table of Iterations

i (iteration)	**X_i**	**$f(x_i)$**	**X_{i+1}**	**$\lvert X_{i+1} - X_i \rvert$**
0	2	-1	2.5	0.5
1	2.5	4.125	2.0976	0.4022
2	2.0976	-0.2587	2.1214	0.0238
3	2.1214	-0.0600	2.1286	0.0072
4	2.1286	0.0016	2.1284	0.0002

Therefore, the root of the equation after 4 iterations correct to 4 decimal places is <u>2.1284</u>

Modified Secant Method

In this modified Secant method, only one initial guess is needed :

$$f'(x_i) \approx \frac{f(x_i + \delta x_i) - f(x_i)}{\delta x_i}$$

$$x_{i+1} = x_i - \frac{f(x_i)}{\dfrac{f(x_i + \delta x_i) - f(x_i)}{\delta x_i}} = x_i - \frac{\delta x_i f(x_i)}{f(x_i + \delta x_i) - f(x_i)}$$

Question 5(o)

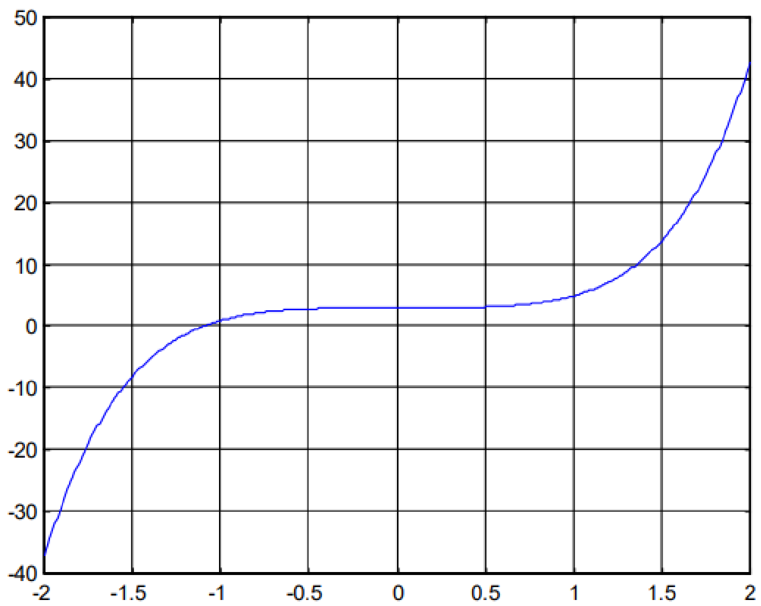
Find the roots of :

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

Initial points

$x_0 = -1 \; and \; x_1 = -1$

$with \;\; error < 0.001$



Fixed-point Iteration Method

- Also known as **one-point iteration** or **successive substitution**

- To find the root for $f(x) = 0$, we reformulate $f(x) = 0$ so that there is an x on one side of the equation.

$$f(x) = 0 \quad \Leftrightarrow \quad g(x) = x$$

- If we can solve $g(x) = x$, we solve $f(x) = 0$.
  - x is known as the fixed point of $g(x)$.
- We solve $g(x) = x$ by computing

$$x_{i+1} = g(x_i) \quad \text{with } x_0 \text{ given}$$

until x_{i+1} converges to x.

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

$$x^2 + 2x - 3 = 0 \quad \Rightarrow \quad 2x = 3 - x^2 \quad \Rightarrow \quad x = \frac{3 - x^2}{2}$$

$$\Rightarrow x_{i+1} = g(x_i) = \frac{3 - x_i^2}{2}$$

Reason: If x converges, i.e. $x_{i+1} \rightarrow x_i$

$$x_{i+1} = \frac{3 - x_i^2}{2} \rightarrow x_i = \frac{3 - x_i^2}{2}$$

$$\Rightarrow x_i^2 + 2x_i - 3 = 0$$

Question 5p

Use fixed point iteration to:

Find root of $f(x) = e^{-x} - x = 0$.

(Answer: α= 0.56714329)

We put $x_{i+1} = e^{-x_i}$

i	x_i	ε_a (%)	ε_t (%)
0	0		100.0
1	1.000000	100.0	76.3
2	0.367879	171.8	35.1
3	0.692201	46.9	22.1
4	0.500473	38.3	11.8
5	0.606244	17.4	6.89
6	0.545396	11.2	3.83
7	0.579612	5.90	2.20
8	0.560115	3.48	1.24
9	0.571143	1.93	0.705
10	0.564879	1.11	0.399

- There are infinite ways to construct $g(x)$ from $f(x)$.

For example, $f(x) = x^2 - 2x - 3 = 0$ (ans: $x = 3$ or -1)

Case a:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x^2 = 2x + 3$$
$$\Rightarrow x = \sqrt{2x+3}$$
$$\Rightarrow g(x) = \sqrt{2x+3}$$

Case b:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x(x-2) - 3 = 0$$
$$\Rightarrow x = \frac{3}{x-2}$$
$$\Rightarrow g(x) = \frac{3}{x-2}$$

Case c:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow 2x = x^2 - 3$$
$$\Rightarrow x = \frac{x^2-3}{2}$$
$$\Rightarrow g(x) = \frac{x^2-3}{2}$$

So which one is better?

Case a	Case b	Case c
$x_{i+1} = \sqrt{2x_i + 3}$	$x_{i+1} = \dfrac{3}{x_i - 2}$	$x_{i+1} = \dfrac{x_i^2 - 3}{2}$

Case a

1. $x_0 = 4$
2. $x_1 = 3.31662$
3. $x_2 = 3.10375$
4. $x_3 = 3.03439$
5. $x_4 = 3.01144$
6. $x_5 = 3.00381$

Converge!

Case b

1. $x_0 = 4$
2. $x_1 = 1.5$
3. $x_2 = -6$
4. $x_3 = -0.375$
5. $x_4 = -1.263158$
6. $x_5 = -0.919355$
7. $x_6 = -1.02762$
8. $x_7 = -0.990876$
9. $x_8 = -1.00305$

Converge, but slower

Case c

1. $x_0 = 4$
2. $x_1 = 6.5$
3. $x_2 = 19.625$
4. $x_3 = 191.070$

Diverge!

s/n	Root finding method	f(a)f(b)<0 assumption	2 initial point	1 initial point	Class	Derivatives needed	Formula
1	**Bisection**	Yes	Yes		Bracket		x3 = (x1 + x2) / 2
2	**Regua Falsi**	Yes	Yes		Bracket		$x_3 = x_1 - f(x_1)\dfrac{x_2 - x_1}{f(x_2) - f(x_1)}$
3	**Newton Raphson**			Yes	Open	Yes	$x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)}$
4	**Secant**		Yes		Open		$x_3 = x_2 - f(x_2)\dfrac{x_2 - x_1}{f(x_2) - f(x_1)}$
5	**Modified Secant**			Yes	Open	Yes	x1 = x – [f(x) – f '(x)]
6	**Fixedpoint Iteration**			Yes	Open		$x_{i+1} = g(x_i)$

COVENANT UNIVERSITY

CANAANLAND, KM 10, IDIROKO ROAD
P.M.B 1023, OTA, OGUN STATE, NIGERIA.

TITLE OF EXAMINATION: TEST 1
COLLEGE: College of Science and Technology
DEPARTMENT: Department of Computer and Information Sciences
SESSION: 2023/2024 **SEMESTER:**
ALPHA
COURSE CODE: CSC431 **CREDIT UNIT:** 3
COURSE TITLE: Computational Science and Numerical Methods
INSTRUCTION: Answer ALL questions **TIME:** 1 HOUR

Question 1 – (24mks: each iteration of each method carries 2 marks)

Given a floating ball with specific gravity 0.6N, radius 5.5cm, and equation for depth x to which the ball is submerged under water: $x^3 - 0.165x^2 + 3.993 \times 10^{-4} = 0$

The ball will be submerged between x = 0 and x = 2R where R is the radius of the ball.

Hint: 1 Newton = 1 kg.m/s², so 5.5cm = 0.055m, $0 \leq x \leq 2R \Rightarrow 0 \leq x \leq 2(0.055) \Rightarrow 0 \leq x \leq 0.11$.

Calculate the depth to which the ball is submerged when floating in water by finding the root of the equation **in 2 iterations** using each of the methods below.

s/n	Root finding method	f(a)f(b)<0 assumption	2 initial points	1 initial point	Type of Method	Derivatives Needed	Formula
A	**Bisection**	Yes	Yes		Bracket		x3 = (x1 + x2) / 2
B	**Regua Falsi**	Yes	Yes		Bracket		$x_3 = x_1 - f(x_1)\dfrac{x_2 - x_1}{f(x_2) - f(x_1)}$
C	**Newton Raphson**			Yes	Open	Yes	$x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)}$
D	**Secant**		Yes		Open		$x_3 = x_2 - f(x_2)\dfrac{x_2 - x_1}{f(x_2) - f(x_1)}$ for [x0, x1], [x1, x2], [x2, x3], …
E	**Modified Secant**			Yes	Open	Yes	x1 = x – [f(x) – f ’(x)]
F	**Fixedpoint Iteration**			Yes	Open		$x_{i+1} = g(x_i)$

Question 2 – (6mks)

A. Give at least two (2) pros and cons each of the following methods of solving non-linear equations: Bisection method, Newton-Raphson’s method and Secant method. [**3 marks]**

B. Discuss the relevance of the course “computational science and numerical methods” (give at least 3 application areas). **[3 marks]**

Sample solution

Question 1

For a function $f(x) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$ with interval ($x_1 = 0$, $x_2 = 0.11$)

Bisection method **[4mks]**

$x_2 = \frac{x_0+x_1}{2}$ for $f(x_0) \times f(x_1) < 0$

1st iteration

$f(x0) = f(0) = 0^3 - 0.165 \times 0^2 + 3.993 \times 10^{-4} = 0.0003993 > 0$

$f(x1) = f(0.11) = 0.11^3 - 0.165 \times 0.11^2 + 3.993 \times 10^{-4}$

$f(x1) = f(0.11) = 0.001331 - 0.0019965 + 0.0003993 = -0.0002662 < 0$

$x_2 = \frac{x0+x1}{2} = \frac{0+0.11}{2} = 0.055$

2nd iteration

$f(x0) = f(0) = 0.0003993 > 0$

$f(x1) = f(0.11) = -0.0002662 < 0$

$f(x2) = f(0.055) = 0.055^3 - 0.165 \times 0.055^2 + 3.993 \times 10^{-4}$

$f(x2) = f(0.055) = 0.000166375 - 0.000499125 + 0.0003993 = -0.00006655 < 0$

$x_3 = \frac{x_2+x_0}{2}$ for $f(x_2) \times f(x_0) < 0$

$x_3 = \frac{x2+x0}{2} = \frac{0.055+0}{2} = 0.0275$

Regula falsi method **[4mks]**

$x_2 = x_0 - [f(x_0)\ \frac{x_1-x_0}{f(x_1)-f(x_0)}]$ for $f(x_0) \times f(x_1) < 0$

1st iteration

$f(x0) = f(0) = 0.0003993 > 0$

$f(x1) = f(0.11) = -0.0002662 < 0$

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

$x_2 = 0 - [0.0003993\ \frac{0.11-0}{-0.0002662 - 0.0003993}]$

$x_2 = 0 - [0.0003993\ \frac{0.11}{-0.0006655}]$

$x_2 = 0 - [0.0003993 \times -165.2892562$

$x_2 = 0 - -0.066$

$x_2 = 0.066$

2nd iteration

$f(x0) = f(0) = 0.0003993 > 0$

$f(x1) = f(0.11) = -0.0002662 < 0$

$f(x2) = f(0.066) = 0.066^3 - 0.165 \times 0.066^2 + 3.993 \times 10^{-4}$

$f(x2) = f(0.066) = 0.000287496 - 0.00071874 + 0.0003993 = -0.000031944 < 0$

$x_3 = x_0 - [f(x_0) \frac{x_2 - x_0}{f(x_2) - f(x_0)}]$ for $f(x_0) \times f(x_2) < 0$

$x_3 = 0 - [0.0003993 \frac{0.066 - 0}{-0.000031944 - 0.0003993}]$

$x_3 = 0 - [0.0003993 \frac{0.066}{-0.000431244}]$

$x_3 = 0 - [0.0003993 \times -153.0456076$

$x_3 = 0 - -0.061111111$

$x_3 = 0.061111111$

Newton Raphson method **[4mks]**

$x_1 = x_0 - [\frac{f(x_0)}{f\prime(x_0)}]$, with radius 0.055m

1st iteration

$f(x_0) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$

$f'(x_0) = 3x^2 - 0.165(2x)$

$f(x_0) = f(0.055) = -0.00006655$

$f'(x_0) = f'(0.055) = 3 \times 0.055^2 - 0.165 \times 2 \times 0.055$

$f'(x_0) = f'(0.055) = 0.009075 - 0.01815 = -0.009075$

$x_1 = x_0 - [\frac{f(x_0)}{f\prime(x_0)}] = 0.055 - [\frac{-0.00006655}{-0.009075}] = [0.055 + 0.00733333]$

$x_1 = 0.062333333$

2nd iteration

$f(x_1) = f(0.062333333) = 0.062333333^3 - 0.165 \times 0.062333333^2 + 3.993 \times 10^{-4}$

$f(x_1) = f(0.06267085) = 0.000242193 - 0.0006410983 + 0.0003993 = -0.0000003947$

$f'(x_1) = f'(0.06267085) = 3 \times 0.062333333^2 - 0.165 \times 2 \times 0.062333333$

$f'(x_1) = f'(0.06267085) = 0.011656333 - 0.020569999$ = -0.008913666

$x_2 = x_1 - [\frac{f(x_1)}{f'(x_1)}] = 0.06267085 - [\frac{-0.0000003947}{-0.008913666}] = [0.06267085 + 0.00004428]$

$x_2 = 0.06271513$

Secant method **[4mks]**

$x_2 = x_1 - [f(x_1)\ \frac{x_1 - x_0}{f(x_1) - f(x_0)}]$ for [x0, x1], [x1, x2], [x2, x3], …

1st iteration

$f(x0) = f(0) = 0.0003993$

$f(x1) = f(0.11)$ = -0.0002662

$x_2 = x_1 - [f(x_1)\ \frac{x_1 - x_0}{f(x_1) - f(x_0)}]$ for [x0, x1]

$x_2 = 0.11 - [-0.0002662 \times \frac{0.11 - 0}{-0.0002662 - 0.0003993}]$

$x_2 = 0.11$ – [-0.0002662 $\times -165.2892562$]

$x_2 = 0.11 - 0.044$

$x_2 = 0.066$

2nd iteration

$f(x0) = f(0) = 0.0003993$

$f(x1) = f(0.11)$ = -0.0002662

$f(x2) = f(0.066)$ = -0.000031944

$x_3 = x_2 - [f(x_2)\ \frac{x_2 - x_1}{f(x_2) - f(x_1)}]$ for [x1, x2]

$x_3 = 0.066 - [0.00031944\ \frac{0.066 - 0.11}{-0.000031944 - 0.0002662}]$

$x_3 = 0.066 - [0.0003993\ \frac{0.044}{-0.000298144}]$

$x_3 = 0.066 - [0.0003993 \times -147.579693]$

$x_3 = 0.066 - -0.058928571$

$x_3 = 0.124$. This solution has diverged. Regula-falsi scheme does not have this problem as the associated sequence always converges.

Modified Secant method **[4mks]**

$x_1 = x_0 - [f(x_0) - f'(x_0)]$, with radius 0.055m

1st iteration

$f(x_0) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$

$f'(x_0) = 3x^2 - 0.165(2x)$

$f(x_0) = f(0.055)$ = -0.00006655

$f'(x_0) = f'(0.055)$ = -0.009075

$x_1 = x_0 - [\, f(x_0) - {}'f(x_0)]$ = $0.055 - [-0.00006655 - -0.009075]$ = $0.055 + 0.00900845$

$x_1 = 0.06400845$

2nd iteration

$f(x_1) = f(0.06400845) = 0.06400845^3 - 0.165 \times 0.06400845^2 + 3.993 \times 10^{-4}$

$f(x_1) = f(0.06400845) = 0.00026225 - 0.00067602 + 0.0003993$ = -0.00001447

$f'(x_1) = f'(0.06267085) = 3 \times 0.06400845^2 - 0.165 \times 2 \times 0.06400845$

$f'(x_1) = f'(0.06267085) = 0.012291245 - 0.021122788$ = -0.008831543

$x_2 = x_1 - [\, f(x_1) - {}'f(x_1)]$ = $0.06267085 - [-0.00001447 - -0.008831543]$

x_2= $0.06267085 - 0.008686843$

$x_2 = 0.053984007$

Fixed-point iteration method **[4mks]**

for $f(x\ \ );\ x\ \ = g(x\ \ )$

if $f(x) = x^3 - 0.165x^2 + 3.993 \times 10^{-4}$ with radius 0.055m

A possible x = g(x) = $[\, 0.165x^2 - 3.993 \times 10^{-4}]^{1/3}$ **at point 0.055**

1st Iteration

x1 = g(x0) = g(0.055) = $\left[0.165 \times 0.055^2 - 3.993 \times 10^{-4}\right]^{1/3}$

= $[0.000499 – 0.0003993]^{1/3}$

= $[0.000099825]^{1/3}$

= 0.046388796

2nd Iteration

x1 = g(x0) = g(0.055) = $[\, 0.165 \times 0.046388796^2 - 3.993 \times 10^{-4}]^{1/3}$

= $[0.0003551 – 0.0003993]^{1/3}$

= $[0.000044233]^{1/3}$

= 0.0353657

Sample solution - Question 2

A. Give at least two (2) pros and cons each of the following methods of solving non-linear equations: Bisection method, Newton-Raphson's method and Secant method. [**3 marks]**

Method	**Pros**	**Cons**
Bisection	• Easy, Reliable, Convergent • One function evaluation per iteration • No knowledge of derivative is needed	• Slow • Needs an interval $[a, b]$ containing the root, i.e., $f(a), f(b) < 0$
Newton-Raphson	• Fast (if near the root) • Two function evaluations per iteration	• May diverge • Needs derivative and an initial guess x0 such that $f'(x_0)$ is nonzero
Secant	• Fast (slower than Newton) • One function evaluation per iteration • No knowledge of derivative is needed	• May diverge • Needs two initial points guess x_0, x_1 such that $f(x_0) - f(x_1)$ is nonzero

[0.5 mk × any 6 correct points = 3 mks]

B. Discuss the relevance of the course "computational science and numerical method" (give at least 3 application areas). **[3 marks]**

Definition

Computational science is a field in mathematics that uses advanced computing capabilities to understand and solve complex problems. It is an area of science that spans many disciplines, but at its core, it involves the development of models and simulations to understand natural systems. It is often thought of as an integration of three disciplines- mathematics, computer science, and science. It also involves the invention, implementation, testing, and application of algorithms and software used to solve large-scale scientific and engineering problems

A **numerical method** is an approximate computer method for solving a mathematical problem which often has no analytical solution. It entails making use of computers to solve problems by step-wise, repeated and iterative solution methods, which would otherwise be tedious or unsolvable by hand-calculations.

Relevance

This course is multi-disciplinary and cuts across various fields such as computational biology, computational chemistry, computational physics, computational finance, economics and engineering. It helps to solve complex problems and optimize various processes across these disciplines.

- Modern scientists increasingly rely on computational modelling and data analysis to explore and understand the natural world. Given the ubiquitous use in science and its critical importance to the future of science and engineering, computational science plays a central role in progress and scientific developments in the 21st Century.

- It aims at educating the next generation of cross-disciplinary science students with the knowledge, skills, and values needed to pose and solve current and new scientific, technological and societal challenges.
- Computational science focuses on the development of predictive computer models of the world around us. As studies of physical phenomena evolved to address increasingly complex systems, traditional experimentation is often infeasible. Thus, the discipline entails the development of new methods that make challenging problems tractable on modern computing platforms, providing scientists and engineers with new windows into the world around us.

Application Areas

1. **Engineering:** It cuts across various engineering disciplines such as Civil engineering, Mechanical engineering, Electrical engineering, etc. For example, numerical methods in Civil Engineering are now used routinely in structural analysis to determine the member forces and moments in structural systems, prior to design.
2. **Optimization:** useful for solving optimization problems which requires trial through numerous iterations e.g., scheduling tasks on processors in a heterogeneous multiprocessor computing network.
3. **Algorithm Trading:** it helps to automate process. Basically, it factors in time, cost and volume to aid decision making.
4. **Computation Biology:** studies biological systems. For example, airflow patterns in the respiratory tract, transport and disposition of chemicals through the body, etc.
5. **Computational Physics:** study and analysis of physical problems through computation and modelling. It covers areas such as fluid dynamics, astrophysics, thermodynamics, electromagnetics, etc.
6. **Computational Chemistry**: entails the study and prediction of chemical reactions. Also entails the understanding of molecular structures and properties. For example, transport and disposition of chemicals, predicting the evolution of crystals growing in an industrial crystallizer, etc.
7. **Computational Finance:** which has gained popularity in recent times. It helps in making the best investments and covers various aspects of finance. For example, calculation of insurance risks and price of insurance.
8. **Computation Economics:** this is similar to Computational Finance but different, as this branch focuses on predictive economic models.

[0.5 mk × any 6 correct points = 3 mks]

Question 3- Removed

Write a Python program that accepts these inputs; order of an equation, coefficients of each x in the equation, and the gradient of the equation **[2.5mk].** The program then imports 'symbols' function from 'sympy' library and generate the symbolic representation of the equation like this: f_of_(x) =+ "…" **[2.5mks].** Finally, the program accepts a value for x, computes f_of_(x), and outputs the value of f_of_(x) for any given function [**1mk].**

```
poly_power = int(input("To what polynomial power is the function? "))
```

```
coeffs = []
for i in range(poly_power):
    co_ef = float(input(f"Enter coefficient value for power {poly_power - i}: "))
    coeffs.append(co_ef)
poly_gradient = float(input("What is the gradient of the function? "))

from sympy import symbols
x = symbols('x')
f_of_x = 0
for i in range(poly_power):
    power = poly_power - i
    f_of_x += coeffs[i] * x**power
f_of_x += poly_gradient
print("f_of_x = ", f_of_x)

x_val = float(input("Please enter a value for x: "))
y = f_of_x.subs(x, x_val)
print(f"f_of_x = {y:.7f}")
```

7 Methods of Interpolation

LINEAR INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2) and x.

$$y = y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1}$$

QUADRATIC INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), and x.

$$y = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

CUBIC INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), and x.

$$y = y_1 \frac{(x - x_2)(x - x_3)(x - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)} + y_2 \frac{(x - x_1)(x - x_3)(x - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)} + y_3 \frac{(x - x_1)(x - x_2)(x - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)} + y_4 \frac{(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)}$$

Nth INTERPOLATION: GENERAL SYSTEM OF INTERPOLATION

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
X = 16	Y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$
\begin{aligned}
y \;=\;& y_1 \frac{(x-x_2)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)\dots(x_1-x_{n-1})(x_1-x_n)} \\
&+ y_2 \frac{(x-x_1)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)\dots(x_2-x_{n-1})(x_2-x_n)} \\
&+ y_3 \frac{(x-x_1)(x-x_2)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)\dots(x_3-x_{n-1})(x_3-x_n)} \\
&+ y_{n-1} \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_n)}{(x_{n-1}-x_1)(x_{n-1}-x_2)(x_{n-1}-x_3)\dots(x_{n-1}-x_{n-2})(x_{n-1}-x_n)} \\
&+ y_n \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)(x_n-x_3)\dots(x_n-x_{n-2})(x_n-x_{n-1})}
\end{aligned}
$$

LAGRANGE METHOD

Table of Velocity as a function of Time

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
x = 16	y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$
\begin{aligned}
y \;=\;& y_1 \frac{(x-x_2)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_1-x_2)(x_1-x_3)(x_1-x_4)\dots(x_1-x_{n-1})(x_1-x_n)} \\
&+ y_2 \frac{(x-x_1)(x-x_3)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_2-x_1)(x_2-x_3)(x_2-x_4)\dots(x_2-x_{n-1})(x_2-x_n)} \\
&+ y_3 \frac{(x-x_1)(x-x_2)(x-x_4)\dots(x-x_{n-1})(x-x_n)}{(x_3-x_1)(x_3-x_2)(x_3-x_4)\dots(x_3-x_{n-1})(x_3-x_n)} + \dots \\
&+ y_{n-1} \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_n)}{(x_{n-1}-x_1)(x_{n-1}-x_2)(x_{n-1}-x_3)\dots(x_{n-1}-x_{n-2})(x_{n-1}-x_n)} \\
&+ y_n \frac{(x-x_1)(x-x_2)(x-x_3)\dots(x-x_{n-2})(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)(x_n-x_3)\dots(x_n-x_{n-2})(x_n-x_{n-1})}
\end{aligned}
$$

$$
\begin{aligned}
y \;=\;& f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x) + f(x_4)L_4(x) \\
&+ \dots + f(x_{n-1})L_{n-1}(x) + f(x_n)L_n(x)
\end{aligned}
$$

$$y = \sum_{i>0}^{n} f(x_i)L_i(x)$$

NEWTON'S DIVIDED DIFFERENCE INTERPOLATION

(s)	(m/s)
0	0
10	227.04
15	362.78
20	517.35
22.5	602.97
30	901.67
…	…
…	…
…	…
x = 16	y = ?

Given (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4), …, (x_{n-1}, y_{n-1}), and x_n.

$$y = f(x_1) + f(x_1, x_2)(x - x_1) + f(x_1, x_2, x_3)(x - x_1)(x - x_2) + f(x_1, x_2, x_3, x_4)(x - x_1)(x - x_2)(x - x_3) + \dots$$

Such that:

The coefficients $f[x_i, x_{i+1}, \dots, x_j]$ are the divided differences obtained from the divided difference table

$f(x_1) = y_1$

$f(x_1, x_2) = \frac{y_2 - y_1}{x_2 - x_1}$ where $y_2 - y_1$ is the linear contribution based on the first two data points

$f(x_1, x_2, x_3) = \frac{(y_3 - y_2)(y_2 - y_1)}{x_3 - x_1}$ where $(y_3 - y_2)(y_2 - y_1)$ is the quadratic contribution based on the first three data points

$f(x_1, x_2, x_3, x_4) = \frac{(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)}{x_4 - x_1}$ where $(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)$ is the cubic polynomial contribution based on the first four data points

Newton's divided difference table

xi	fi/yi	f(xi, xj) for FDD	f(xi, xj, xk) for SDD	f(xi, xj, xk, xL) for TDD
x1	f1			
		$f(x_2, x_1) = \frac{f_2 - f_1}{x_2 - x_1}$		
x2	f2		$f(x_3, x_2, x_1) = \frac{(y_3 - y_2)(y_2 - y_1)}{x_3 - x_1}$	
		$f(x_3, x_2) = \frac{f_3 - f_2}{x_3 - x_2}$		$f(x_4, x_3, x_2, x_1) = \frac{(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)}{x_4 - x_1}$
x3	f3		$f(x_4, x_3, x_2) = \frac{(y_4 - y_3)(y_3 - y_2)}{x_4 - x_2}$	
		$f(x_4, x_3) = \frac{f_4 - f_3}{x_4 - x_3}$		$f(x_5, x_4, x_3, x_2) = \frac{(y_5 - y_4)(y_4 - y_3)(y_3 - y_2)}{x_5 - x_2}$
x4	f4		$f(x_5, x_4, x_3) = \frac{(y_5 - y_4)(y_4 - y_3)}{x_5 - x_3}$	
		$f(x_5, x_4) = \frac{f_5 - f_4}{x_5 - x_4}$		

Question 7 - Assignment

Given a table of Velocity (m/s) as function of Time (s) below:

(s)	10	15	16	20	22.5	30
(m/s)	227.04	362.78	??	517.35	602.97	901.67

Find the velocity at t=16 seconds using

A. The Lagrangian method of

i. Linear interpolation;

$$y = f(x_1)L_1(x) + f(x_2)L_2(x) = y_1\frac{x - x_2}{x_1 - x_2} + y_2\frac{x - x_1}{x_2 - x_1}$$

ii. Quadratic interpolation;

$$y = f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x) = y_1\frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2\frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3\frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

iii. Cubic interpolation;

$$y = f(x_1)L_1(x) + f(x_2)L_2(x) + f(x_3)L_3(x) + f(x_4)L_4(x)$$

$$= y_1\frac{(x - x_2)(x - x_3)(x - x_4)}{(x_1 - x_2)(x_1 - x_3)(x_1 - x_4)} + y_2\frac{(x - x_1)(x - x_3)(x - x_4)}{(x_2 - x_1)(x_2 - x_3)(x_2 - x_4)} + y_3\frac{(x - x_1)(x - x_2)(x - x_4)}{(x_3 - x_1)(x_3 - x_2)(x_3 - x_4)} + y_4\frac{(x - x_1)(x - x_2)(x - x_3)}{(x_4 - x_1)(x_4 - x_2)(x_4 - x_3)}$$

B. The values from Newton's divided difference table for the coefficients in

i. First divided difference;

$$y = f(x_1) + f(x_1, x_2)(x - x_1)$$

ii. Second divided difference;

$$y = f(x_1) + f(x_1, x_2)(x - x_1) + f(x_1, x_2, x_3)(x - x_1)(x - x_2)$$

iii. Third divided difference;

$$y = f(x_1) + f(x_1, x_2)(x - x_1) + f(x_1, x_2, x_3)(x - x_1)(x - x_2) + f(x_1, x_2, x_3, x_4)(x - x_1)(x - x_2)(x - x_3)$$

xi	fi/yi	f(xi, xj) for FDD	f(xi, xj, xk) for SDD	f(xi, xj, xk, xL) for TDD
x1	f1			
		$f(x_2, x_1) = \frac{f_2 - f_1}{x_2 - x_1}$		
x2	f2		$f(x_3, x_2, x_1) = \frac{(y_3 - y_2)(y_2 - y_1)}{x_3 - x_1}$	
		$f(x_3, x_2) = \frac{f_3 - f_2}{x_3 - x_2}$		$f(x_4, x_3, x_2, x_1) = \frac{(y_4 - y_3)(y_3 - y_2)(y_2 - y_1)}{x_4 - x_1}$
x3	f3		$f(x_4, x_3, x_2) = \frac{(y_4 - y_3)(y_3 - y_2)}{x_4 - x_2}$	
		$f(x_4, x_3) = \frac{f_4 - f_3}{x_4 - x_3}$		$f(x_5, x_4, x_3, x_2) = \frac{(y_5 - y_4)(y_4 - y_3)(y_3 - y_2)}{x_5 - x_2}$
x4	f4		$f(x_5, x_4, x_3) = \frac{(y_5 - y_4)(y_4 - y_3)}{x_5 - x_3}$	
		$f(x_5, x_4) = \frac{f_5 - f_4}{x_5 - x_4}$		

Sample solution

0 INTRO TO COMPUTATIONAL SCIENCE AND NUMERICAL METHODS

Numerical analysis ---> Numerical methods -----> Computational numerical methods ----> Computational numerical analysis

Numerical Analysis

Numerical Analysis is the study of numerical methods. Numerical analysis finds application in all fields of engineering and the physical sciences, and in the 21st century also the life and social sciences, medicine, business and even the arts. The GOAL of numerical analysis is the design and analysis of techniques/METHODS to give approximate but accurate solutions to hard problems.

Numerical methods

Numerical methods are mathematical attempts at finding approximate solutions of problems rather than the exact ones.

Computational numerical methods

Before modern computers, numerical methods often relied on hand formulas, using data from large printed tables. Since the mid20th century, computers calculate the required functions instead, but many of the same formulas continue to be used in software algorithms.

Computational numerical analysis

Current growth in computing power has enabled the use of more complex numerical analysis, providing detailed and realistic mathematical models in science and engineering. Numerical analysis continues this long tradition: rather than giving exact symbolic answers translated into digits and applicable only to real-world measurements, approximate solutions within specified error bounds are used.

1 Methods of Approximations

- Rounding off to significant figures
- Rounding off to decimal places
    - Working with arithmetic precision

2 Methods of Errors

    - Sources of errors
- Rounding errors
- Inherent errors
- Truncation errors
- True errors
- Relative true errors
- Absolute errors
- Relative absolute errors
- Approximate errors
- Relative approximate errors
- Absolute relative errors
- Percentage errors
    - Propagation of errors

3 Methods of Drawing the Lines of best fit

- Linearization
- Least squares curve fitting
- Group averages grouped averages curve fitting

4 Methods of solving Linear Systems of Equations

- Gaussian elimination
- Gauss-Jordan elimination
- LU decomposition
- Jacobi iteration
- Gauss-Seidal iteration

5 Methods of finding the roots of Algebraic and Transcendental Equations

- Bisection
- Newton-Raphson
- Regula-falsi
- Secant
- Modified Secant
- Fixed point Iteration

6 Methods of Finite Differences

- First forward difference
- First backward difference
- First central difference

7 Methods of Interpolation

- Newton's forward interpolation formula
- Newton's backward interpolation formula

8 Methods of Numerical Integration

- Newton-coates Quadrature
- Trapezoidal rule
- Simpson's one-third rule
- Simpson's three-eight rule
- Romberg's method

9 Methods of Solving First Order Ordinary Differential Equations

- Picard's Method
- Euler Method
- Modified Euler Method
- Runge-Kutta first order method
- Runge-Kutta second order method
- Runge-Kutta third order method
- Runge-Kutta fourth order method

10 Methods of describing grouped and ungrouped statistical data

- Measures of central tendency (mean, median, mode)
- Measures of position (quartiles, percentiles, deciles)
- Measures of dispersion (range, interquartile range, standard deviation, variance)

1 Methods of Approximations

Question 1

Without calling any in-built library or function, write a new function from scratch to

a. round-off any number to a stated precision of decimal place
b. return the absolute value of any number
c. find the natural log of any number
d. approximate any number to its nearest_integer
e. take the magnitude of any number
f. round off any number to a stated amount of significant figures
g. re-write ‘f’ using built-in libraries and functions.

Solution to Question 1f

ALGORITHM - To round off numbers to certain amount of significant figures

```
Given any ‘number’, with the ‘num_sig_figs’ to approximate the number to
If ‘number’ == 0,
        Return 0.0
Else
        Take the absolute value of ‘number’
        Take the natural log of the absolute value
        Take ‘number’ to its nearest_integer
        Take the magnitude of the natural log
        Calculate rounding_factor = 10** (num_sig_figs – magnitude - 1)
        Find rounded_number = nearest_integer / rounding_factor
        If ‘number’ > 0
                Return rounded_number
        If ‘number’ < 0
                Return –rounded_number
```

PSEUDOCODE - To round off numbers to certain amount of significant figures

Without calling any in-built library or function, this is a new function from scratch to round off numbers to certain amount of significant figures

```
def round_to_significant_figures(number, num_sig_figs):
    if number == 0:
        return 0.0


    # Calculate the absolute value of 'number'
    def absolute_value(number):
        if number < 0:
            return -number
        else:
            return number
    abs_number = absolute_value(number)


    # Calculate the natural logarithm of the absolute value
    def custom_ln(number, num_terms=100):
        if number == 1:
            return 0.0
        elif number < 1:
            number = 1 / number
            num_terms = -num_terms
        nat_log = 0.0
        for n in range(1, num_terms + 1):
            term = ((number - 1) ** n) / n
            if n % 2 == 0:
                nat_log -= term
            else:
                nat_log += term
        return nat_log


    # Calculate the magnitude of the natural log
```

```
    def custom_floor(nat_log):
        if nat_log >= 0:
            return int(nat_log)
        else:
            integer_part = int(nat_log)
            if integer_part == nat_log:
                return integer_part
            else:
                return integer_part - 1
    magnitude = custom_floor(custom_ln(abs_number))

    # Calculate rounding_factor
    rounding_factor = 10 ** (num_sig_figs - magnitude - 1)

    # Use rounding factor to round number to the specified significant figures
    def custom_roundoff(number):
        decimal_part = number - int(number)
        if decimal_part < 0.5:
            return int(number)
        else:
            return int(number) + 1
    rounded_number = custom_roundoff(abs_number * rounding_factor) / rounding_factor

    # Restore the sign
    if number > 0:
        return rounded_number
    else:
        return -rounded_number

#round_to_significant_figures(number, num_sig_figs)
```

Arithmetic precision

It might be necessary to round off numbers to make them useful for numerical computation, moreso as it would require an infinite computer memory to store an unending number.

The precision of a number is an indication of the number of digits that have been used to express it. In scientific computing, it is the number of significant digits or numbers, while in management and financial systems, it is the number of decimal places. We are quite aware that most currencies in the world are quoted to two decimal places.

Arithmetic precision (often referred to simply as precision) is the specified number of significant figures or digits to which the number of interest is to be rounded.

2 Methods of errors

Rounding Errors

These are errors incurred by truncating a sequence of digits representing a number, as we saw in the case of representing the rational number 3/7 by 2.3333, instead of 2.3333….., which is an unending number. Apart from being unable to write this number in an exact form by hand, our instruments of calculation, be it the calculator or the computer, can only handle a finite string of digits. Rounding errors can be reduced if we change the calculation procedure in such a way as to avoid the subtraction of nearly equal numbers or division by a small number. It can also be reduced by retaining at least one more significant figure at each step than the one given in the data, and then rounding off at the last step.

Inherent Errors

As the name implies, these are errors that are inherent in the statement of the problem itself. This could be due to the limitations of the means of calculation, for instance, the calculator or the computer. This error could be reduced by using a higher precision of calculation.

Truncation Errors

If we truncate Taylor's series, which should be an infinite series, then some error is incurred. This is the error associated with truncating a sequence or by terminating an iterative process. This kind of error also results when, for instance, we carry out numerical differentiation or integration, because we are replacing an infinitesimal process with a finite one. In either case, we would have required that the elemental value of the independent variable tend to zero in order to get the exact value.

Absolute Error, Relative True Error, Relative Approximate Error and Percentage Error etc.

Question 2

a. A student measured the length of a string of actual length 72.5 cm as 72.4 cm.
   i. Calculate the absolute error and the percentage error
   ii. Write a function that accepts measured length and actual length to output absolute error and the percentage error.

b. The derivative of a function f(x) can be approximated by the equation

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}, \qquad \text{if } f(x) = 7e^{0.5x}, \text{ and } h = 0.3,$$

   i. Find the true value, the approximate value, true error, and relative error of f'(2)
   ii. If true values are not known or are very difficult to obtain,
   then Approximate error (E_a) = Present Approximation – Previous approximation.
   For $f(x) = 7e^{0.5x}$ at x = 2, find
      - f'(2) using h = 0.3
      - f'(2) using h = 0.15

- approximate error, and relative approximate error of f'(2)

iii. Write a function that takes in any value of x for the derivative of a function f(x) approximated by the equation f'(x) = [f(x + h) - f(x)] / h for f(x) = 7e^0.5x, h1 = 0.3, h2 = 0.15, and returns true value, approximate value, true error, relative error, approximate error, and relative approximate error of f'(x)

Solution to Question 2a

(2ai)

Absolute error = | actual value – measured value |
Relative absolute error = | actual value – measured value | / actual value
The percentage error = Relative absolute error x 100
Absolute error = | 72.5 – 72.4 |= 0.01.
The percentage error = (0.1 / 72.5) x 100 = 0.1379

(2aii)

```
def calculate_errors(actual_length, measured_length):
    absolute_error = abs(actual_length - measured_length)
    relative_error = absolute_error / actual_length
    percentage_error = relative_error * 100

    return {
        "Absolute Error": absolute_error,
        "Relative Absolute Error": relative_error,
        "Percentage Error": percentage_error
    }

calculate_errors(actual_length, measured_length)
```

Solution to Question 2b

(2bi)
Approximate value of f'(x), for x = 2, and h = 0.3

$$f'(2) \approx \frac{f(2+0.3) - f(2)}{0.3}$$

$$= \frac{f(2.3) - f(2)}{0.3}$$

$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$

$$= \frac{22.107 - 19.028}{0.3} \quad = 10.263$$

(2bi cont'd)
The exact or true value of f'(2) can be found by using our knowledge of differential calculus

$$f(x) = 7e^{0.5x}$$

$$f'(x) = 7 \times 0.5 \times e^{0.5x}$$

$$= 3.5e^{0.5x}$$

$$f'(2) = 3.5e^{0.5(2)}$$

$$= 9.5140$$

(2bi cont'd)

True Error = True Value – Approximate Value
E_t = 9.5140 – 10263 = -0.722

Relative true error = (True value – Approximate value) / True value

= (9.5140 – 10263) / 9.5140= -0.722 / 9.5140

(2bii)

For x = 2, and h = 0.3

- Approximate value of f'(x) = 10.263

For x = 2, and h = 0.15

- Approximate value of f'(x) =

$$f'(2) \approx \frac{f(2+0.15)-f(2)}{0.15}$$

$$= \frac{f(2.15)-f(2)}{0.15}$$

$$= \frac{7e^{0.5(2.15)} - 7e^{0.5(2)}}{0.15}$$

$$= \frac{20.50-19.028}{0.15} = 9.8800$$

Approximate error (E_a) = Present Approximation – Previous approximation

= 9.8800 – 10.263

= -0.38300

Relative approximate error = Approximate error / Previous approximation

$$= \frac{-0.38300}{9.8800} = -0.038765$$

(Question 2biii)

```
import math
def f(x):
    return 7 * math.exp(0.5 * x)
def derivative_of_f(x):
    return 3.5 * math.exp(0.5 * x)

def calculate_derivative_error(x, first_h, f, second_h=None):
    true_value = derivative_of_f(x)
    first_approximate_value = (f(x + first_h) - f(x)) / first_h
    true_error = abs(true_value - first_approximate_value)
    relative_true_error = true_error / true_value
    if second_h is not None:
        second_approximate_value = (f(x + second_h) - f(x)) / second_h
        approximate_error = abs(second_approximate_value - first_approximate_value)
        relative_approximate_error = approximate_error / second_approximate_value
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "Second Approximate Value": second_approximate_value,
            "Approximate Error": approximate_error,
            "Relative Approximate Error": relative_approximate_error
        }
    else:
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "True Error": true_error,
            "Relative True Error": relative_true_error
        }
#calculate_derivative_error(x, first_h, f)
#calculate_derivative_error(x, first_h, f, second_h)
```

Propagation of errors

In numerical methods, the calculations are not made with exact numbers. How do these inaccuracies propagate through the calculations?

Question 2c

Find the bounds for the propagation in adding two numbers. For example if one is calculating $X + Y$ where

$$X = 1.5 \pm 0.05$$

$$Y = 3.4 \pm 0.04$$

Solution

Maximum possible value of $X = 1.55$

Maximum possible value of $Y = 3.44$

Maximum possible value of $X + Y = 1.55 + 3.44 = 4.99$

Minimum possible value of $X = 1.45$.

Minimum possible value of $Y = 3.36$.

Minimum possible value of $X + Y = 1.45 + 3.36 = 4.81$

Hence

$$4.81 \leq X + Y \leq 4.99.$$

Propagation of Errors In Formula

If f is a function of several variables $X_1, X_2, X_3, \ldots\ldots., X_{n-1}, X_n$

then the maximum possible value of the error in f is

$$\Delta f \approx \left|\frac{\partial f}{\partial X_1}\Delta X_1\right| + \left|\frac{\partial f}{\partial X_2}\Delta X_2\right| + \ldots\ldots. + \left|\frac{\partial f}{\partial X_{n-1}}\Delta X_{n-1}\right| + \left|\frac{\partial f}{\partial X_n}\Delta X_n\right|$$

Question 2d

The strain in an axial member of a square cross-section is given by

$$\in = \frac{F}{h^2 E}$$

Given F = 72

h = 4 x 10^{-3}

E = 70 x 10^{9}

Find the maximum possible error in the measured strain.

Solution

$$\in = \frac{72}{(4\times10^{-3})^2(70\times10^9)}$$
$$= 64.286\times10^{-6}$$
$$= 64.286\mu$$

$$\Delta \in = \left|\frac{\partial \in}{\partial F}\Delta F\right| + \left|\frac{\partial \in}{\partial h}\Delta h\right| + \left|\frac{\partial \in}{\partial E}\Delta E\right|$$

$$\frac{\partial \in}{\partial F} = \frac{1}{h^2 E} \qquad \frac{\partial \in}{\partial h} = -\frac{2F}{h^3 E} \qquad \frac{\partial \in}{\partial E} = -\frac{F}{h^2 E^2}$$

Thus

$$\Delta E = \left|\frac{1}{h^2 E}\Delta F\right| + \left|\frac{2F}{h^3 E}\Delta h\right| + \left|\frac{F}{h^2 E^2}\Delta E\right|$$
$$= \left|\frac{1}{(4\times10^{-3})^2(70\times10^9)}\times0.9\right| + \left|\frac{2\times72}{(4\times10^{-3})^3(70\times10^9)}\times0.0001\right|$$
$$+ \left|\frac{72}{(4\times10^{-3})^2(70\times10^9)^2}\times1.5\times10^9\right|$$

Hence

$\in = (64.286\mu \pm 5.3955\mu)$

Question 2e

Subtraction of numbers that are nearly equal can create unwanted inaccuracies. Using the formula for error propagation, show that this is true.

Solution

Let z = x – y, Then

$$\begin{aligned} |\Delta z| &= \left|\frac{\partial z}{\partial x}\Delta x\right| + \left|\frac{\partial z}{\partial y}\Delta y\right| \\ &= |(1)\Delta x| + |(-1)\Delta y| \\ &= |\Delta x| + |\Delta y| \end{aligned}$$

So the relative error or relative change is

$$\left|\frac{\Delta z}{z}\right| = \frac{|\Delta x| + |\Delta y|}{|x-y|}$$

Check: $x = 2 \pm 0.001$

$$y = 2.003 \pm 0.001$$

$$\left|\frac{\Delta z}{z}\right| = \frac{|0.001| + |0.001|}{|2 - 2.003|}$$

$$= 0.667$$ Percentage error = 66.67%

Taylor series

Some examples of common Taylor series

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

The general form of Taylor series is given as

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots$$

provided that all derivatives of f(x) are continuous and exist in the interval [x,x+h]

What does this mean in plain English?

As Archimedes would have said, "*Give me the value of the function at a single point, and the (first, second, and so on) values of all its derivatives at that single point, and I can give you the value of the function at any other point*"

Question 2f

Find the value of f(6) given that f(4) =125, f'(4) = 74, f''(4) = 30, f'''(4) = 6 and all other higher order derivatives of f(x) at x=4 are zero.

Solution

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + \cdots$$

x = 4

h = 6 – 4 = 2

Since the higher order derivatives are zero,

$$f(4+2) = f(4) + f'(4)2 + f''(4)\frac{2^2}{2!} + f'''(4)\frac{2^3}{3!}$$

$$f(6) = 125 + 74(2) + 30\left(\frac{2^2}{2!}\right) + 6\left(\frac{2^3}{3!}\right)$$

= 125 + 148 + 60 + 8

= 341

Note that to find f(6) exactly, we only need the value of the function and all its derivatives at some other point, in this case x = 4.

Error in Taylor series

The Taylor polynomial of order n of a function f(x) with (n+1) continuous derivatives in the domain [x,x+h] is given by

$$f(x+h)= f(x)+ f'(x)h+ f''(x)\frac{h^2}{2!}+\cdots+ f^{(n)}(x)\frac{h^n}{n!}+R_n(x)$$

where the remainder is given by

$$R_n(x)=\frac{(x-h)^{n+1}}{(n+1)!}f^{(n+1)}(c)$$

Such that

$$x<c<x+h$$

that is, c is some point in the domain [x,x+h]

Derivation for Maclaurin Series for e^x

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots$$

The Maclaurin series is simply the Taylor series about the point x=0

$$f(x+h)= f(x)+ f'(x)h+ f''(x)\frac{h^2}{2!}+ f'''(x)\frac{h^3}{3!}+ f''''(x)\frac{h^4}{4}+ f'''''(x)\frac{h^5}{5}+\cdots$$

$$f(0+h)= f(0)+ f'(0)h+ f''(0)\frac{h^2}{2!}+ f'''(0)\frac{h^3}{3!}+ f''''(0)\frac{h^4}{4}+ f'''''(0)\frac{h^5}{5}+\cdots$$

Since f(x) = e^x, f'(x) = e^x, f''(x) = e^x, $f^n(x) = e^x$, and $f^n(0) = e^0 = 1$;

The Maclaurin series is then:

$$f(h)=(e^0)+(e^0)h+\frac{(e^0)}{2!}h^2+\frac{(e^0)}{3!}h^3\ldots$$

$$=1+h+\frac{1}{2!}h^2+\frac{1}{3!}h^3\ldots$$

Therefore

$$f(x)=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots$$

It can be seen that as the number of terms used increases, the error bound decreases and hence a better estimate of the function can be found.

Question 2g

How many terms would it require to get an approximation of e^1 within a magnitude of true error of less than $10^{-6.}$

Solution

Using (n + 1) terms of Taylor seriesgives error bound of

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c) \qquad x = 0, h = 1, f(x) = e^x$$

$$R_n(0) = \frac{(0-1)^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

$$= \frac{(-1)^{n+1}}{(n+1)!} e^c$$

Since

x < c < x+h

0 < c < 0+1

0 < c < 1

$$\frac{1}{(n+1)!} < |R_n(0)| < \frac{e}{(n+1)!}$$

So if we want to find out how many terms it would require to get an approximation of e^1 magnitude of true error of less than 10^6,

$$\frac{e}{(n+1)!} < 10^{-6}$$

$$(n+1)! > 10^6 e$$

$$(n+1)! > 10^6 \times 3$$

$$n \geq 9$$

So 9 terms or more are needed to get a true error less than 10^{-6}.

3 Drawing line of best fit

The process of fitting a curve to a set of data is called curve-fitting.

Linearisation

A nonlinear relationship can be linearised and the resulting graph analysed to bring out the relationship between variables.

$y = ix + j$ ----------> linear or straight line graph, i=slope, j=intercept

$y = ix^2 + jx + k$ --------> quadratic graph or curve

$y = ix^n + jx + k$: $n >= 3$ --------> polynomial or sinusoidal wave form graph

$y = ie^x$ ---------------> ?? non-linear graph

$y = 2\log_x i3$ --------> ?? non-linear graph

Remember:

ln(x) is the natural logarithm to the base 'e' ≈ 2.71828, often referred to simply as "log."

$\log_{10}(x)$ is the common logarithm to the base 10, often referred to simply as "log."

In mathematical notation, the distinction is clear:

$\ln(x) = \log_e(x)$, where 'e' is the base of the natural logarithm.

$\log(x) = \log_{10}(x)$ where 10 is the base of common logarithm.

Case 1: $y = ae^x$.

(i) We could take the logarithm of both sides to base e:

$$\ln y = \ln(ae^x) = \ln a + \ln e^x = x + \ln a,$$

since $\ln e^x = x$. Thus, a plot of $\ln y$ against x gives a linear graph with slope unity and a y-intercept of $\ln a$.

(ii) We could also have plotted y against e^x. The result is a linear graph through the origin, with slope equal to a.

Case 2: $T = 2\pi\sqrt{\frac{l}{g}}$

We can write this expression in three different ways:

(i) $\ln T = \ln(2\pi) + \frac{1}{2}\ln\left(\frac{l}{g}\right) = \ln(2\pi) + \frac{1}{2}(\ln l - \ln g)$.

Rearranging, we obtain,

$$\ln T = \frac{1}{2}\ln l + \left(\ln(2\pi) - \frac{1}{2}\ln g\right)$$

writing this in the form $y = mx + c$, we see that a plot of $\ln T$ against $\ln l$ gives a slope of 0.5 and a $\ln T$ intercept of $\left(\ln(2\pi) - \frac{1}{2}\ln g\right)$. Once the intercept is read of the graph, you can then calculate the value of g.

(ii) $T = \frac{2\pi}{\sqrt{g}}\sqrt{l}$

A plot of T versus $\sqrt{l}$ gives a linear graph through the origin (as the intercept is zero). The slope of the graph is $\frac{2\pi}{\sqrt{g}}$, from which the value of g can be recovered.

Case 3: $N = N_0 e^{-\lambda t}$

The student can show that a plot of $\ln N$ versus t will give a linear graph with slope $-\lambda$, and $\ln N$ intercept is $\ln N_0$.

What other functions of N and t could you plot in order to get λ and N_0?

Case 4: $\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$

We rearrange the equation:

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$$

A plot of v^{-1} (y-axis) versus u^{-1} (x-axis) gives a slope of -1 and a vertical intercept of $\frac{1}{f}$.

Question 3a

A student obtained the following reading with a mirror in the laboratory.

u	10	20	30	40	50
v	-7	-10	-14	-15	-17

Linearise the relationship $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$. Plot the graph of v^{-1} versus u^{-1} and draw the line of best fit. Hence, find the focal length of the mirror. All distances are in cm.

Solution

u	v	1/u	1/v
10	-7	0.1	-0.14286
20	-10	0.05	-0.1
30	-14	0.033333	-0.07143
40	-15	0.025	-0.06667
50	-17	0.02	-0.05882

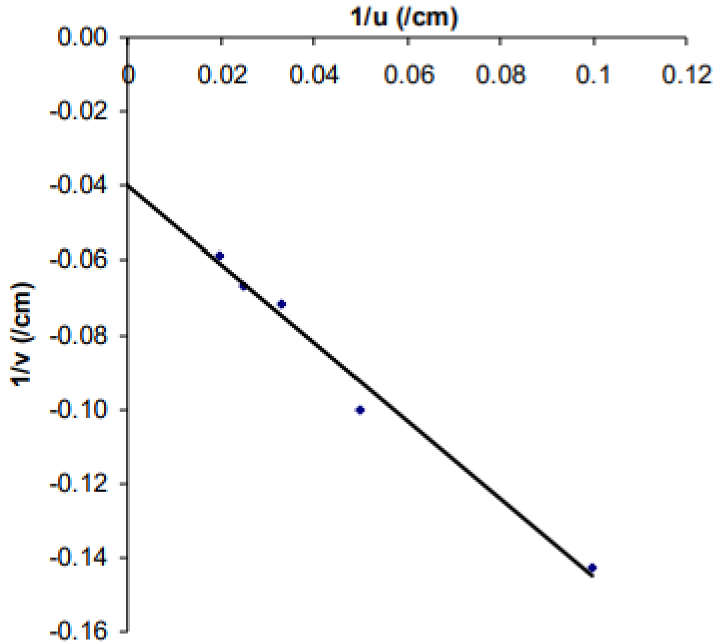


Fig. 1.1: Linear graph of the function $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$

The slope is -1.05 and the intercept -0.04. From $\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$, we see that the intercept is $\frac{1}{f} = -0.04$, or $f = -\frac{1}{0.04} = -25$ cm.

Method of least squares curve fitting

The least square method entails minimizing the sum of the squares of the difference between the measured value and the one predicted by the assumed equation.

$$m = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2}$$

$$c = \bar{y} - m\bar{x}$$

Question 3b

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

Thus, for the following set of readings:

t	5	12	19	26	33
x	23	28	32	38	41

The table can be extended to give

t	5	12	19	26	33	Σ=95	$\bar{t}$ =19
x	23	28	32	38	41	Σ=162	$\bar{x}$ =32.4
tx	115	336	608	988	1353	Σ=3400	$\overline{tx}$=680
t^2	25	144	361	676	1089	Σ=2295	$\overline{x^2}$ =459

$$m = \frac{\overline{tx} - \bar{t}\,\bar{x}}{\overline{t^2} - \bar{t}^2} = \frac{680 - 19 \times 32.4}{459 - 19^2} = 0.6571$$

$$c = \bar{x} - m\bar{t} = 32.4 - 0.6571 \times 19 = 19.9151$$

Hence, the relationship between x and t is,

$$x = 0.6571t + 19.9151$$

Method of group averages curve fitting

As the name implies, a set of data is divided into two groups, each of which is assumed to have a zero sum of residuals.

$$\bar{y}_1 = m\bar{x}_1 + c$$
$$\bar{y}_2 = m\bar{x}_2 + c$$

Subtracting,

$$\bar{y}_1 - \bar{y}_2 = m(\bar{x}_1 - \bar{x}_2)$$
$$m = \frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1 - \bar{x}_2}$$

and

$$c = \bar{y}_1 - m\bar{x}_1$$

Question 3c

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

First divide the data into 2 groups

t	5	12	19
x	23	28	32

and

t	26	33
x	38	41

The tables can be extended to give, for Table 3:

t	5	12	19	Σ=36	$\bar{t}_1$=12
x	23	28	32	Σ=83	$\bar{x}_1$=27.666667

and for Table 4:

t	26	33	Σ=59	$\bar{t}_2$=29.5
x	38	41	Σ=79	$\bar{x}_2$=39.5

$$m = \frac{\bar{x}_1 - \bar{x}_2}{\bar{t}_1 - \bar{t}_2} = \frac{27.666667 - 39.5}{12 - 29.5} = 0.67619$$

and

$$c = \bar{x}_1 - m\bar{t}_1 = 27.666667 - (0.67619 \times 12)$$
$$= 19.552387$$

Thus, the equation of best fit is,

$$x = 0.67619t + 19.552387$$

4 Methods of Linear Systems of Equation

Let us consider a linear system of equations

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

This can be written in the form

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

Gaussian Elimination involving 2 variables

Question 4a: find x and y given that

2x + 3y = 13

x – y = -1

The augmented matrix representing our system of two equations is

$$\left[\begin{array}{cc|c} 2 & 3 & 13 \\ 1 & -1 & -1 \end{array}\right]$$

By Gaussian elimination, we seek to make every entry below the main diagonal zero. This we achieve by reducing 1 to zero, making use of the first row.

Thus,

$$5y = 15 \Rightarrow y = 3$$

Substituting this in the first row gives

$$2x + 3(3) = 13$$

from which we obtain $x = 2$.

The process of reducing every element below the main diagonal to zero (row echelon form) is called Gaussian Elimination. That of substituting obtained values to calculate other variables is called Back Substitution.

Gaussian Elimination involving 2 variables

The same process can be carried over to the case of a system of three equations.

Question 4b:

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

The augmented matrix is

$$
\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]
$$

This yields (by Gaussian elimination)

$$
\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]
\begin{array}{c} \xrightarrow{(ii)'=(i)-2(ii)} \\ \xrightarrow{(iii)=(i)-(2/3)(iii)} \end{array}
\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 7/3 & 5/3 & 3 \end{array}\right]
$$

$$
\xrightarrow{(iii)''=(ii)'+(15/7)(iii)'}
\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right]
$$

Upon back substitution,

$$-10z = 10 \text{ or } z = -1$$

$$z = -1;\ y + z = 1 \Rightarrow y = 2;\ 2x + y - z = 5 \Rightarrow x = 1$$

Traditionally, in Mathematics, it is usual to use indices such as x_1, x_2, etc. instead of x, y, z. Do you have any idea why this is so? It is because if we stay with the alphabets, we shall soon run out of symbols. Bear in mind that not all the alphabets can be employed as variables; as an example, a, b, c are commonly used as constants. In addition, it makes it easy to associate the coefficients a_{11}, a_{12}, etc. with x_1, x_2, etc. respectively. More importantly in numerical work, it makes programming easier. For instance for our system of three equations, we could use the more general notation:

The general Gaussian elimination for linear system of 3 variables is thus given as:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] \begin{array}{c} \xrightarrow{(ii)'=a_{12}(i)-a_{11}(ii)} \\ \xrightarrow{(iii)'=a_{11}(i)-a_{11}(iii)} \end{array} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & a_{32}' & a_{33}' & a_{34}' \end{array}\right]$$

$$\xrightarrow{(iii)''=a_{32}(ii)'+a_{22}(iii)'} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & 0 & a_{33}'' & a_{34}'' \end{array}\right]$$

Question 4b-2:

$$\begin{aligned} -3x1 + 2x2 - x3 &= -1, \\ 6x1 - 6x2 + 7x3 &= -7, \\ 3x1 - 4x2 + 4x3 &= -6. \end{aligned}$$

First write out the augmented matrix

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 6 & -6 & 7 & -7 \\ 3 & -4 & 4 & -6 \end{pmatrix}$$

Perform row reduction by multiplying the first row by 2 (the lcm of all x1's), then add first row to both second and third row

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & -2 & 3 & -7 \end{pmatrix}$$

Perform row reduction by multiplying the second row by -1 (the lcm of x2's in rows 2 and 3), then add second row to third row

$$\begin{pmatrix} -3 & 2 & -1 & -1 \\ 0 & -2 & 5 & -9 \\ 0 & 0 & -2 & 2 \end{pmatrix}$$

$$-2x3 = 2 \rightarrow x3 = -1$$
$$-2x2 = -9 - 5x3 = -4 \rightarrow x2 = 2,$$
$$-3x1 = -1 - 2x2 + x3 = -6 \rightarrow x1 = 2.$$

Therefore, we have:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_n \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$$

Gauss-Jordan Elimination

This entails eliminating in addition to the entries below the major diagonal, the entries above it, so that the main matrix is a diagonal matrix. In that case, the solution to the system is given by dividing the element in the augmented part of the matrix by the diagonal element for that row.

$$2x + y - z = 5$$
$$x + 3y + 2z = 5$$
$$3x - 2y - 4z = 3$$

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

In solving the same problem using Gauss-Jordan elimination, we continue from completion of the Gaussian elimination part.

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right] \xrightarrow[(ii)'=(iii)-2(ii)]{(i)'=(iii)-10(i)} \left[\begin{array}{ccc|c} -20 & 10 & 0 & -40 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

$$\xrightarrow{(i)''=(ii)'+(i)'} \left[\begin{array}{ccc|c} -20 & 0 & 0 & -20 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

It follows that,

$-20x = -20$ or $x = 1$; $-10y = 20$ or $y = 2$; and $-10z = 10$ or $z = -1$

(lower and upper echelon) - LU decomposition

Suppose we could write the matrix

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

This implies that

$$l_{11}u_{11} = a_{11},\ l_{11}u_{12} = a_{12},\ l_{11}u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11},\ a_{22} = l_{21}u_{12} + l_{22}u_{22},\ a_{23} = l_{21}u_{13} + l_{22}u_{23}$$
$$a_{31} = l_{31}u_{11},\ a_{32} = l_{31}u_{12} + l_{32}u_{22},\ a_{33} = l_{31}u_{13} + l_{32}u_{23} + l_{33}u_{33}$$

Without loss of generality, we could set the diagonal elements of the L matrix equal to 1.

Then,

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

Multiplying out the right side of equation 3.19,

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

From the equality of matrices, this requires that,

$$u_{11} = a_{11}$$

$u_{12} = a_{12}$
$u_{13} = a_{13}$

$a_{21} = l_{21}u_{11} \qquad \Rightarrow l_{21} = a_{21} / u_{11} = a_{21} / a_{11}$

$a_{31} = l_{31}u_{11} \qquad \Rightarrow l_{31} = a_{31} / u_{11} = a_{31} / a_{11}$

$a_{22} = l_{21}u_{12} + u_{22}$, or $u_{22} = a_{22} - l_{21}u_{12} = a_{22} - \dfrac{a_{21}}{u_{11}}u_{12}$

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

$a_{23} = l_{21}u_{13} + u_{23}$, or $u_{23} = a_{23} - l_{21}u_{13} = a_{23} - \dfrac{a_{21}}{u_{11}}u_{13}$

$$\Rightarrow u_{23} = a_{23} - \frac{a_{21}}{a_{11}}a_{13}$$

$$l_{32} = \frac{a_{32} - l_{31}u_{12}}{u_{22}} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{u_{11}}u_{12}\right]$$

$$\Rightarrow l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}}a_{12}\right]$$

$a_{32} = l_{31}u_{12} + l_{32}u_{22}$

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

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

You can see that we have determined all the nine elements of the two matrices in terms of the elements of the original matrix.

Once we have obtained L and U, then we can write the original equation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

as

$$LU\mathbf{x} = \mathbf{y}$$

where $\mathbf{x}$ and $\mathbf{y}$ are column vectors.

We shall write $\mathbf{w} = U\mathbf{x}$
Then,

$$L\mathbf{w} = \mathbf{y}$$

Now we continue to solving **Question5b** again using LU decomposition

The corresponding matrix is

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

$$u_{11} = a_{11} = 2$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = -1$$
$$l_{21} = a_{21} / a_{11} = 1/2$$
$$l_{31} = a_{31} / a_{11} = 3/2$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 3 - \frac{1}{2}(1) = 5/2$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - \frac{1}{2}(-1) = 2 + \frac{1}{2} = 5/2$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{5/2} \left[-2 - \frac{3}{2}(1) \right] = -7/5$$
$$u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23} = -4 - (3/2)(-1) - (-7/5)(5/2) = 1$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 5/2 & 5/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 2 \\ 3 & -2 & -4 \end{bmatrix}$$

The above decomposition is correct as the multiplication of L and U gives the original matrix.

The original equation is equivalent to

$$LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$$

$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 3 \end{bmatrix}$$

Solving,

$$w_1 = \quad 5$$

$$\frac{1}{2}w_1 + w_2 = 5 \text{ or } w_2 = 5 - \frac{1}{2}w_1 = 5 - \frac{1}{2}(5) = \frac{5}{2}$$

$$\frac{3}{2}w_1 - \frac{7}{5}w_2 + w_3 = 3, \text{ or } w_3 = 3 + \frac{7}{5}w_2 - \frac{3}{2}w_1 = 3 + \frac{7}{5}\left(\frac{5}{2}\right) - \frac{3}{2}(5) = -1$$

$U\mathbf{x} = \mathbf{w}$ implies:

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

By back substitution,

$$x_3 = -1$$

$$\frac{5}{2}x_2 + \frac{5}{2}x_3 = \frac{5}{2} \Rightarrow \frac{5}{2}x_2 = \frac{5}{2} - \frac{5}{2}x_3 = \frac{5}{2} - \frac{5}{2}(-1) = 5$$

$$x_2 = 2$$

$$2x_1 + x_2 - x_3 = 5$$

$$x_1 = \frac{5 - x_2 + x_3}{2} = \frac{5 - 2 + (-1)}{2} = 1$$

The solution set is therefore,

$$x_1 = 1,\ y = 2,\ z = -1.$$

Question 4c

Solve the system of linear equations $x+y+z=-1$, $x+2y+2z=-4$, $9x+6y+z=7$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

(i) Gaussian elimination

Initial augmented matrix

1	1	1	-1
1	2	2	-4
9	6	1	7

First round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	-3	-8	16

Second round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

Last matrix for Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

First round of Jordan elimination

5	5	0	2
0	5	0	-8
0	0	-5	7

Second round of Jordan elimination

-25	0	0	-50
0	5	0	-8
0	0	-5	7

(iii) LU decomposition

$x+y+z=-1$

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

$9x+6y+z=7$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7\text{, or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$x_3 = -7/5 = -1.4$

$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$

$x_2 = -8/5 = -1.6$

$x_1 + x_2 + x_3 = -1$

$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \dfrac{10}{5} = 2$

The solution set is therefore,

$x_1 = 2$, $y = -1.6$, $z = -1.4$.

Question 4d

Solve the system of linear equations $x + 2y + 2z = -2$, $2x + 2y + z = -4$, $9x + 6y + 2z = -14$ using the method of

(i) Gaussian elimination
(ii) Gauss-Jordan elimination
(iii) LU decomposition

Gaussian elimination

Initial augmented matrix

1	2	2	-2
2	2	1	-4
9	6	2	-14

First round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	-12	-16	4

Second round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

Answers

x	0
y	-3
z	2

Gauss-Jordan elimination

Last matrix for Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

First round of Jordan elimination

4	8	0	-24
0	-8	0	24
0	0	-4	-8

Second round of Jordan elimination

32	0	0	0
0	-8	0	24
0	0	-4	-8

LU decomposition

$$x + 2y + 2z = -4$$
$$9x + 6y + z = 7$$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

W got the decomposition right, as the multiplication of the L and U gives the original matrix.

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7, \text{ or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$$x_3 = -7/5 = -1.4$$
$$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$$
$$x_2 = -8/5 = -1.6$$
$$x_1 + x_2 + x_3 = -1$$
$$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \frac{10}{5} = 2$$

The solution set is therefore,

$$x_1 = 2,\ y = -1.6,\ z = -1.4.$$

It is usually a good practice to revert to fractions to avoid incurring rounding errors.

5 Methods for finding Roots of Algebraic and Transcendental equations

In all scientific fields, there's always the need to find the root of an equation, equivalently the zero of a function. Numerical methods allow for more complicated cases of handling roots of quadratic and polynomial equations.

Bisection method

As the name implies, we obtain the points x_1 and x_2, such that $f(x_2)\, f(x_1) < 0$, meaning that the value of f has opposite signs at the two points, which points to the fact that a root exists between x_1 and x_2. We approximate this root by the average of the two, i.e., $(x_1 + x_2) / 2$. Let this be x_3. Then we evaluate $f(x_3)$. x_3 is then combined with x_1 or x_2, depending on the one at which the sign of the function is opposite $f(x_3)$. This gives x_4. This process is repeated until f(x) attains the prescribed tolerance. The convergence of the Bisection method is slow and steady.

Bisection Method

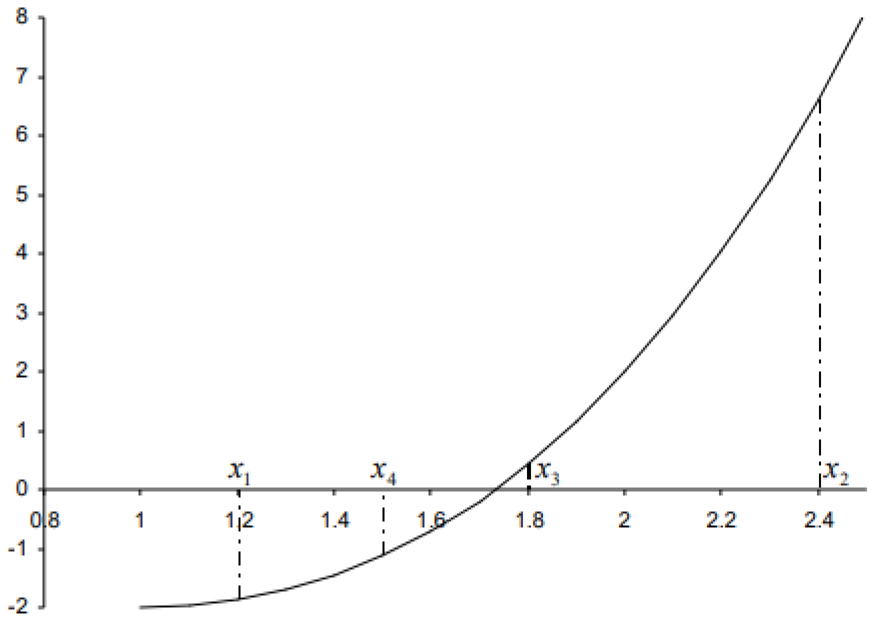


Fig.

- The **Bisection method** is one of the simplest methods to find a zero of a nonlinear function.
- It is also called **interval halving** method.
- To use the Bisection method, one needs an initial interval that is known to contain a zero of the function.
- The method systematically reduces the interval. It does this by dividing the interval into two equal parts, performs a simple test and based on the result of the test, half of the interval is thrown away.
- The procedure is repeated until the desired interval size is obtained.

Bisection Algorithm

Assumptions:

- $f(x)$ is continuous on $[a,b]$
- $f(a)\, f(b) < 0$

Loop

1. Compute the mid point $c=(a+b)/2$
2. Evaluate $f(c)$
3. If $f(a)\, f(c) < 0$ then new interval $[a, c]$
   If $f(a)\, f(c) > 0$ then new interval $[c, b]$

End loop

Question 5a

Can you use Bisection method to find a zero of:

$f(x) = x^3 - 3x + 1$ in the interval [0,2]?

Answer:

$f(x)$ is continuous on [0,2]

and $\mathrm{f}(0) * \mathrm{f}(2) = (1)(3) = 3 > 0$

$\Rightarrow$ Assumptions are not satisfied

$\Rightarrow$ Bisection method can not be used

Question 5b

Can you use Bisection method to find a zero of:

$f(x) = x^3 - 3x + 1$ in the interval [0,1]?

Answer:

$f(x)$ is continuous on [0,1]

and $\mathrm{f}(0) * \mathrm{f}(1) = (1)(-1) = -1 < 0$

$\Rightarrow$ Assumptions are satisfied

$\Rightarrow$ Bisection method can be used

Question 5c

Find the root of:

$$f(x) = x^3 - 3x + 1 \text{ in the interval } :[0,1]$$

* $f(x)$ is continuous
* $f(0) = 1,\ f(1) = -1 \Rightarrow f(a)\, f(b) < 0$

$\Rightarrow$ Bisection method can be used to find the root

Iteration	a	b	c= (a+b)/2	f(c)	(b-a)/2
1	0	1	0.5	-0.375	0.5
2	0	0.5	0.25	0.266	0.25
3	0.25	0.5	.375	-7.23E-3	0.125
4	0.25	0.375	0.3125	9.30E-2	0.0625
5	0.3125	0.375	0.34375	9.37E-3	0.03125

Question 5d

Find a zero of the function $f(x) = 2x^3 - 3x^2 - 2x + 3$ between the points 1.4 and 1.7, using the bisection method. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(1.4) = -0.192$$
$$f(1.7) = 0.756$$
$$x_3 = \frac{1.4 + 1.7}{2} = 1.55$$

$$f(1.55) = 1.4025 \times 10^{-1}$$
$$x_4 = \frac{1.55 + 1.4}{2} = 1.475$$
$$f(1.475) = -0.0588$$
$$x_5 = \frac{1.55 + 1.475}{2} = 1.5125$$

This confirm that the Table for Bisection method is indeed true

n	x	$f(x)$
1	1.55	0.14025
2	1.475	-5.88E-02
3	1.5125	3.22E-02
4	1.49375	-1.54E-02
5	1.503125	7.87E-03
6	1.498437	-3.89E-03
7	1.500781	1.96E-03
8	1.499609	-9.76E-04
9	1.500195	4.89E-04
10	1.499902	-2.44E-04
11	1.500049	1.22E-04
12	1.499976	-6.10E-05

CONVERGENCE ANALYSIS OF BISECTION METHOD

Bisection method obtains an interval that is guaranteed to contain a zero of the function.

Questions:

- What is the best estimate of the zero of $f(x)$?
- What is the error level in the obtained estimate?

The best estimate of the zero of the function $f(x)$ after the first iteration of the Bisection method is the mid point of the initial interval:

$$\textit{Estimate of the zero}: \ r = \frac{b+a}{2}$$

$$\textit{Error} \le \frac{b-a}{2}$$

$$\textit{Given} \quad f(x),\ a,\ b,\ \textit{and} \ \varepsilon$$

How many iterations are needed such that: $|x - r| \le \varepsilon$ where r is the zero of $f(x)$ and x is the bisection estimate (i.e., $x = c_k$)?

$$\boxed{n \ge \frac{\log(b-a) - \log(\varepsilon)}{\log(2)}}$$

$$\boxed{n \ge \log_2\left(\frac{\text{width of initial interval}}{\text{desired error}}\right) = \log_2\left(\frac{b-a}{\varepsilon}\right)}$$

Two common stopping criteria

1. Stop after a fixed number of iterations
2. Stop when the absolute error is less than a specified value

After n iterations :

$$\left|error\right| = \left|r - c_n\right| \le E_a^n = \frac{b-a}{2^n} = \frac{\Delta x^0}{2^n}$$

Question 5e

$a = 6,\ b = 7,\ \varepsilon = 0.0005$

How many iterations are needed such that: $\left|x - r\right| \le \varepsilon$?

$$n \ge \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1) - \log(0.0005)}{\log(2)} = 10.9658$$

$$\Rightarrow n \ge 11$$

Question 5f

- Use Bisection method to find a root of the equation x = cos(x) with absolute error <0.02
  (assume the initial interval [0.5, 0.9])

What is *f (x)* ?

Are the assumptions satisfied ?

How many iterations are needed ?

How to compute the new estimate ?

Question 5f (i) – what is f(x)?

x = cos(x)

f(x) = x – cos(x)

Question 5f (ii) – Are the assumptions satisfied?

Assuming interval [0.5, 0.9]

f(0.5) = 0.5 – cos(0.5) = -0.3776; This is a negative value

f(0.9) = 0.9 – cos(0.9) = 0.2784; This is a positive value

f(0.5)*f(0.9) = -0.3776 * 0.2764 < 0; Assumption is therefore satisfied.

Bisection method can be used.

Question 5f (iii) – How many iterations are needed?

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

a = 0.5, b = 0.9, ε = 0.02

n $\geq$ [log(0.9 – 0.5) – log(0.02)] / log(2)

n $\geq$ [-0.3979 - -1.6990] / 0.3010

n $\geq$ 1.3011 / 0.3010

n $\geq$ 4.3226

n $\geq$ 5

Question 5f (iii) – How to compute the new estimate?

$$\textit{Estimate of the zero}: \ r = \frac{b+a}{2} \quad , \quad Error \leq \frac{b-a}{2}$$

r1 = (0.9 + 0.5) / 2 = 0.7; **Error < (0.9 - 0.5)/2 $\leq$ 0.2;**

f(0.7) = 0.7 – cos(0.7) = 0.7 – 0.9999 = -0.2999

f(0.5) = -0.3776; f(0.9) = 0.2784; f(0.7) = -0.2999

r2 = (0.7 + 0.9) / 2 = 0.8, **Error < (0.9 - 0.7) / 2 ≤ 0.1**

f(0.8) = 0.8 – cos(0.8) = 0.8 – 0.9999 = -0.1999

f(0.7) = -0.2999; f(0.9) = 0.2784; f(0.8) = -0.1999

r3 = (0.8 + 0.9) / 2 = 0.85, **Error < (0.9 - 0.8) / 2 ≤ 0.5**

f(0.85) = 0.85 – cos(0.85) = 0.85 – 0.9999 = -0.1499

f(0.8) = -0.1999; f(0.9) = 0.2784; f(0.85) = -0.1499

r4 = (0.85 + 0.9) / 2 = 0.875, **Error < (0.9 - 0.85) / 2 ≤ 0.025**

f(0.875) = 0.875 – cos(0.875) = 0.875 – 0.9999 = -0.1249

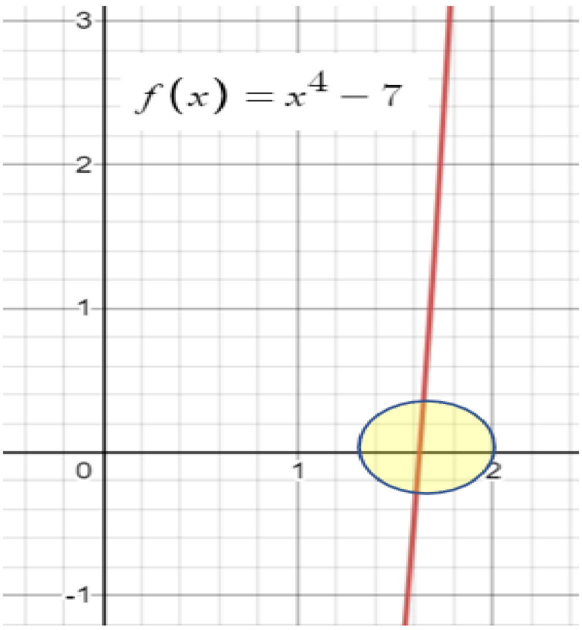
f(0.85) = -0.1499; f(0.9) = 0.2784; f(0.875) = -0.1249

r5 = (0.875 + 0.9) / 2 = 0.8875, **Error < (0.9 - 0.875) / 2 ≤ 0.02**

f(0.8875) = 0.8875 – cos(0.8875) = 0.8875 – 0.9999 = -0.1124

f(0.875) = -0.1249; f(0.9) = 0.2784; f(0.8875) = -0.1124

Question 5g



Find the 3rd approximation of the root of $f(x) = x^4 – 7$ using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(left)	f(mid)	f(right)	New Interval	Midpoint
f(1) = -6	f(1.5) = -2	f(2) = 9	(1.5, 2)	1.75
f(1.5) = -2	f(1.75) = 2.4	f(2) = 9	(1.5, 1.75)	1.625
f(1.5) = -2	f(1.625) = -.03	f(1.75) = 2.4	(1.625, 1.75)	1.6875

$f(x) = x^4 - 7$

$f(2) = (2)^4 - 7 = 9$; this is positive

$f(1) = (1)^4 - 7 = -6$; this is negative

$f(2)*f(1) = 9 * -6 < 0$; Assumption is therefore satisfied. Bisection method can be used.

for;

Starting interval (1, 2)

mid x = [2+1] / 2 = 1.5; Initial estimate

$f(mid) = f(1.5) = (1.5)^4 - 7 = 5.0625 - 7 = -1.9375$.

for;

f(2) = 9, f(1) = -6, **f(1.5) = -1.9375**

Next interval (2, 1.5)

mid x = [2+1.5]/2 = 1.75; first approximation

$f(mid) = f(1.75) = (1.75)^4 - 7 = 9.3789 - 7 = 2.3789$.

for;

f(2) = 9**,** **f(1.5) = -1.9375**, **f(1.75) = 2.3789**

Next interval (1.75, 1.5)

mid x = [1.75+1.5]/2 = 1.625; second approximation

$f(mid) = f(1.625) = (1.625)^4 - 7 = 6.9729 - 7 = -0.0271$.

for;

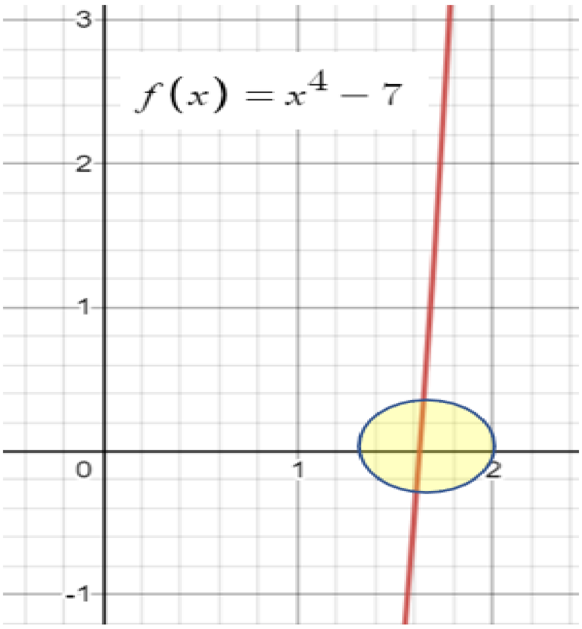
f(1.75) = 2.3789, f(1.5) = -1.9375, **f(1.625) = -0.0271**

Next interval (1.75, 1.625)

mid x = [1.75+1.625]/2 = 1.6875; third approximation

$f(mid) = f(1.6875) = (1.6875)^4 - 7 = 8.1091 - 7 = 1.1091$. **Stop.**

Question 5h



Find the 3rd approximation of the root of f(x) = 10 – x^2 using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(x) = 10 – x^2

f(2) = 10 - $(2)^2$ = 6; this is positive

f(1) = 10 - $(1)^2$ = 9; this is also positive

f(2)*f(1) = 6 * 9 !< 0; Assumption is NOT satisfied. Bisection method cannot be used.

Question 5h

Given a floating ball with a specific gravity of 0.6 and has a radius of 5.5 cm. You are asked to find the depth to which the ball is submerged when floating in water.

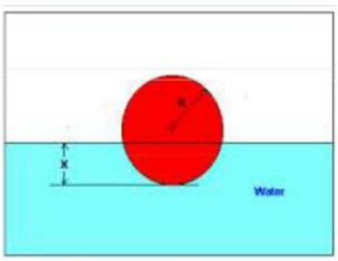


Diagram of the floating ball

The equation that gives the depth x to which the ball is submerged under water is given by:

$$x^3 - 0.165x^2 + 3.993x10^{-4} = 0$$

1. Use the Bisection method of finding roots of equations to find the depth x to which the ball is submerged under water. Conduct three iterations to estimate the root of the above equation.
2. Find the absolute relative approximate error at the end of each iteration.
3. Use both false position and newton methods to solve the roots of the equations.

Hint: From the Physics point of view, the ball would be submerged between $x = 0$ and $x = 2R$, where R = radius of the ball.

That is, $0 \le x \le 2R$ ==> $0 \le x \le 2(0.055)$ ==> $0 \le x \le 0.11$

Newton-Raphson Method

(Also known as Newton's Method)

Given an initial guess of the root x_0, Newton-Raphson method uses information about the function and its derivative at that point to find a better guess of the root.

Assumptions:

- $f(x)$ is continuous and the first derivative is known
- An initial guess x_0 such that $f'(x_0) \neq 0$ is given

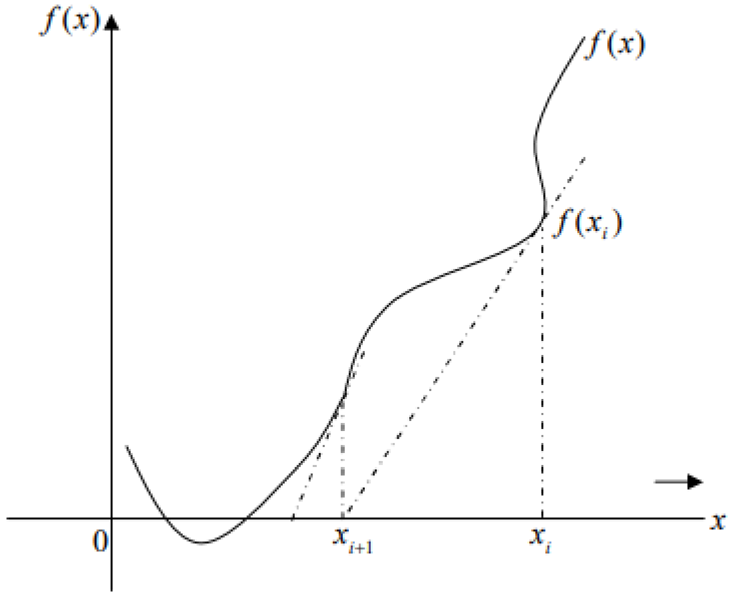
It is quite clear that the function $f(x)$ must be differentiable for you to be able apply the Newton-Raphson method.

More generally,

$$x_{i+1} = x_i + \Delta x = x_i - \frac{f(x_i)}{f'(x_i)}$$

With an initial guess of x_0, we can then get a sequence x_1, x_2, …, which we expect to converge to the root of the equation.

Newton-Raphson method is equivalent to taking the slope of the function f(x) at the i^{th} iterative point, and the next approximation is the point where the slope intersects the x axis.



Question 5i

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ starting with the point 1.4, using the **Newton-Raphson Method**. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(x) = 2x^3 - 3x^2 - 2x + 3$$

$$f'(x) = 6x^2 - 6x - 2$$

$$x_0 = 1.4$$

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

$$= \frac{6x_0^3 - 6x_0^2 - 2x_0 - 2x_0^3 + 3x_0^2 + 2x_0 - 3}{6x_0^2 - 6x_0 - 2}$$

$$= \frac{4x_0^3 - 3x_0^2 - 3}{6x_0^2 - 6x_0 - 2}$$

$$= \frac{4(1.4)^3 - 3(1.4)^2 - 3}{6(1.4)^2 - 6(1.4) - 2}$$

$$= 1.5412$$

$x_1 = 1.5412, \; |x_1 - x_0| = 0.1412$

$x_2 = 1.5035, \; |x_2 - x_1| = 0.0377$

$x_3 = 1.5, \quad |x_3 - x_2| = 0.0035$

$x_4 = 1.5, \quad |x_4 - x_3| = 0$

Question 5j-2

Find a zero of the function $f(x) = x^3 - 2x^2 + x - 3$, $x_0 = 4$

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

Iteration 1: $x_1 = x_0 - \dfrac{f(x_0)}{f'(x_0)} = 4 - \dfrac{33}{33} = 3$

Iteration 2: $x_2 = x_1 - \dfrac{f(x_1)}{f'(x_1)} = 3 - \dfrac{9}{16} = 2.4375$

Iteration 3: $x_3 = x_2 - \dfrac{f(x_2)}{f'(x_2)} = 2.4375 - \dfrac{2.0369}{9.0742} = 2.2130$

k (Iteration)	$\mathbf{x_k}$	$\mathbf{f(x_k)}$	$\mathbf{f'(x_k)}$	$\mathbf{x_{k+1}}$	$\mathbf{\lvert x_{k+1} - x_k \rvert}$
0	**4**	**33**	**33**	**3**	1
1	**3**	**9**	**16**	**2.4375**	0.5625
2	**2.4375**	**2.0369**	**9.0742**	2.2130	0.2245
3	2.2130	0.2564	6.8404	2.1756	0.0384
4	2.1756	0.0065	6.4969	2.1746	0.0010

Question 5j-2

Use Newton's Method to find a root of:

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

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

Five Iterations of the Solution

k	x_k	$f(x_k)$	$f'(x_k)$	ERROR
0	1.0000	-1.0000	2.0000	
1	1.5000	0.8750	5.7500	0.1522
2	1.3478	0.1007	4.4499	0.0226
3	1.3252	0.0021	4.2685	0.0005
4	1.3247	0.0000	4.2646	0.0000
5	1.3247	0.0000	4.2646	0.0000

Question 5j-3

Use Newton's Method to find a root of:

$f(x) = e^{-x} - x$

Use the initial point: $x_0 = 1$.

Stop after three iterations, or

if $|x_{k+1} - x_k| < 0.001$, or

if $|f(x_k)| < 0.0001$.

x_k	$f(x_k)$	$f'(x_k)$	$\dfrac{f(x_k)}{f'(x_k)}$
1.0000	-0.6321	-1.3679	0.4621
0.5379	0.0461	-1.5840	-0.0291
0.5670	0.0002	-1.5672	-0.0002
0.5671	0.0000	-1.5671	-0.0000

Question 5k

Estimates of the root of: x-cos(x)=0.

0.60000000000000		**Initial guess**
0.744017319445980	1	correct digit
0.73909047688624	4	correct digits
0.73908513322147	10	correct digits
0.73908513321516	14	correct digits

Question 5k-2

Given the equation: $\boldsymbol{f(x) = x^3 - 10 = 0}$ which root lies between 2 and 3. Find the real root using the Newton Raphson method (Up to 3 iterations and correct to 4 decimal places).

Taking $x_0 = 2$.

Let

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

and

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

Given $x_0 = 2$

$$\boldsymbol{x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}} \qquad (1)$$

Where i=0,1,2,3...

Let $f(x_i) = f(x)$ and $f'(x_i) = f'(x)$

First Iteration:

Hence, substituting $i = 0$ into equation (1) to get the first approximation of the root

We have

$$\boldsymbol{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}} \qquad \textbf{(2)}$$

$x_0 = 2;\ f(x_0) = -2;\ f'(x_0) = 12$

Note: $f(x_0)\ and\ f'(x_0)$ are derived by substituting $x_0 = 2$ into $f(x_i)\ and\ f'(x_i)$ where $i = 0$

Substituting the values into equation (2)

$x_1 = 2 - \frac{(-2)}{12} = 2.1667$

$x_1 = 2.1667$

Second Iteration:

Hence, substituting $i = 1$ into equation (1) to get the second approximation of the root

We have

$$\boldsymbol{x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}} \qquad \textbf{(3)}$$

$x_1 = 2.1667;\ f(x_1) = 0.1718;\ f'(x_1) = 14.0838$

Note: $f(x_1)\ and\ f'(x_1)$ are derived by substituting $x_1 = 2.1667$ into $f(x_i)\ and\ f'(x_i)$ where $i = 1$

Substituting the values into equation (3)

$x_2 = 2.1667 - \frac{0.1718}{14.0838} = 2.1545$

$x_2 = 2.1545$

Third Iteration:

Hence, substituting $i = 2$ into equation (1) to get the third approximation of the root

We have

$$\boldsymbol{x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}} \qquad \textbf{(4)}$$

$x_2 = 2.1545;\ f(x_2) = 0.0009;\ f'(x_2) = 13.9256$

Note: $f(x_1)\ and\ f'(x_1)$ are derived by substituting $x_2 = 2.1545$ into $f(x_i)\ and\ f'(x_i)$ where $i = 2$

Substituting the values into equation (4)

$x_3 = 2.1545 - \frac{0.0009}{13.9256} = 2.1544$

$x_3 = 2.1544$

Summary Table of Iterations

i (iteration)	X_i	$f(x_i)$	$f'(x_i)$	X_{i+1}	$\lvert X_{i+1} - X_i \rvert$
0	2	-2	12	2.1667	0.1667
1	2.1667	0.1718	14.0838	2.1545	0.0122
2	2.1545	0.0009	13.9256	2.1544	0.0001
3	2.1544	-0.0005	13.9243	2.1544	0

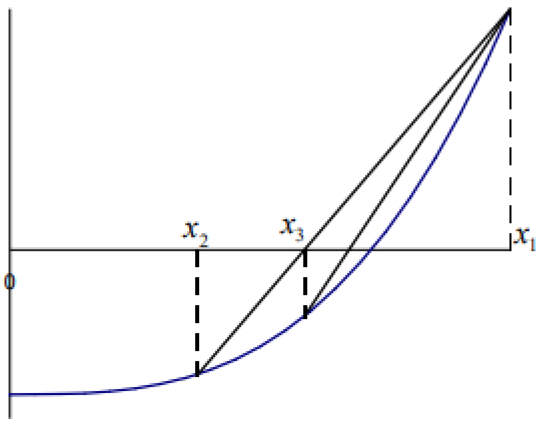
Thus, the root of the equation: $f(x) = x^3 - 10 = 0$ is 2.1544

Regula-falsi method

- Also known as the false-position method, or linear interpolation method.

A regula-falsi or a method of false position assumes a test value for the solution of the equation.

- The *regula falsi* method start with two point, (a, f(a)) and (b,f(b)), satisfying the condition that f(a)f(b)<0.



Then, for an arbitrary x and the corresponding y,

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

gives the equation of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Setting $y=0$, that is, where the chord crosses the x-axis,

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

Then, we evaluate $f(x_3)$. Just as in the case of root-bisection, if the sign is opposite that of $f(x_1)$, then a root lies in-between x_1 and x_3. Then, we replace x_2 by x_3 in equation

In just the same way, if the root lies between x_1 and x_3, we replace x_2 by x_1. We shall repeat this procedure until we are as close to the root as desired.

Question 5k-3

Find a zero of the function $f(x) = 2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the regula-falsi method.

$$f(1.4) = -0.192\,,\ \ f(1.7) = 0.756$$

A solution lies between x = 1.4 and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.4607595$

$$f(1.4607595) = -0.088983$$

The root lies between 1.46076 and 1.7. Let $x_1 = 1.46076$ and $x_2 = 1.7$.

$$x_4 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4607595 - (-0.088983)\frac{1.7 - 1.46076}{0.756 - (-0.088983)}$$

$= 1.485953$

Table for Regula-falsi method

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.495149	-0.011985
4	1.498346	-0.004118
5	1.499439	-0.001401
6	1.499810	-0.000475
7	1.499936	-0.000161
8	1.499978	-0.000055

Question 5L

- **Finding the Cube Root of 2 Using Regula Falsi**

- Since f(1)= -1, f(2)=6, we take as our starting bounds on the zero a=1 and b=2.
- Our first approximation to the zero is

$$x = b - \frac{b-a}{f(b) - f(a)}(f(b)) = 2 - \frac{2-1}{6+1}(6)$$
$$= 2 - 6/7 = 8/7 \approx 1.1429$$

- We then find the value of the function:

$$y = f(x) = (8/7)^3 - 2 \approx -0.5073$$

- Since f(a) and y are both negative, but y and f(b) have opposite signs

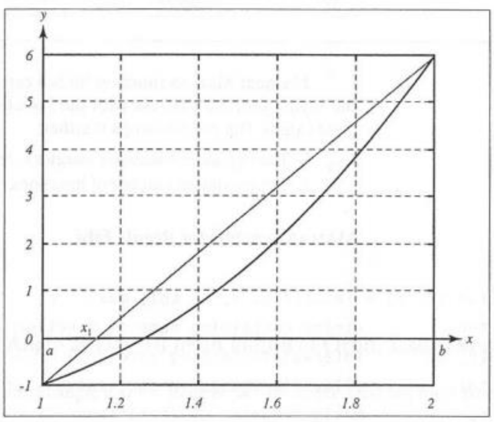


FIGURE 2.5 Graph of $y = x^3 - 2$ and approximation line on the interval [1, 2].

- Calculation of $\sqrt[3]{2}$ using *regula falsi*.

Step	a	b	x	y
1	1	2	1.1429	−0.50729
2	1.1429	2	1.2097	−0.22986
3	1.2097	2	1.2388	−0.098736
4	1.2388	2	1.2512	−0.041433
5	1.2512	2	1.2563	−0.017216
6	1.2563	2	1.2584	−0.0071239
7	1.2584	2	1.2593	−0.0029429
8	1.2593	2	1.2597	−0.0012148
9	1.2597	2	1.2598	−0.00050134
10	1.2598	2	1.2599	−0.00020687

Question 5L-2

Using the method of false position, find the real root of the equation $\boldsymbol{x^3 - 2x - 5 = 0}$. Where the real root lies between 2 and 2.1. (Up to 3 iterations and correct to 3 decimal places).

Let

$$f(x) = x^3 - 2x - 5$$

Given the roots as 2 and 2.1, therefore $a = 2 \text{ } and \text{ } b = 2.1$

To find our approximation, we use the formula:

$$\boldsymbol{x_i = \frac{af(b) - bf(a)}{f(b) - f(a)}} \qquad \textbf{(1)}$$

$$\boldsymbol{where \ i = 1,2,3 \ldots}$$

First Iteration:

Where $i = 1; a = 2; b = 2.1; f(a) = -1; \ f(b) = 0.061$

Note: $f(a) \ and \ f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 1$

Therefore, substituting in the values into equation (1) we have

$$x_1 = \frac{2(0.061) - 2.1(-1)}{0.061 - (-1)}$$

$$x_1 = \frac{0.122 + 2.1}{1.061} = \frac{2.222}{1.061} = 2.0942$$

$$x_1 = 2.0942$$

Thus,

$$f(x_1) = f(2.0942) = (2.0942)^3 - 2(2.0942) - 5$$

$$f(x_1) = -0.0039$$

Since $f(x_1)$ is a negative value, therefore, the new root lies between (2.0942, 2.1) and $a = 2.0942$; $b = 2.1$

Second Iteration:

Where $i = 2; a = 2.0942; b = 2.1; f(a) = -0.0039; f(b) = 0.061$

Note: $f(a)$ *and* $f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 2$

Therefore, substituting in the values into equation (1) we have

$$x_2 = \frac{2.0942(0.061)-2.1(-0.0039)}{0.061-(-0.0039)}$$

$$x_2 = \frac{0.12775 + 0.00819}{0.0649} = \frac{0.13594}{0.0649} = 2.0946$$

$$x_2 = 2.0946$$

Thus,

$$f(x_2) = f(2.0946) = (2.0946)^3 - 2(2.0946) - 5$$

$$f(x_2) = 0.0005$$

Since $f(x_2)$ is a positive value, therefore, the new root lies between (2.0942, 2.0946) and $a = 2.0942$; $b = 2.0946$

Third Iteration:

Where $i = 3; a = 2.0942; b = 2.0946; f(a) = -0.0039; f(b) = 0.0005$

Note: $f(a)$ *and* $f(b)$ are derived by substituting the values of a and b respectively into $f(x_i)$ where $i = 3$

Therefore, substituting in the values into equation (1) we have

$$x_3 = \frac{2.0942(0.0005)-2.0946(-0.0039)}{0.0005-(-0.0039)}$$

$$x_3 = \frac{0.00105 + 0.00817}{0.0044} = \frac{0.00922}{0.0044} = 2.0952$$

$$x_3 = 2.0952$$

Summary Table of Iterations

i (iteration)	a	b	X_i	$f(x_i)$
1	2	2.1	2.0942	-0.0039
2	2.0942	2.1	2.0946	0.0005
3	2.0942	2.0946	2.0952	0.007

Therefore, after three iterations, the required approximate root correct to 3 decimal places is <u>2.095</u>

Secant Method

In the case of the secant method, it is not necessary that the root lie between the two initial points. As such, the condition $f(x_1)f(x_2) < 0$ is not needed. Following the same analysis with the case of the regula-falsi method,

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

Setting $y = 0$ gives

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

Thus, having found x_n, we can obtain x_{n+1} as,

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})},\ n = 2, 3, \ldots$$

By inspection, if $f(x_n) - f(x_{n-1}) = 0$, the sequence does not converge, because the formula fails to work for x_{n+1}. The regula-falsi scheme does not have this problem as the associated sequence always converges.

Question 5m

Find the roots of the equation f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the secant method.

$x_1 = 1.4,\ x_2 = 1.7$

$f(1.4) = -0.192$, $f(1.7) = 0.756$

A solution lies between $x = 1.4$ and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.460759$

$f(x_3) = -0.088983$

$$x_4 = x_3 - f(x_3)\frac{x_3 - x_2}{f(x_3) - f(x_2)} = 1.460759 - (-0.088983) \times \frac{1.460759 - 1.7}{-0.088983 - 0.756}$$

$= 1.485953$

If the scheme continues, the table for secant method will be

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.501487	0.003730
4	1.499949	-0.000129
5	1.500000	0.000000

Question 5n

Find the roots of the equation by the secant method:

$$f(x) = x^2 - 2x + 0.5$$

$$x_0 = 0$$

$$x_1 = 1$$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

Question 5n-i

Find the roots of the equation by the secant method:

Use Secant method to find the root of:

$f(x) = x^6 - x - 1$

Two initial points $x_0 = 1$ *and* $x_1 = 1.5$

$$x_{i+1} = x_i - f(x_i)\frac{(x_i - x_{i-1})}{f(x_i) - f(x_{i-1})}$$

k	x_k	$f(x_k)$
0	1.0000	-1.0000
1	1.5000	8.8906
2	1.0506	-0.7062
3	1.0836	-0.4645
4	1.1472	0.1321
5	1.1331	-0.0165
6	1.1347	-0.0005

Question 5n-ii

Given the equation: $\boldsymbol{f(x) = x^3 - 5x + 1}$, where x_0 and x_1 are 2 and 2.5 respectively. Find the real root using the Secant method. (Up to 4 iterations and correct to 4 decimal places).

Let

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

Given $x_0 = 2\ and\ x_1 = 2.5$

To find our approximation, we use the formula:

$$\boldsymbol{x_{i+1} = x_i - f(x_i)\left(\frac{x_i - x_{i-1}}{f(x_i) - f(x_{i-1})}\right)} \qquad (\mathbf{1})$$

where i=1,2,3 …

Let $f(x_i) = f(x)$, where i=1,2,3 …

First Iteration:

Hence, substituting $i = 1$ into equation (1) to get the first approximation of the root

We have

$$\boldsymbol{x_2 = x_1 - f(x_1)\left(\frac{x_1 - x_0}{f(x_1) - f(x_0)}\right)} \qquad (\mathbf{2})$$

$x_0 = 2; x_1 = 2.5;\ f(x_0) = -1;\ f(x_1) = 4.125$

Note: $f(x_0)\ and\ f(x_1)$ are derived by substituting $x_0 = 2,\ x_1 = 2.5$ into $f(x_0)\ and\ f(x_1)$ respectively.

Substituting the values into equation (2)

$$x_2 = 2.5 - (4.125)\left(\frac{2.5 - 2}{4.125 - (-1)}\right)$$

$$x_2 = 2.5 - (4.125)\left(\frac{0.5}{5.125}\right)$$

$$x_2 = 2.5 - (4.125)(0.09756)$$

$$x_2 = 2.5 - 0.4024$$

$$x_2 = 2.0976$$

Thus,

$$f(x_2) = f(2.0976) = (2.0976)^3 - 5(2.0976) + 1$$

$$f(x_2) = -0.2587$$

Second Iteration:

Hence, substituting $i = 2$ into equation (1) to get the second approximation of the root

We have

$$\boldsymbol{x_3 = x_2 - f(x_2)\left(\frac{x_2 - x_1}{f(x_2) - f(x_1)}\right)} \qquad (\mathbf{3})$$

$x_1 = 2.5; x_2 = 2.0976;\ f(x_1) = 4.125;\ f(x_2) = -0.2587$

Note: $f(x_1)\ and\ f(x_2)$ are derived by substituting $x_1 = 2.5,\ x_2 = 2.0976$ into $f(x_1)\ and\ f(x_2)$ respectively.

Substituting the values into equation (3)

$$x_3 = 2.0976 - (-0.2587)\left(\frac{2.0976 - 2.5}{(-0.2587) - 4.125}\right)$$

$$x_3 = 2.0976 + (0.2587)\left(\frac{-0.4025}{-4.3837}\right)$$

$$x_3 = 2.0976 + (0.2587)(0.09182)$$

$$x_3 = 2.0976 + 0.0238$$

$$x_3 = 2.1214$$

Thus,

$$f(x_3) = f(2.1214) = (2.1214)^3 - 5(2.1214) + 1$$

$$f(x_3) = -0.0600$$

Third Iteration:

Hence, substituting $i = 3$ into equation (1) to get the third approximation of the root

We have

$$\boldsymbol{x_4 = x_3 - f(x_3)\left(\frac{x_3 - x_2}{f(x_3) - f(x_2)}\right)} \qquad (\mathbf{4})$$

$x_2 = 2.0976; x_3 = 2.1214;\ f(x_2) = -0.2587;\ f(x_3) = -0.0600$

Note: $f(x_2)$ *and* $f(x_3)$ are derived by substituting $x_2 = 2.0976,\ x_3 = 2.1214$ into $f(x_2)$ *and* $f(x_3)$ respectively.

Substituting the values into equation (4)

$$x_4 = 2.1214 - (-0.0600)\left(\frac{2.1214 - 2.0976}{(-0.0600) - (-0.2587)}\right)$$

$$x_4 = 2.1214 + (0.0600)\left(\frac{0.0238}{0.1987}\right)$$

$$x_4 = 2.1214 + (0.0600)(0.1198)$$

$$x_4 = 2.1214 + 0.0072$$

$$x_4 = 2.1286$$

Thus,

$$f(x_4) = f(2.1286) = (2.1286)^3 - 5(2.1286) + 1$$

$$f(x_4) = 0.0016$$

Fourth Iteration:

Hence, substituting $i = 4$ into equation (1) to get the fourth approximation of the root

We have

$$\boldsymbol{x_5 = x_4 - f(x_4)\left(\frac{x_4 - x_3}{f(x_4) - f(x_3)}\right)} \qquad (\mathbf{5})$$

$x_3 = 2.1214; x_4 = 2.1286;\ f(x_3) = -0.0600;\ f(x_4) = 0.0016$

Note: $f(x_3)$ *and* $f(x_4)$ are derived by substituting $x_3 = 2.1214,\ x_4 = 2.1286$ into $f(x_3)$ *and* $f(x_4)$ respectively.

Substituting the values into equation (5)

$$x_5 = 2.1286 - (0.0016)\left(\frac{2.1286 - 2.1214}{0.0016 - (-0.0600)}\right)$$

$$x_5 = 2.1286 - (0.0016)\left(\frac{0.0072}{0.0616}\right)$$

$$x_5 = 2.1286 - (0.0016)(0.1169)$$

$$x_5 = 2.1286 - 0.0002$$

$$x_5 = 2.1284$$

Thus,

$$f(x_5) = f(2.1284) = (2.1284)^3 - 5(2.1284) + 1$$

$$f(x_5) = -0.0002$$

Summary Table of Iterations

i (iteration)	**X_i**	**$f(x_i)$**	**X_{i+1}**	**$\|X_{i+1} - X_i\|$**
0	2	-1	2.5	0.5
1	2.5	4.125	2.0976	0.4022
2	2.0976	-0.2587	2.1214	0.0238
3	2.1214	-0.0600	2.1286	0.0072
4	2.1286	0.0016	2.1284	0.0002

Therefore, the root of the equation after 4 iterations correct to 4 decimal places is <u>2.1284</u>

Modified Secant Method

In this modified Secant method, only one initial guess is needed :

$$f'(x_i) \approx \frac{f(x_i + \delta x_i) - f(x_i)}{\delta x_i}$$

$$x_{i+1} = x_i - \frac{f(x_i)}{\dfrac{f(x_i + \delta x_i) - f(x_i)}{\delta x_i}} = x_i - \frac{\delta x_i f(x_i)}{f(x_i + \delta x_i) - f(x_i)}$$

Question 5(o)

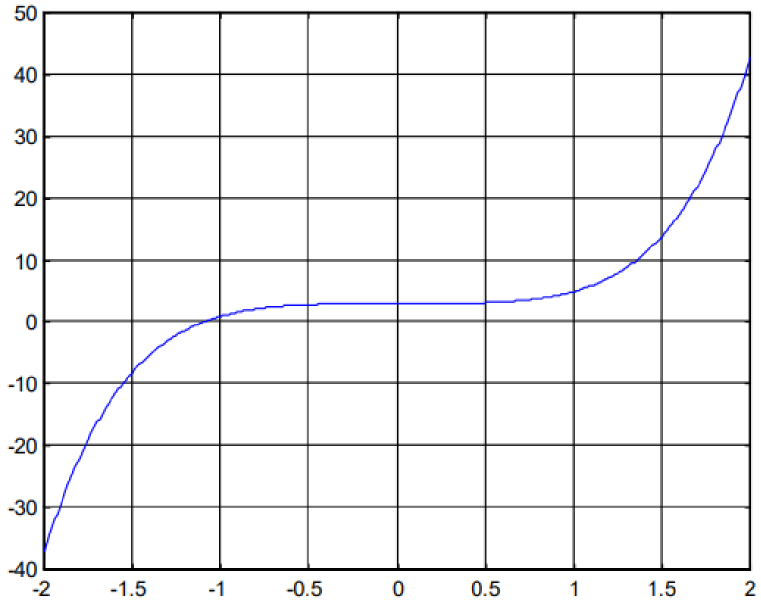
Find the roots of :

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

Initial points

$x_0 = -1$ *and* $x_1 = -1$

with error < 0.001



Fixed-point Iteration Method

- Also known as **one-point iteration** or **successive substitution**

- To find the root for $f(x) = 0$, we reformulate $f(x) = 0$ so that there is an x on one side of the equation.

$$f(x) = 0 \quad \Leftrightarrow \quad g(x) = x$$

- If we can solve $g(x) = x$, we solve $f(x) = 0$.
  - x is known as the fixed point of $g(x)$.
- We solve $g(x) = x$ by computing

$$x_{i+1} = g(x_i) \quad \text{with } x_0 \text{ given}$$

until x_{i+1} converges to x.

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

$$x^2 + 2x - 3 = 0 \quad \Rightarrow \quad 2x = 3 - x^2 \quad \Rightarrow \quad x = \frac{3 - x^2}{2}$$

$$\Rightarrow x_{i+1} = g(x_i) = \frac{3 - x_i^2}{2}$$

Reason: If x converges, i.e. $x_{i+1} \rightarrow x_i$

$$x_{i+1} = \frac{3 - x_i^2}{2} \rightarrow x_i = \frac{3 - x_i^2}{2}$$

$$\Rightarrow x_i^2 + 2x_i - 3 = 0$$

Question 5p

Use fixed point iteration to:

Find root of $f(x) = e^{-x} - x = 0$.
(Answer: $\alpha = 0.56714329$)

We put $x_{i+1} = e^{-x_i}$

i	x_i	ε_a (%)	ε_t (%)
0	0		100.0
1	1.000000	100.0	76.3
2	0.367879	171.8	35.1
3	0.692201	46.9	22.1
4	0.500473	38.3	11.8
5	0.606244	17.4	6.89
6	0.545396	11.2	3.83
7	0.579612	5.90	2.20
8	0.560115	3.48	1.24
9	0.571143	1.93	0.705
10	0.564879	1.11	0.399

- There are infinite ways to construct $g(x)$ from $f(x)$.

For example, $f(x) = x^2 - 2x - 3 = 0$ (ans: $x = 3$ or -1)

Case a:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x^2 = 2x + 3$$
$$\Rightarrow x = \sqrt{2x+3}$$
$$\Rightarrow g(x) = \sqrt{2x+3}$$

Case b:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow x(x-2) - 3 = 0$$
$$\Rightarrow x = \frac{3}{x-2}$$
$$\Rightarrow g(x) = \frac{3}{x-2}$$

Case c:

$$x^2 - 2x - 3 = 0$$
$$\Rightarrow 2x = x^2 - 3$$
$$\Rightarrow x = \frac{x^2 - 3}{2}$$
$$\Rightarrow g(x) = \frac{x^2 - 3}{2}$$

So which one is better?

Case a

$$x_{i+1} = \sqrt{2x_i + 3}$$

1. $x_0 = 4$
2. $x_1 = 3.31662$
3. $x_2 = 3.10375$
4. $x_3 = 3.03439$
5. $x_4 = 3.01144$
6. $x_5 = 3.00381$

Converge!

Case b

$$x_{i+1} = \frac{3}{x_i - 2}$$

1. $x_0 = 4$
2. $x_1 = 1.5$
3. $x_2 = -6$
4. $x_3 = -0.375$
5. $x_4 = -1.263158$
6. $x_5 = -0.919355$
7. $x_6 = -1.02762$
8. $x_7 = -0.990876$
9. $x_8 = -1.00305$

Converge, but slower

Case c

$$x_{i+1} = \frac{x_i^2 - 3}{2}$$

1. $x_0 = 4$
2. $x_1 = 6.5$
3. $x_2 = 19.625$
4. $x_3 = 191.070$

Diverge!

s/n	Root finding method	f(a)f(b)<0 assumption	2 initial point	1 initial point	Class	Derivatives needed	Formula
1	**Bisection**	Yes	Yes		Bracket		x3 = (x1 + x2) / 2
2	**Regua Falsi**	Yes	Yes		Bracket		$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)}$
3	**Newton Raphson**			Yes	Open	Yes	$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$
4	**Secant**		Yes		Open		$x_3 = x_2 - f(x_2)\frac{x_2 - x_1}{f(x_2) - f(x_1)}$
5	**Modified Secant**			Yes	Open	Yes	x1 = x – [f(x) – f ’(x)]
6	**Fixedpoint Iteration**			Yes	Open		$x_{i+1} = g(x_i)$

0 INTRO TO COMPUTATIONAL SCIENCE AND NUMERICAL METHODS

Numerical analysis ---> Numerical methods -----> Computational numerical methods ----> Computational numerical analysis

Numerical Analysis

Numerical Analysis is the study of numerical methods. Numerical analysis finds application in all fields of engineering and the physical sciences, and in the 21st century also the life and social sciences, medicine, business and even the arts. The GOAL of numerical analysis is the design and analysis of techniques/METHODS to give approximate but accurate solutions to hard problems.

Numerical methods

Numerical methods are mathematical attempts at finding approximate solutions of problems rather than the exact ones.

Computational numerical methods

Before modern computers, numerical methods often relied on hand formulas, using data from large printed tables. Since the mid20th century, computers calculate the required functions instead, but many of the same formulas continue to be used in software algorithms.

Computational numerical analysis

Current growth in computing power has enabled the use of more complex numerical analysis, providing detailed and realistic mathematical models in science and engineering. Numerical analysis continues this long tradition: rather than giving exact symbolic answers translated into digits and applicable only to real-world measurements, approximate solutions within specified error bounds are used.

1 Methods of Approximations

- Rounding off to significant figures
- Rounding off to decimal places
    - Working with arithmetic precision

2 Methods of Errors

    - Sources of errors
- Rounding errors
- Inherent errors
- Truncation errors
- True errors
- Relative true errors
- Absolute errors
- Relative absolute errors
- Approximate errors
- Relative approximate errors
- Absolute relative errors
- Percentage errors
    - Propagation of errors

3 Methods of Drawing the Lines of best fit

- Linearization
- Least squares curve fitting
- Group averages grouped averages curve fitting

4 Methods of solving Linear Systems of Equations

- Gaussian elimination
- Gauss-Jordan elimination
- LU decomposition
- Jacobi iteration
- Gauss-Seidal iteration

5 Methods of finding the roots of Algebraic and Transcendental Equations

- Bisection
- Newton-Raphson
- Regula-falsi
- Secant

6 Methods of Finite Differences

- First forward difference
- First backward difference
- First central difference

7 Methods of Interpolation

- Newton's forward interpolation formula

- Newton’s backward interpolation formula

8 Methods of Numerical Integration

- Newton-coates Quadrature
- Trapezoidal rule
- Simpson’s one-third rule
- Simpson’s three-eight rule
- Romberg’s method

9 Methods of Solving First Order Ordinary Differential Equations

- Picard’s Method
- Euler Method
- Modified Euler Method
- Runge-Kutta first order method
- Runge-Kutta second order method
- Runge-Kutta third order method
- Runge-Kutta fourth order method

10 Methods of describing grouped and ungrouped statistical data

- Measures of central tendency (mean, median, mode)
- Measures of position (quartiles, percentiles, deciles)
- Measures of dispersion (range, interquartile range, standard deviation, variance)

1 Methods of Approximations

Question 1

Without calling any in-built library or function, write a new function from scratch to

a. round-off any number to a stated precision of decimal place
b. return the absolute value of any number
c. find the natural log of any number
d. approximate any number to its nearest_integer
e. take the magnitude of any number
f. round off any number to a stated amount of significant figures
g. re-write ‘f’ using built-in libraries and functions.

Solution to Question 1f

ALGORITHM - To round off numbers to certain amount of significant figures

```
Given any ‘number’, with the ‘num_sig_figs’ to approximate the number to
If ‘number’ == 0,
        Return 0.0
Else
        Take the absolute value of ‘number’
        Take the natural log of the absolute value
        Take ‘number’ to its nearest_integer
        Take the magnitude of the natural log
        Calculate rounding_factor = 10** (num_sig_figs – magnitude - 1)
        Find rounded_number = nearest_integer / rounding_factor
        If ‘number’ > 0
                Return rounded_number
        If ‘number’ < 0
                Return –rounded_number
```

PSEUDOCODE - To round off numbers to certain amount of significant figures

Without calling any in-built library or function, this is a new function from scratch to round off numbers to certain amount of significant figures

```
def round_to_significant_figures(number, num_sig_figs):
    if number == 0:
        return 0.0

    # Calculate the absolute value of 'number'
    def absolute_value(number):
        if number < 0:
            return -number
        else:
            return number
    abs_number = absolute_value(number)

    # Calculate the natural logarithm of the absolute value
    def custom_ln(number, num_terms=100):
        if number == 1:
            return 0.0
        elif number < 1:
            number = 1 / number
            num_terms = -num_terms
        nat_log = 0.0
        for n in range(1, num_terms + 1):
            term = ((number - 1) ** n) / n
            if n % 2 == 0:
                nat_log -= term
            else:
                nat_log += term
        return nat_log

    # Calculate the magnitude of the natural log
```

```
    def custom_floor(nat_log):
        if nat_log >= 0:
            return int(nat_log)
        else:
            integer_part = int(nat_log)
            if integer_part == nat_log:
                return integer_part
            else:
                return integer_part - 1
    magnitude = custom_floor(custom_ln(abs_number))

    # Calculate rounding_factor
    rounding_factor = 10 ** (num_sig_figs - magnitude - 1)

    # Use rounding factor to round number to the specified significant figures
    def custom_roundoff(number):
        decimal_part = number - int(number)
        if decimal_part < 0.5:
            return int(number)
        else:
            return int(number) + 1
    rounded_number = custom_roundoff(abs_number * rounding_factor) / rounding_factor

    # Restore the sign
    if number > 0:
        return rounded_number
    else:
        return -rounded_number

#round_to_significant_figures(number, num_sig_figs)
```

Arithmetic precision

It might be necessary to round off numbers to make them useful for numerical computation, moreso as it would require an infinite computer memory to store an unending number.

The precision of a number is an indication of the number of digits that have been used to express it. In scientific computing, it is the number of significant digits or numbers, while in management and financial systems, it is the number of decimal places. We are quite aware that most currencies in the world are quoted to two decimal places.

Arithmetic precision (often referred to simply as precision) is the specified number of significant figures or digits to which the number of interest is to be rounded.

2 Methods of errors

Rounding Errors

These are errors incurred by truncating a sequence of digits representing a number, as we saw in the case of representing the rational number 3/7 by 2.3333, instead of 2.3333….., which is an unending number. Apart from being unable to write this number in an exact form by hand, our instruments of calculation, be it the calculator or the computer, can only handle a finite string of digits. Rounding errors can be reduced if we change the calculation procedure in such a way as to avoid the subtraction of nearly equal numbers or division by a small number. It can also be reduced by retaining at least one more significant figure at each step than the one given in the data, and then rounding off at the last step.

Inherent Errors

As the name implies, these are errors that are inherent in the statement of the problem itself. This could be due to the limitations of the means of calculation, for instance, the calculator or the computer. This error could be reduced by using a higher precision of calculation.

Truncation Errors

If we truncate Taylor's series, which should be an infinite series, then some error is incurred. This is the error associated with truncating a sequence or by terminating an iterative process. This kind of error also results when, for instance, we carry out numerical differentiation or integration, because we are replacing an infinitesimal process with a finite one. In either case, we would have required that the elemental value of the independent variable tend to zero in order to get the exact value.

Absolute Error, Relative True Error, Relative Approximate Error and Percentage Error etc.

Question 2

a. A student measured the length of a string of actual length 72.5 cm as 72.4 cm.
   i. Calculate the absolute error and the percentage error
   ii. Write a function that accepts measured length and actual length to output absolute error and the percentage error.

b. The derivative of a function f(x) can be approximated by the equation

$$f'(x) \approx \frac{f(x+h)-f(x)}{h}, \qquad \text{if } f(x) = 7e^{0.5x}, \text{ and } h = 0.3,$$

   i. Find the true value, the approximate value, true error, and relative error of f'(2)
   ii. If true values are not known or are very difficult to obtain, then Approximate error (E_a) = Present Approximation – Previous approximation. For $f(x) = 7e^{0.5x}$ at x = 2, find
      - f'(2) using h = 0.3
      - f'(2) using h = 0.15

- approximate error, and relative approximate error of f’(2)

iii. Write a function that takes in any value of x for the derivative of a function f(x) approximated by the equation f'(x) = [f(x + h) - f(x)] / h for f(x) = 7e^0.5x, h1 = 0.3, h2 = 0.15, and returns true value, approximate value, true error, relative error, approximate error, and relative approximate error of f’(x)

Solution to Question 2a

(2ai)

Absolute error = | actual value – measured value |
Relative absolute error = | actual value – measured value | / actual value
The percentage error = Relative absolute error x 100
Absolute error = | 72.5 – 72.4 |= 0.01.
The percentage error = (0.1 / 72.5) x 100 = 0.1379

(2aii)

```
def calculate_errors(actual_length, measured_length):
    absolute_error = abs(actual_length - measured_length)
    relative_error = absolute_error / actual_length
    percentage_error = relative_error * 100

    return {
        "Absolute Error": absolute_error,
        "Relative Absolute Error": relative_error,
        "Percentage Error": percentage_error
    }

calculate_errors(actual_length, measured_length)
```

Solution to Question 2b

(2bi)
Approximate value of f'(x), for x = 2, and h = 0.3

$$f'(2) \approx \frac{f(2+0.3)-f(2)}{0.3}$$

$$= \frac{f(2.3)-f(2)}{0.3}$$

$$= \frac{7e^{0.5(2.3)} - 7e^{0.5(2)}}{0.3}$$

$$= \frac{22.107-19.028}{0.3} \quad = 10.263$$

(2bi cont'd)
The exact or true value of f'(2) can be found by using our knowledge of differential calculus

$$f(x) = 7e^{0.5x}$$

$$f'(x) = 7 \times 0.5 \times e^{0.5x}$$

$$= 3.5e^{0.5x}$$

$$f'(2) = 3.5e^{0.5(2)}$$

$$= 9.5140$$

(2bi cont'd)

True Error = True Value – Approximate Value
E_t = 9.5140 – 10263 = -0.722

Relative true error = (True value – Approximate value) / True value

= (9.5140 – 10263) / 9.5140= -0.722 / 9.5140

(2bii)

For x = 2, and h = 0.3

- Approximate value of f'(x) = 10.263

For x = 2, and h = 0.15

- Approximate value of f'(x) =

$$f^{'}(2) \approx \frac{f(2+0.15)-f(2)}{0.15}$$

$$= \frac{f(2.15)-f(2)}{0.15}$$

$$= \frac{7e^{0.5(2.15)}-7e^{0.5(2)}}{0.15}$$

$$= \frac{20.50-19.028}{0.15} = 9.8800$$

Approximate error (E_a) = Present Approximation – Previous approximation

= 9.8800 – 10.263

= -0.38300

Relative approximate error = Approximate error / Previous approximation

$$= \frac{-0.38300}{9.8800} = -0.038765$$

(Question 2biii)

```
import math
def f(x):
    return 7 * math.exp(0.5 * x)
def derivative_of_f(x):
    return 3.5 * math.exp(0.5 * x)

def calculate_derivative_error(x, first_h, f, second_h=None):
    true_value = derivative_of_f(x)
    first_approximate_value = (f(x + first_h) - f(x)) / first_h
    true_error = abs(true_value - first_approximate_value)
    relative_true_error = true_error / true_value
    if second_h is not None:
        second_approximate_value = (f(x + second_h) - f(x)) / second_h
        approximate_error = abs(second_approximate_value - first_approximate_value)
        relative_approximate_error = approximate_error / second_approximate_value
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "Second Approximate Value": second_approximate_value,
            "Approximate Error": approximate_error,
            "Relative Approximate Error": relative_approximate_error
        }
    else:
        return {
            "True Value": true_value,
            "First Approximate Value": first_approximate_value,
            "True Error": true_error,
            "Relative True Error": relative_true_error
        }
#calculate_derivative_error(x, first_h, f)
#calculate_derivative_error(x, first_h, f, second_h)
```

Propagation of errors

In numerical methods, the calculations are not made with exact numbers. How do these inaccuracies propagate through the calculations?

Question 2c

Find the bounds for the propagation in adding two numbers. For example if one is calculating $X + Y$ where

$$X = 1.5 \pm 0.05$$

$$Y = 3.4 \pm 0.04$$

Solution

Maximum possible value of $X = 1.55$

Maximum possible value of $Y = 3.44$

Maximum possible value of $X + Y = 1.55 + 3.44 = 4.99$

Minimum possible value of $X = 1.45$.

Minimum possible value of $Y = 3.36$.

Minimum possible value of $X + Y = 1.45 + 3.36 = 4.81$

Hence

$$4.81 \leq X + Y \leq 4.99.$$

Propagation of Errors In Formula

If f is a function of several variables $X_1, X_2, X_3, \ldots\ldots, X_{n-1}, X_n$

then the maximum possible value of the error in f is

$$\Delta f \approx \left|\frac{\partial f}{\partial X_1}\Delta X_1\right| + \left|\frac{\partial f}{\partial X_2}\Delta X_2\right| + \ldots\ldots + \left|\frac{\partial f}{\partial X_{n-1}}\Delta X_{n-1}\right| + \left|\frac{\partial f}{\partial X_n}\Delta X_n\right|$$

Question 2d

The strain in an axial member of a square cross-section is given by

$$\in = \frac{F}{h^2 E}$$

Given F = 72

h = 4 x 10^{-3}

E = 70 x 10^{9}

Find the maximum possible error in the measured strain.

Solution

$$\in = \frac{72}{(4\times10^{-3})^2(70\times10^9)}$$
$$= 64.286\times10^{-6}$$
$$= 64.286\mu$$

$$\Delta\in = \left|\frac{\partial\in}{\partial F}\Delta F\right| + \left|\frac{\partial\in}{\partial h}\Delta h\right| + \left|\frac{\partial\in}{\partial E}\Delta E\right|$$

$$\frac{\partial\in}{\partial F} = \frac{1}{h^2E} \qquad \frac{\partial\in}{\partial h} = -\frac{2F}{h^3E} \qquad \frac{\partial\in}{\partial E} = -\frac{F}{h^2E^2}$$

Thus

$$\Delta E = \left|\frac{1}{h^2E}\Delta F\right| + \left|\frac{2F}{h^3E}\Delta h\right| + \left|\frac{F}{h^2E^2}\Delta E\right|$$
$$= \left|\frac{1}{(4\times10^{-3})^2(70\times10^9)}\times0.9\right| + \left|\frac{2\times72}{(4\times10^{-3})^3(70\times10^9)}\times0.0001\right|$$
$$+ \left|\frac{72}{(4\times10^{-3})^2(70\times10^9)^2}\times1.5\times10^9\right|$$

Hence

$\in = (64.286\mu \pm 5.3955\mu)$

Question 2e

Subtraction of numbers that are nearly equal can create unwanted inaccuracies. Using the formula for error propagation, show that this is true.

Solution

Let z = x – y, Then

$$|\Delta z| = \left|\frac{\partial z}{\partial x}\Delta x\right| + \left|\frac{\partial z}{\partial y}\Delta y\right|$$
$$= |(1)\Delta x| + |(-1)\Delta y|$$
$$= |\Delta x| + |\Delta y|$$

So the relative error or relative change is

$$\left|\frac{\Delta z}{z}\right| = \frac{|\Delta x| + |\Delta y|}{|x-y|}$$

Check: $x = 2 \pm 0.001$

$$y = 2.003 \pm 0.001$$
$$\left|\frac{\Delta z}{z}\right| = \frac{|0.001| + |0.001|}{|2-2.003|}$$
$$= 0.667$$

Percentage error = 66.67%

Taylor series

Some examples of common Taylor series

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

The general form of Taylor series is given as

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \frac{f'''(x)}{3!}h^3 + \cdots$$

provided that all derivatives of f(x) are continuous and exist in the interval [x,x+h]

What does this mean in plain English?

As Archimedes would have said, "*Give me the value of the function at a single point, and the (first, second, and so on) values of all its derivatives at that single point, and I can give you the value of the function at any other point*"

Question 2f

Find the value of f(6) given that f(4) =125, f'(4) = 74, f''(4) = 30, f'''(4) = 6 and all other higher order derivatives of f(x) at x=4 are zero.

Solution

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2!} + f'''(x)\frac{h^3}{3!} + \cdots$$

x = 4

h = 6 – 4 = 2

Since the higher order derivatives are zero,

$$f(4+2) = f(4) + f'(4)2 + f''(4)\frac{2^2}{2!} + f'''(4)\frac{2^3}{3!}$$

$$f(6) = 125 + 74(2) + 30\left(\frac{2^2}{2!}\right) + 6\left(\frac{2^3}{3!}\right)$$

= 125 + 148 + 60 + 8

= 341

Note that to find f(6) exactly, we only need the value of the function and all its derivatives at some other point, in this case x = 4.

Error in Taylor series

The Taylor polynomial of order n of a function f(x) with (n+1) continuous derivatives in the domain [x,x+h] is given by

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2!}+\cdots+f^{(n)}(x)\frac{h^n}{n!}+R_n(x)$$

where the remainder is given by

$$R_n(x)=\frac{(x-h)^{n+1}}{(n+1)!}f^{(n+1)}(c)$$

Such that

$$x<c<x+h$$

that is, c is some point in the domain [x,x+h]

Derivation for Maclaurin Series for e^x

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots$$

The Maclaurin series is simply the Taylor series about the point x=0

$$f(x+h)=f(x)+f'(x)h+f''(x)\frac{h^2}{2!}+f'''(x)\frac{h^3}{3!}+f''''(x)\frac{h^4}{4}+f'''''(x)\frac{h^5}{5}+\cdots$$

$$f(0+h)=f(0)+f'(0)h+f''(0)\frac{h^2}{2!}+f'''(0)\frac{h^3}{3!}+f''''(0)\frac{h^4}{4}+f'''''(0)\frac{h^5}{5}+\cdots$$

Since f(x) = e^x, f'(x) = e^x, f''(x) = e^x, $f^n(x)$ = e^x, and $f^n(0)$ = e^0 = 1;

The Maclaurin series is then:

$$f(h)=(e^0)+(e^0)h+\frac{(e^0)}{2!}h^2+\frac{(e^0)}{3!}h^3\ldots$$

$$=1+h+\frac{1}{2!}h^2+\frac{1}{3!}h^3\ldots$$

Therefore

$$f(x)=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots$$

It can be seen that as the number of terms used increases, the error bound decreases and hence a better estimate of the function can be found.

Question 2g

How many terms would it require to get an approximation of e^1 within a magnitude of true error of less than $10^{-6.}$

Solution

Using (n + 1) terms of Taylor seriesgives error bound of

$$R_n(x) = \frac{(x-h)^{n+1}}{(n+1)!} f^{(n+1)}(c) \qquad x = 0, h = 1, f(x) = e^x$$

$$R_n(0) = \frac{(0-1)^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

$$= \frac{(-1)^{n+1}}{(n+1)!} e^c$$

Since

x < c < x+h

0 < c < 0+1

0 < c < 1

$$\frac{1}{(n+1)!} < |R_n(0)| < \frac{e}{(n+1)!}$$

So if we want to find out how many terms it would require to get an approximation of e^1 magnitude of true error of less than 10^6,

$$\frac{e}{(n+1)!} < 10^{-6}$$

$$(n+1)! > 10^6 e$$

$$(n+1)! > 10^6 \times 3$$

$$n \geq 9$$

So 9 terms or more are needed to get a true error less than 10^{-6}.

3 Drawing line of best fit

The process of fitting a curve to a set of data is called curve-fitting.

Linearisation

A nonlinear relationship can be linearised and the resulting graph analysed to bring out the relationship between variables.

y = ix + j ----------> linear or straight line graph, i=slope, j=intercept

$y = ix^2 + jx + k$ --------> quadratic graph or curve

$y = ix^n + jx + k$: n >= 3 --------> polynomial or sinusoidal wave form graph

$y = ie^x$ ---------------> ?? non-linear graph

$y = 2\log_x i3$ --------> ?? non-linear graph

Remember:

ln(x) is the natural logarithm to the base 'e' ≈ 2.71828, often referred to simply as "log."

$\log_{10}(x)$ is the common logarithm to the base 10, often referred to simply as "log."

In mathematical notation, the distinction is clear:

$\ln(x) = \log_e(x)$, where 'e' is the base of the natural logarithm.

$\log(x) = \log_{10}(x)$ where 10 is the base of common logarithm.

Case 1: $y = ae^x$.

(i) We could take the logarithm of both sides to base e:

$$\ln y = \ln(ae^x) = \ln a + \ln e^x = x + \ln a,$$

since $\ln e^x = x$. Thus, a plot of $\ln y$ against x gives a linear graph with slope unity and a y-intercept of ln a.

(ii) We could also have plotted y against e^x. The result is a linear graph through the origin, with slope equal to a.

Case 2: $T = 2\pi\sqrt{\frac{l}{g}}$

We can write this expression in three different ways:

(i) $\ln T = \ln(2\pi) + \frac{1}{2}\ln\left(\frac{l}{g}\right) = \ln(2\pi) + \frac{1}{2}(\ln l - \ln g)$.

Rearranging, we obtain,

$$\ln T = \frac{1}{2}\ln l + \left(\ln(2\pi) - \frac{1}{2}\ln g\right)$$

writing this in the form $y = mx + c$, we see that a plot of $\ln T$ against $\ln l$ gives a slope of 0.5 and a $\ln T$ intercept of $\left(\ln(2\pi) - \frac{1}{2}\ln g\right)$. Once the intercept is read of the graph, you can then calculate the value of g.

(ii) $T = \frac{2\pi}{\sqrt{g}}\sqrt{l}$

A plot of T versus $\sqrt{l}$ gives a linear graph through the origin (as the intercept is zero). The slope of the graph is $\frac{2\pi}{\sqrt{g}}$, from which the value of g can be recovered.

Case 3: $N = N_0 e^{-\lambda t}$

The student can show that a plot of $\ln N$ versus t will give a linear graph with slope $-\lambda$, and $\ln N$ intercept is $\ln N_0$.

What other functions of N and t could you plot in order to get λ and N_0?

Case 4: $\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$

We rearrange the equation:

$$\frac{1}{v} = \frac{1}{f} - \frac{1}{u}$$

A plot of v^{-1} (y-axis) versus u^{-1} (x-axis) gives a slope of -1 and a vertical intercept of $\frac{1}{f}$.

Question 3a

A student obtained the following reading with a mirror in the laboratory.

u	10	20	30	40	50
v	-7	-10	-14	-15	-17

Linearise the relationship $\frac{1}{v}=\frac{1}{f}-\frac{1}{u}$. Plot the graph of v^{-1} versus u^{-1} and draw the line of best fit. Hence, find the focal length of the mirror. All distances are in cm.

Solution

u	v	1/u	1/v
10	-7	0.1	-0.14286
20	-10	0.05	-0.1
30	-14	0.033333	-0.07143
40	-15	0.025	-0.06667
50	-17	0.02	-0.05882

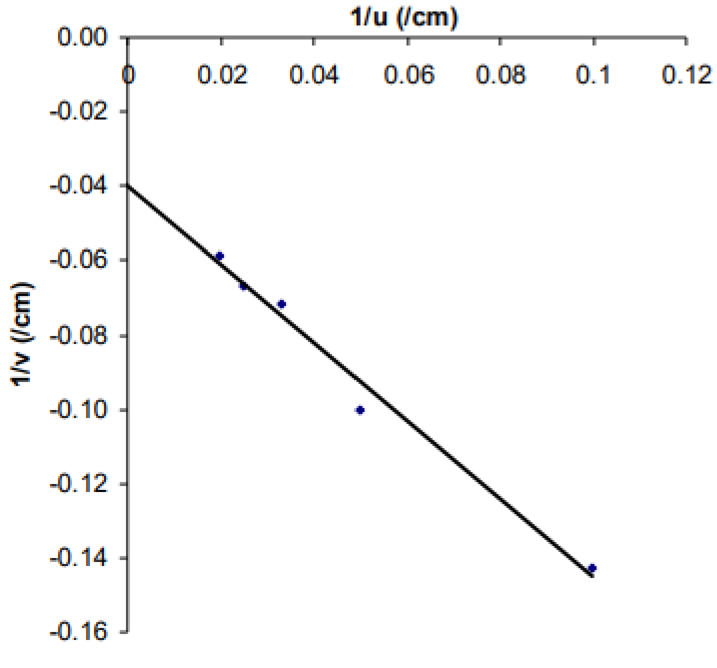


Fig. 1.1: Linear graph of the function $\frac{1}{v}=\frac{1}{f}-\frac{1}{u}$

The slope is -1.05 and the intercept -0.04. From $\frac{1}{v}=\frac{1}{f}-\frac{1}{u}$, we see that the intercept is $\frac{1}{f}=-0.04$, or $f=-\frac{1}{0.04}=-25$ cm.

Method of least squares curve fitting

The least square method entails minimizing the sum of the squares of the difference between the measured value and the one predicted by the assumed equation.

$$m = \frac{\overline{xy} - \bar{x}\,\bar{y}}{\overline{x^2} - \bar{x}^2}$$

$$c = \bar{y} - m\bar{x}$$

Question 3b

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

Thus, for the following set of readings:

t	5	12	19	26	33
x	23	28	32	38	41

The table can be extended to give

t	5	12	19	26	33	Σ=95	$\bar{t}$=19
x	23	28	32	38	41	Σ=162	$\bar{x}$=32.4
tx	115	336	608	988	1353	Σ=3400	$\overline{tx}$=680
t^2	25	144	361	676	1089	Σ=2295	$\overline{x^2}$=459

$$m = \frac{\overline{tx} - \bar{t}\,\bar{x}}{\overline{t^2} - \bar{t}^2} = \frac{680 - 19 \times 32.4}{459 - 19^2} = 0.6571$$
$$c = \bar{x} - m\bar{t} = 32.4 - 0.6571 \times 19 = 19.9151$$

Hence, the relationship between x and t is,

$$x = 0.6571t + 19.9151$$

Method of group averages curve fitting

As the name implies, a set of data is divided into two groups, each of which is assumed to have a zero sum of residuals.

$$\bar{y}_1 = m\bar{x}_1 + c$$
$$\bar{y}_2 = m\bar{x}_2 + c$$

Subtracting,

$$\bar{y}_1 - \bar{y}_2 = m(\bar{x}_1 - \bar{x}_2)$$
$$m = \frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1 - \bar{x}_2}$$

and

$$c = \bar{y}_1 - m\bar{x}_1$$

Question 3c

A student obtained the following data in the laboratory. By making use of the method of least squares, find the relationship between x and t.

t	5	12	19	26	33
x	23	28	32	38	41

Solution:

First divide the data into 2 groups

t	5	12	19
x	23	28	32

and

t	26	33
x	38	41

The tables can be extended to give, for Table 3:

t	5	12	19	Σ=36	$\bar{t}_1$=12
x	23	28	32	Σ=83	$\bar{x}_1$=27.666667

and for Table 4:

t	26	33	Σ=59	$\bar{t}_2$=29.5
x	38	41	Σ=79	$\bar{x}_2$=39.5

$$m = \frac{\bar{x}_1 - \bar{x}_2}{\bar{t}_1 - \bar{t}_2} = \frac{27.666667 - 39.5}{12 - 29.5} = 0.67619$$

and

$$c = \bar{x}_1 - m\bar{t}_1 = 27.666667 - (0.67619 \times 12)$$
$$= 19.552387$$

Thus, the equation of best fit is,

$$x = 0.67619t + 19.552387$$

4 Methods of Linear Systems of Equation

Let us consider a linear system of equations

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

This can be written in the form

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

Gaussian Elimination involving 2 variables

Question 4a: find x and y given that

2x + 3y = 13

x – y = -1

The augmented matrix representing our system of two equations is

$$\left[\begin{array}{cc|c} 2 & 3 & 13 \\ 1 & -1 & -1 \end{array}\right]$$

By Gaussian elimination, we seek to make every entry below the main diagonal zero. This we achieve by reducing 1 to zero, making use of the first row.

Thus,

$$5y = 15 \Rightarrow y = 3$$

Substituting this in the first row gives

$$2x + 3(3) = 13$$

from which we obtain $x = 2$.

The process of reducing every element below the main diagonal to zero (row echelon form) is called Gaussian Elimination. That of substituting obtained values to calculate other variables is called Back Substitution.

The same process can be carried over to the case of a system of three equations.

Question 4b:

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

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

This yields (by Gaussian elimination)

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right] \xrightarrow[(iii)=(i)-(2/3)(iii)]{(ii)'=(i)-2(ii)} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 7/3 & 5/3 & 3 \end{array}\right]$$

$$\xrightarrow{(iii)''=(ii)'+(15/7)(iii)'} \left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

Upon back substitution,

$$-10z = 10 \text{ or } z = -1$$

$$z = -1;\ y + z = 1 \Rightarrow y = 2;\ 2x + y - z = 5 \Rightarrow x = 1$$

Traditionally, in Mathematics, it is usual to use indices such as x_1, x_2, etc. instead of x, y, z. Do you have any idea why this is so? It is because if we stay with the alphabets, we shall soon run out of symbols. Bear in mind that not all the alphabets can be employed as variables; as an example, a, b, c are commonly used as constants. In addition, it makes it easy to associate the coefficients a_{11}, a_{12}, etc. with x_1, x_2, etc. respectively. More importantly in numerical work, it makes programming easier. For instance for our system of three equations, we could use the more general notation:

The general Gaussian elimination for linear system of 3 variables is thus given as:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right] \xrightarrow[(iii)=a_{11}(i)-a_{11}(iii)]{(ii)'=a_{12}(i)-a_{11}(ii)} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & a_{32}' & a_{33}' & a_{34}' \end{array}\right]$$

$$\xrightarrow{(iii)''=a_{32}(ii)'+a_{22}(iii)'} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22}' & a_{23}' & a_{24}' \\ 0 & 0 & a_{33}'' & a_{34}'' \end{array}\right]$$

Gauss-Jordan Elimination

This entails eliminating in addition to the entries below the major diagonal, the entries above it, so that the main matrix is a diagonal matrix. In that case, the solution to the system is given by dividing the element in the augmented part of the matrix by the diagonal element for that row.

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

The augmented matrix is

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 1 & 3 & 2 & 5 \\ 3 & -2 & -4 & 3 \end{array}\right]$$

In solving the same problem using Gauss-Jordan elimination, we continue from completion of the Gaussian elimination part.

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 5 \\ 0 & -5 & -5 & -5 \\ 0 & 0 & -10 & 10 \end{array}\right] \xrightarrow[(ii)'=(iii)-2(ii)]{(i)'=(iii)-10(i)} \left[\begin{array}{ccc|c} -20 & 10 & 0 & -40 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

$$\xrightarrow{(i)''=(ii)'+(i)'} \left[\begin{array}{ccc|c} -20 & 0 & 0 & -20 \\ 0 & -10 & 0 & 20 \\ 0 & 0 & -10 & 10 \end{array}\right]$$

It follows that,

$-20x = -20$ or $x = 1$; $-10y = 20$ or $y = 2$; and $-10z = 10$ or $z = -1$

(lower and upper echelon) - LU decomposition

Suppose we could write the matrix

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

This implies that

$$l_{11}u_{11} = a_{11},\ l_{11}u_{12} = a_{12},\ l_{11}u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11},\ a_{22} = l_{21}u_{12} + l_{22}u_{22},\ a_{23} = l_{21}u_{13} + l_{22}u_{23}$$
$$a_{31} = l_{31}u_{11},\ a_{32} = l_{31}u_{12} + l_{32}u_{22},\ a_{33} = l_{31}u_{13} + l_{32}u_{23} + l_{33}u_{33}$$

Without loss of generality, we could set the diagonal elements of the L matrix equal to 1.

Then,

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

Multiplying out the right side of equation 3.19,

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

From the equality of matrices, this requires that,

$$u_{11} = a_{11}$$

$$u_{12} = a_{12}$$
$$u_{13} = a_{13}$$
$$a_{21} = l_{21}u_{11} \qquad \Rightarrow l_{21} = a_{21}/u_{11} = a_{21}/a_{11}$$
$$a_{31} = l_{31}u_{11} \qquad \Rightarrow l_{31} = a_{31}/u_{11} = a_{31}/a_{11}$$
$$a_{22} = l_{21}u_{12} + u_{22} \text{ , or } u_{22} = a_{22} - l_{21}u_{12} = a_{22} - \frac{a_{21}}{u_{11}}u_{12}$$
$$\Rightarrow u_{22} = a_{22} - \frac{a_{21}}{a_{11}}a_{12}$$
$$a_{23} = l_{21}u_{13} + u_{23} \text{, or } u_{23} = a_{23} - l_{21}u_{13} = a_{23} - \frac{a_{21}}{u_{11}}u_{13}$$
$$\Rightarrow u_{23} = a_{23} - \frac{a_{21}}{a_{11}}a_{13}$$
$$l_{32} = \frac{a_{32} - l_{31}u_{12}}{u_{22}} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{u_{11}}u_{12}\right]$$
$$\Rightarrow l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}}a_{12}\right]$$
$$a_{32} = l_{31}u_{12} + l_{32}u_{22}$$
$$a_{33} = l_{31}u_{13} + l_{32}u_{23} + u_{33}$$
$$\Rightarrow u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23}$$

You can see that we have determined all the nine elements of the two matrices in terms of the elements of the original matrix.

Once we have obtained L and U, then we can write the original equation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

as

$$LU\mathbf{x} = \mathbf{y}$$

where $\mathbf{x}$ and $\mathbf{y}$ are column vectors.

We shall write $\mathbf{w} = U\mathbf{x}$

Then,

$$L\mathbf{w} = \mathbf{y}$$

Now we continue to solving **Question5b** again using LU decomposition

The corresponding matrix is

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

$$u_{11} = a_{11} = 2$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = -1$$
$$l_{21} = a_{21} / a_{11} = 1/2$$
$$l_{31} = a_{31} / a_{11} = 3/2$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 3 - \frac{1}{2}(1) = 5/2$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - \frac{1}{2}(-1) = 2 + \frac{1}{2} = 5/2$$
$$l_{32} = \frac{1}{u_{22}}\left[a_{32} - \frac{a_{31}}{a_{11}} a_{12}\right] = \frac{1}{5/2}\left[-2 - \frac{3}{2}(1)\right] = -7/5$$
$$u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = -4 - (3/2)(-1) - (-7/5)(5/2) = 1$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 5/2 & 5/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 2 \\ 3 & -2 & -4 \end{bmatrix}$$

The above decomposition is correct as the multiplication of L and U gives the original matrix.

The original equation is equivalent to

$$LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$$

$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ 3/2 & -7/5 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 3 \end{bmatrix}$$

Solving,

$$w_1 = 5$$

$$\frac{1}{2}w_1 + w_2 = 5 \text{ or } w_2 = 5 - \frac{1}{2}w_1 = 5 - \frac{1}{2}(5) = \frac{5}{2}$$

$$\frac{3}{2}w_1 - \frac{7}{5}w_2 + w_3 = 3, \text{ or } w_3 = 3 + \frac{7}{5}w_2 - \frac{3}{2}w_1 = 3 + \frac{7}{5}\left(\frac{5}{2}\right) - \frac{3}{2}(5) = -1$$

$U\mathbf{x} = \mathbf{w}$ implies:

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

By back substitution,

$$x_3 = -1$$

$$\frac{5}{2}x_2 + \frac{5}{2}x_3 = \frac{5}{2} \Rightarrow \frac{5}{2}x_2 = \frac{5}{2} - \frac{5}{2}x_3 = \frac{5}{2} - \frac{5}{2}(-1) = 5$$

$$x_2 = 2$$

$$2x_1 + x_2 - x_3 = 5$$

$$x_1 = \frac{5 - x_2 + x_3}{2} = \frac{5 - 2 + (-1)}{2} = 1$$

The solution set is therefore,

$$x_1 = 1,\ y = 2,\ z = -1.$$

Question 4c

Solve the system of linear equations $x+y+z=-1$, $x+2y+2z=-4$, $9x+6y+z=7$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

(i) Gaussian elimination

Initial augmented matrix

1	1	1	-1
1	2	2	-4
9	6	1	7

First round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	-3	-8	16

Second round of Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

Last matrix for Gaussian elimination

1	1	1	-1
0	1	1	-3
0	0	-5	7

First round of Jordan elimination

5	5	0	2
0	5	0	-8
0	0	-5	7

Second round of Jordan elimination

-25	0	0	-50
0	5	0	-8
0	0	-5	7

(iii) LU decomposition

$x+y+z=-1$

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

$9x+6y+z=7$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1}\left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31}u_{13} - l_{32}u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7 \text{, or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$x_3 = -7/5 = -1.4$

$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$

$x_2 = -8/5 = -1.6$

$x_1 + x_2 + x_3 = -1$

$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \dfrac{10}{5} = 2$

The solution set is therefore,

$x_1 = 2$, $y = -1.6$, $z = -1.4$.

Question 4d

Solve the system of linear equations $x + 2y + 2z = -2$, $2x + 2y + z = -4$, $9x + 6y + 2z = -14$ using the method of

(i) Gaussian elimination

(ii) Gauss-Jordan elimination

(iii) LU decomposition

Gaussian elimination

Initial augmented matrix

1	2	2	-2
2	2	1	-4
9	6	2	-14

First round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	-12	-16	4

Second round of Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

Answers

x	0
y	-3
z	2

Gauss-Jordan elimination

Last matrix for Gaussian elimination

1	2	2	-2
0	-2	-3	0
0	0	-4	-8

First round of Jordan elimination

4	8	0	-24
0	-8	0	24
0	0	-4	-8

Second round of Jordan elimination

32	0	0	0
0	-8	0	24
0	0	-4	-8

LU decomposition

$$x + 2y + 2z = -4$$
$$9x + 6y + z = 7$$

The corresponding matrix is

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

$$u_{11} = a_{11} = 1$$
$$u_{12} = a_{12} = 1$$
$$u_{13} = a_{13} = 1$$
$$l_{21} = a_{21} / a_{11} = 1/1 \quad = 1$$
$$l_{31} = a_{31} / a_{11} = 9/1 \quad = 9$$
$$u_{22} = a_{22} - \frac{a_{21}}{a_{11}} a_{12} = 2 - (1)(1) = 1$$
$$u_{23} = a_{23} - \frac{a_{21}}{a_{11}} a_{13} = 2 - (1)(1) = 1$$
$$l_{32} = \frac{1}{u_{22}} \left[a_{32} - \frac{a_{31}}{a_{11}} a_{12} \right] = \frac{1}{1} \left[6 - \frac{9}{1}(1) \right] = -3$$
$$u_{33} = a_{33} - l_{31} u_{13} - l_{32} u_{23} = 1 - (9)(1) - (-3)(1) = -5$$

Thus,

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 9 & 6 & 1 \end{bmatrix}$$

W got the decomposition right, as the multiplication of the L and U gives the original matrix.

The original equation is equivalent to $LU\mathbf{x} = L\mathbf{w} = \mathbf{y}$,
$L\mathbf{w} = \mathbf{y}$ implies

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 9 & -3 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -4 \\ 7 \end{bmatrix}$$

Solving,

$$w_1 = \quad -1$$
$$w_1 + w_2 = -4 \text{ or } w_2 = -4 - w_1 = -4 - (-1) = -3$$
$$9w_1 - 3w_2 + w_3 = 7\text{, or } w_3 = 7 + 3w_2 - 9w_1 = 7 + 3(-3) - 9(-1) = 7$$

$U\mathbf{x} = \mathbf{w}$ implies:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ -3 \\ 7 \end{bmatrix}$$

By back substitution,

$$x_3 = -7/5 = -1.4$$
$$x_2 + x_3 = -3 \Rightarrow x_2 = -3 - x_3 = -3 - (-7/5)$$
$$x_2 = -8/5 = -1.6$$
$$x_1 + x_2 + x_3 = -1$$
$$x_1 = -1 - x_2 - x_3 = -1 - (-8/5) - (-7/5) = \frac{10}{5} = 2$$

The solution set is therefore,

$$x_1 = 2,\ y = -1.6,\ z = -1.4.$$

It is usually a good practice to revert to fractions to avoid incurring rounding errors.

5 Methods for finding Roots of Algebraic and Transcendental equations

In all scientific fields, there's always the need to find the root of an equation, equivalently the zero of a function. Numerical methods allow for more complicated cases of handling roots of quadratic and polynomial equations.

Bisection method

As the name implies, we obtain the points x_1 and x_2, such that $f(x_2)\,f(x_1) < 0$, meaning that the value of f has opposite signs at the two points, which points to the fact that a root exists between x_1 and x_2. We approximate this root by the average of the two, i.e., $(x_1 + x_2) / 2$. Let this be x_3. Then we evaluate $f(x_3)$. x_3 is then combined with x_1 or x_2, depending on the one at which the sign of the function is opposite $f(x_3)$. This gives x_4. This process is repeated until f(x) attains the prescribed tolerance. The convergence of the Bisection method is slow and steady.

Bisection Method

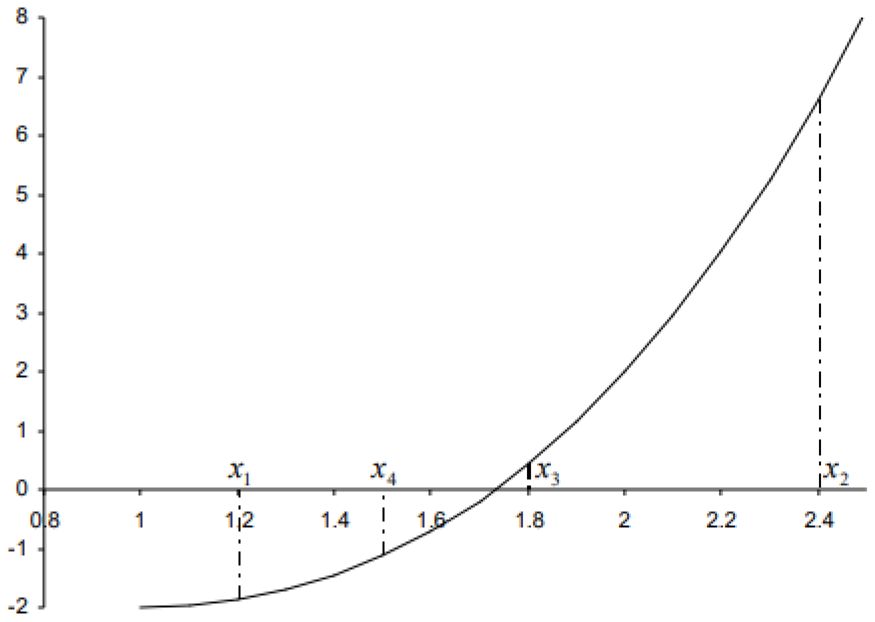


Fig.

- The **Bisection method** is one of the simplest methods to find a zero of a nonlinear function.
- It is also called **interval halving** method.
- To use the Bisection method, one needs an initial interval that is known to contain a zero of the function.
- The method systematically reduces the interval. It does this by dividing the interval into two equal parts, performs a simple test and based on the result of the test, half of the interval is thrown away.
- The procedure is repeated until the desired interval size is obtained.

Bisection Algorithm

Assumptions:

- $f(x)$ is continuous on $[a,b]$
- $f(a)\, f(b) < 0$

Loop

1. Compute the mid point $c=(a+b)/2$
2. Evaluate $f(c)$
3. If $f(a)\, f(c) < 0$ then new interval $[a, c]$

   If $f(a)\, f(c) > 0$ then new interval $[c, b]$

End loop

Question 5a

Can you use Bisection method to find a zero of:

$f(x) = x^3 - 3x + 1$ in the interval $[0,2]$?

Answer:

$f(x)$ is continuous on $[0,2]$

and $f(0) * f(2) = (1)(3) = 3 > 0$

$\Rightarrow$ Assumptions are not satisfied

$\Rightarrow$ Bisection method can not be used

Question 5b

Can you use Bisection method to find a zero of:

$f(x) = x^3 - 3x + 1$ in the interval $[0,1]$?

Answer:

$f(x)$ is continuous on $[0,1]$

and $f(0) * f(1) = (1)(-1) = -1 < 0$

$\Rightarrow$ Assumptions are satisfied

$\Rightarrow$ Bisection method can be used

Question 5c

Find the root of:

$$f(x) = x^3 - 3x + 1 \text{ in the interval } :[0,1]$$

* $f(x)$ is continuous

* $f(0) = 1,\ f(1) = -1 \Rightarrow f(a)\, f(b) < 0$

$\Rightarrow$ Bisection method can be used to find the root

Iteration	a	b	c= (a+b) / 2	f(c)	(b-a) / 2
1	0	1	0.5	-0.375	0.5
2	0	0.5	0.25	0.266	0.25
3	0.25	0.5	.375	-7.23E-3	0.125
4	0.25	0.375	0.3125	9.30E-2	0.0625
5	0.3125	0.375	0.34375	9.37E-3	0.03125

Question 5d

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ between the points 1.4 and 1.7, using the bisection method. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(1.4) = -0.192$$

$$f(1.7) = 0.756$$

$$x_3 = \frac{1.4 + 1.7}{2} = 1.55$$

$$f(1.55) = 1.4025 \times 10^{-1}$$

$$x_4 = \frac{1.55 + 1.4}{2} = 1.475$$

$$f(1.475) = -0.0588$$

$$x_5 = \frac{1.55 + 1.475}{2} = 1.5125$$

This confirm that the Table for Bisection method is indeed true

n	x	$f(x)$
1	1.55	0.14025
2	1.475	-5.88E-02
3	1.5125	3.22E-02
4	1.49375	-1.54E-02
5	1.503125	7.87E-03
6	1.498437	-3.89E-03
7	1.500781	1.96E-03
8	1.499609	-9.76E-04
9	1.500195	4.89E-04
10	1.499902	-2.44E-04
11	1.500049	1.22E-04
12	1.499976	-6.10E-05

CONVERGENCE ANALYSIS OF BISECTION METHOD

Bisection method obtains an interval that is guaranteed to contain a zero of the function.

Questions:

- What is the best estimate of the zero of $f(x)$?
- What is the error level in the obtained estimate?

The best estimate of the zero of the function $f(x)$ after the first iteration of the Bisection method is the mid point of the initial interval:

$$\textit{Estimate of the zero}: \; r = \frac{b+a}{2}$$

$$\textit{Error} \leq \frac{b-a}{2}$$

$$\textit{Given} \quad f(x),\ a,\ b,\ \textit{and} \ \varepsilon$$

How many iterations are needed such that: $|x - r| \leq \varepsilon$ where r is the zero of $f(x)$ and x is the bisection estimate (i.e., $x = c_k$)?

$$\boxed{n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)}}$$

$$\boxed{n \geq \log_2\left(\frac{\text{width of initial interval}}{\text{desired error}}\right) = \log_2\left(\frac{b-a}{\varepsilon}\right)}$$

Two common stopping criteria

1. Stop after a fixed number of iterations
2. Stop when the absolute error is less than a specified value

After n iterations :

$$|error| = |r - c_n| \leq E_a^n = \frac{b-a}{2^n} = \frac{\Delta x^0}{2^n}$$

Question 5e

$a = 6,\ b = 7,\ \varepsilon = 0.0005$

How many iterations are needed such that: $|x - r| \leq \varepsilon$?

$$n \geq \frac{\log(b-a) - \log(\varepsilon)}{\log(2)} = \frac{\log(1) - \log(0.0005)}{\log(2)} = 10.9658$$

$$\Rightarrow n \geq 11$$

Question 5f

- Use Bisection method to find a root of the equation x = cos(x) with absolute error <0.02
  (assume the initial interval [0.5, 0.9])

What is *f (x)* ?

Are the assumptions satisfied ?

How many iterations are needed ?

How to compute the new estimate ?

Question 5f (i) – what is f(x)?

x = cos(x)

f(x) = x – cos(x)

Question 5f (ii) – Are the assumptions satisfied?

Assuming interval [0.5, 0.9]

f(0.5) = 0.5 – cos(0.5) = -0.3776; This is a negative value

f(0.9) = 0.9 – cos(0.9) = 0.2784; This is a positive value

f(0.5)*f(0.9) = -0.3776 * 0.2764 < 0; Assumption is therefore satisfied.

Bisection method can be used.

Question 5f (iii) – How many iterations are needed?

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

a = 0.5, b = 0.9, ε = 0.02

n $\geq$ [log(0.9 – 0.5) – log(0.02)] / log(2)

n $\geq$ [-0.3979 - -1.6990] / 0.3010

n $\geq$ 1.3011 / 0.3010

n $\geq$ 4.3226

n $\geq$ 5

Question 5f (iii) – How to compute the new estimate?

$$\textit{Estimate of the zero}: \ r = \frac{b+a}{2}, \quad \textit{Error} \leq \frac{b-a}{2}$$

r1 = (0.9 + 0.5) / 2 = 0.7; Error < (0.9 - 0.5)/2 $\leq$ 0.2;

f(0.7) = 0.7 – cos(0.7) = 0.7 – 0.9999 = -0.2999

f(0.5) = -0.3776; f(0.9) = 0.2784; f(0.7) = -0.2999

r2 = (0.7 + 0.9) / 2 = 0.8, **Error < (0.9 - 0.7) / 2 ≤ 0.1**

f(0.8) = 0.8 – cos(0.8) = 0.8 – 0.9999 = -0.1999

f(0.7) = -0.2999; f(0.9) = 0.2784; f(0.8) = -0.1999

r3 = (0.8 + 0.9) / 2 = 0.85, **Error < (0.9 - 0.8) / 2 ≤ 0.5**

f(0.85) = 0.85 – cos(0.85) = 0.85 – 0.9999 = -0.1499

f(0.8) = -0.1999; f(0.9) = 0.2784; f(0.85) = -0.1499

r4 = (0.85 + 0.9) / 2 = 0.875, **Error < (0.9 - 0.85) / 2 ≤ 0.025**

f(0.875) = 0.875 – cos(0.875) = 0.875 – 0.9999 = -0.1249

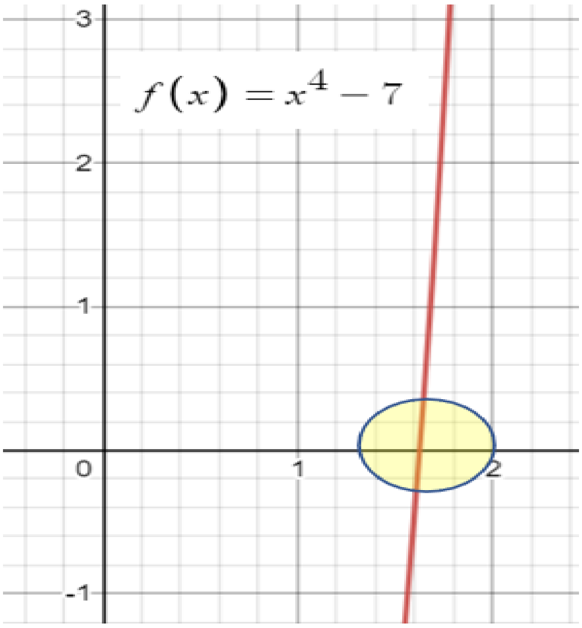
f(0.85) = -0.1499; f(0.9) = 0.2784; f(0.875) = -0.1249

r5 = (0.875 + 0.9) / 2 = 0.8875, **Error < (0.9 - 0.875) / 2 ≤ 0.02**

f(0.8875) = 0.8875 – cos(0.8875) = 0.8875 – 0.9999 = -0.1124

f(0.875) = -0.1249; f(0.9) = 0.2784; f(0.8875) = -0.1124

Question 5g



Find the 3rd approximation of the root of f(x) = x^4–7 using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(left)	f(mid)	f(right)	New Interval	Midpoint
f(1) = -6	f(1.5) = -2	f(2) = 9	(1.5, 2)	1.75
f(1.5) = -2	f(1.75) = 2.4	f(2) = 9	(1.5, 1.75)	1.625
f(1.5) = -2	f(1.625) = -.03	f(1.75) = 2.4	(1.625, 1.75)	1.6875

$f(x) = x^4 – 7$

$f(2) = (2)^4 – 7 = 9$; this is positive

$f(1) = (1)^4 – 7 = -6$; this is negative

$f(2)*f(1) = 9 * -6 < 0$; Assumption is therefore satisfied. Bisection method can be used.

for;

Starting interval (1, 2)

mid x = [2+1] / 2 = 1.5; Initial estimate

$f(mid) = f(1.5) = (1.5)^4 – 7 = 5.0625 – 7 = -1.9375$.

for;

f(2) = 9, f(1) = -6, **f(1.5) = -1.9375**

Next interval (2, 1.5)

mid x = [2+1.5]/2 = 1.75; first approximation

$f(mid) = f(1.75) = (1.75)^4 – 7 = 9.3789 – 7 = 2.3789$.

for;

f(2) = 9**,** **f(1.5) = -1.9375,** **f(1.75) = 2.3789**

Next interval (1.75, 1.5)

mid x = [1.75+1.5]/2 = 1.625; second approximation

$f(mid) = f(1.625) = (1.625)^4 – 7 = 6.9729 – 7 = -0.0271$.

for;

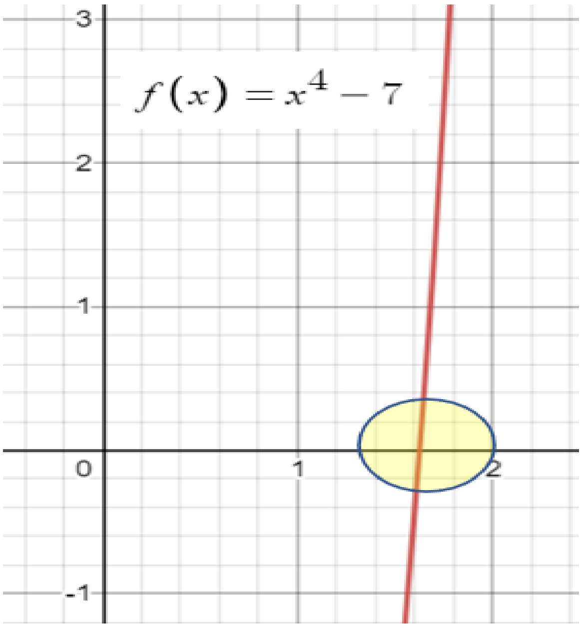
f(1.75) = 2.3789, f(1.5) = -1.9375, **f(1.625) = -0.0271**

Next interval (1.75, 1.625)

mid x = [1.75+1.625]/2 = 1.6875; third approximation

$f(mid) = f(1.6875) = (1.6875)^4 – 7 = 8.1091 – 7 = 1.1091$. **Stop.**

Question 5g



Find the 3rd approximation of the root of f(x) = 10 – x^2 using the bisection method

Solution

The function changes from – to + somewhere in the interval x = 1 to x = 2. So we use [1, 2] as the starting interval.

f(x) = 10 – x^2

f(2) = 10 - $(2)^2$ = 6; this is positive

f(1) = 10 - $(1)^2$ = 9; this is also positive

f(2)*f(1) = 6 * 9 !< 0; Assumption is NOT satisfied. Bisection method cannot be used.

Newton-Raphson Method

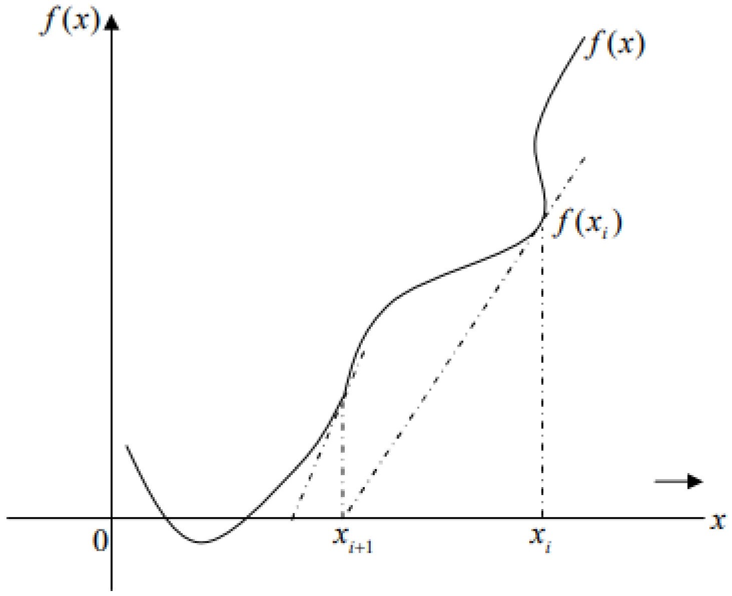
It is quite clear that the function $f(x)$ must be differentiable for you to be able apply the Newton-Raphson method.

More generally,

$$x_{i+1} = x_i + \Delta x = x_i - \frac{f(x_i)}{f'(x_i)}$$

With an initial guess of x_0, we can then get a sequence x_1, x_2, ..., which we expect to converge to the root of the equation.

Newton-Raphson method is equivalent to taking the slope of the function f(x) at the i^{th} iterative point, and the next approximation is the point where the slope intersects the x axis.



Question 5g

Find a zero of the function $f(x) = 2x^3 - 3x^2 - 2x + 3$ starting with the point 1.4, using the **Newton-Raphson Method**. Take the tolerance to be $|x_{j+1} - x_j| \leq 10^{-5}$.

Solution

$$f(x) = 2x^3 - 3x^2 - 2x + 3$$

$$f'(x) = 6x^2 - 6x - 2$$

$$x_0 = 1.4$$

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

$$= \frac{6x_0^{\ 3} - 6x_0^{\ 2} - 2x_0 - 2x_0^{\ 3} + 3x_0^{\ 2} + 2x_0 - 3}{6x_0^{\ 2} - 6x_0 - 2}$$

$$= \frac{4x_0^{\ 3} - 3x_0^{\ 2} - 3}{6x_0^{\ 2} - 6x_0 - 2}$$

$$= \frac{4(1.4)^3 - 3(1.4)^2 - 3}{6(1.4)^2 - 6(1.4) - 2}$$

$$= 1.5412$$

$x_1 = 1.5412$, $|x_1 - x_0| = 0.1412$

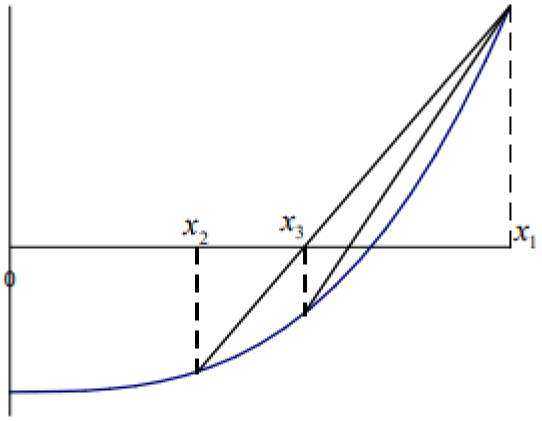
$x_2 = 1.5035$, $|x_2 - x_1| = 0.0377$

$x_3 = 1.5$, $|x_3 - x_2| = 0.0035$

$x_4 = 1.5$, $|x_4 - x_3| = 0$

Regula-falsi method

A regula-falsi or a method of false position assumes a test value for the solution of the equation.



Then, for an arbitrary x and the corresponding y,

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

gives the equation of the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Setting $y = 0$, that is, where the chord crosses the x-axis,

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

Then, we evaluate $f(x_3)$. Just as in the case of root-bisection, if the sign is opposite that of $f(x_1)$, then a root lies in-between x_1 and x_3. Then, we replace x_2 by x_3 in equation

In just the same way, if the root lies between x_1 and x_3, we replace x_2 by x_1. We shall repeat this procedure until we are as close to the root as desired.

Question 5h

Find a zero of the function f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the regula-falsi method.

$f(1.4) = -0.192$, $f(1.7) = 0.756$

A solution lies between x = 1.4 and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$= 1.4607595$

$f(1.4607595) = -0.088983$

The root lies between 1.46076 and 1.7. Let $x_1 = 1.46076$ and $x_2 = 1.7$.

$$x_4 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4607595 - (-0.088983)\frac{1.7 - 1.46076}{0.756 - (-0.088983)}$$
$$= 1.485953$$

Table for Regula-falsi method

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.495149	-0.011985
4	1.498346	-0.004118
5	1.499439	-0.001401
6	1.499810	-0.000475
7	1.499936	-0.000161
8	1.499978	-0.000055

Secant Method

In the case of the secant method, it is not necessary that the root lie between the two initial points. As such, the condition $f(x_1)f(x_2) < 0$ is not needed. Following the same analysis with the case of the regula-falsi method,

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

Setting $y = 0$ gives

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

Thus, having found x_n, we can obtain x_{n+1} as,

$$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})},\ n = 2, 3, \ldots$$

By inspection, if $f(x_n) - f(x_{n-1}) = 0$, the sequence does not converge, because the formula fails to work for x_{n+1}. The regula-falsi scheme does not have this problem as the associated sequence always converges.

Question 5i

Find the roots of the equation f(x) = $2x^3 - 3x^2 - 2x + 3$ between x = 1.4 and 1.7 by the secant method.

$x_1 = 1.4,\ x_2 = 1.7$

$f(1.4) = -0.192$, $f(1.7) = 0.756$

A solution lies between $x = 1.4$ and 1.7. Let $x_1 = 1.4$ and $x_2 = 1.7$. Then,

$$x_3 = x_1 - f(x_1)\frac{x_2 - x_1}{f(x_2) - f(x_1)} = 1.4 - (-0.192)\frac{1.7 - 1.4}{0.756 - (-0.192)}$$

$$= 1.460759$$

$f(x_3) = -0.088983$

$$x_4 = x_3 - f(x_3)\frac{x_3 - x_2}{f(x_3) - f(x_2)} = 1.460759 - (-0.088983) \times \frac{1.460759 - 1.7}{-0.088983 - 0.756}$$

$$= 1.485953$$

If the scheme continues, the table for secant method will be

n	x	$f(x)$
1	1.460759	-0.088983
2	1.485953	-0.033938
3	1.501487	0.003730
4	1.499949	-0.000129
5	1.500000	0.000000