Chapter 15A: Numerical Methods for Solving Equations

SYLLABUS INCLUDES

H2 Further Mathematics:

- Approximation of roots of equations using linear interpolation and Newton-Raphson method.
- Iterations involving recurrence relations of the form $x_{n+1} = \mathrm{F}(x_n)$.

PRE-REQUISITES

- Differentiation Techniques

CONTENT

1 INTRODUCTION

In mathematics we often encounter problems that cannot be solved by any analytical method, i.e. obtaining exact solutions from methods of symbolic manipulation. For example, there are many integrals that cannot be performed analytically, many equations that cannot be easily solved algebraically, and many differential equations whose solution cannot be expressed in a simple form.

In certain branches of applied mathematics, it is not always necessary to know the analytical solution to a problem, even if it exists. It is usually more important to an engineer to have some sufficiently good numerical approximation, so that he may use it in his calculations, and eventually in his constructions.

This, then, is the reason for studying numerical methods: we study it so that we may provide numerical solutions to problems which are either unsolvable by analytical methods, or for which numerical solutions are needed in practical situations.

In this chapter, we will be first investigating some methods of approximating a root of an equation. In the next two chapters, we shall be looking at some techniques for approximating the definite integral and the numerical solutions to first order differential equations.

2 APPROXIMATION OF ROOTS

The roots of linear and quadratic equations can be easily obtained. Some cubic and quartic equations may be solved by the remainder theorem. However, there are many other equations which are impossible to solve exactly, such as $x-\cos x=0$ or $xe^x-1=0$.
We can only find an approximation to the roots of these equations.

In this chapter, we will discuss the use of
(i) Bisection
(ii) Linear Interpolation
(iii) Newton-Raphson (Newton's) Method
to approximate such roots to the required accuracy.

2.1 LOCATING THE ROOTS OF AN EQUATION

Suppose we want to solve the equation f(x) = g(x), the first step is to get an idea of where the roots are. In other words, we try to locate the roots. This can be done graphically as follows:

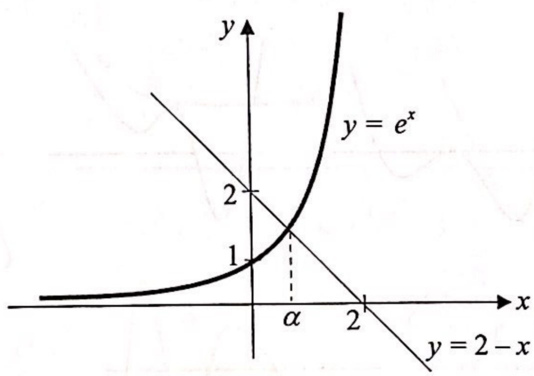
1. Sketch the two graphs $y = \text{f}(x)$ and $y = \text{g}(x)$ on the same axes.
2. The number of roots of the equation f(x) = g(x) corresponds to the number of points of intersection of the above two graphs.
3. The rough positions of the roots can be obtained by looking at the x-coordinate of the points of the intersection of the above two graphs.

Example 1

By sketching the graphs of $y = \text{e}^x$ and $y = 2 - x$ on the same diagram, show that the equation $\text{e}^x + x - 2 = 0$ has only one real root, α. Find the integer N such that $N < \alpha < N + 1$.

Solution

$\text{e}^x + x - 2 = 0 \Rightarrow e^x = 2 - x$



The graphs of $y = e^x$ and $y = 2 - x$ intersect at exactly one point.
Hence, the equation has only one real root, α.

Let $f(x) = e^x + x - 2$ $e + 1 - 2 = 1.71$

We note that $f(0) = -1 < 0$ and $f(1) = 1.71 > 0$

$f(0)f(1) < 0$

$0 < \alpha < 1$

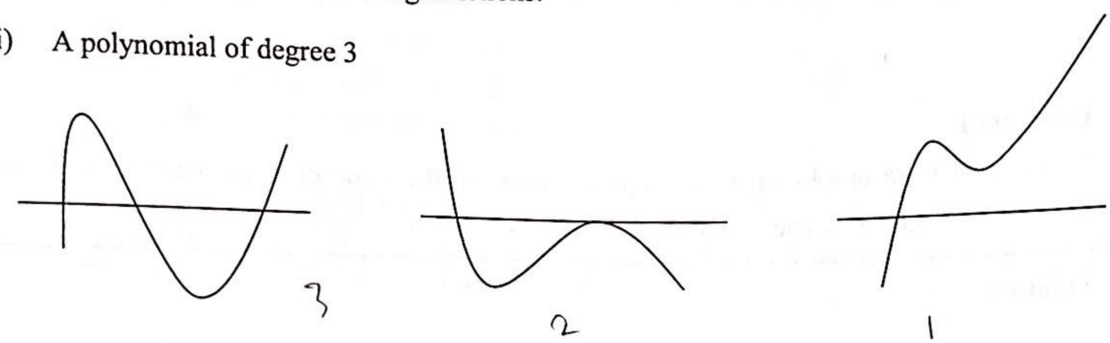
$N = 0$

To find the roots of an equation, the first step is to get some idea of their whereabouts, i.e., we try to locate the roots. In what follows, we discuss 3 results that are helpful in determining the number and locality of roots.

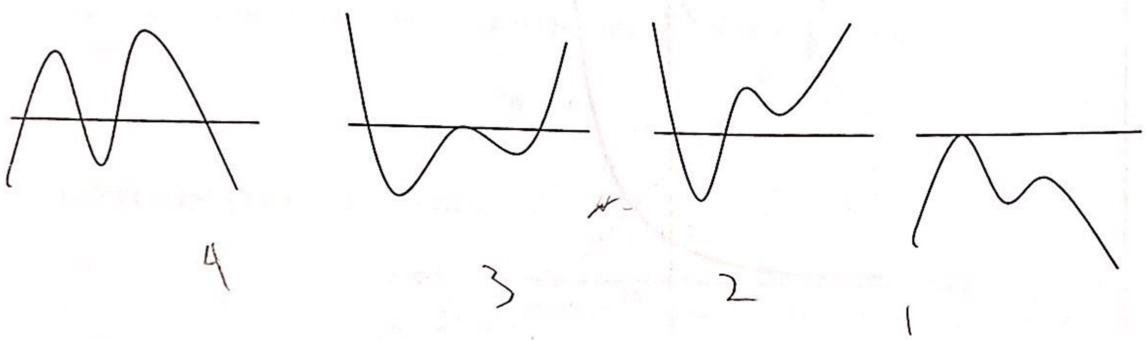
Discussion 1

Consider the curves of the following functions:

(i) A polynomial of degree 3



(ii) A polynomial of degree 4



Hence, we have

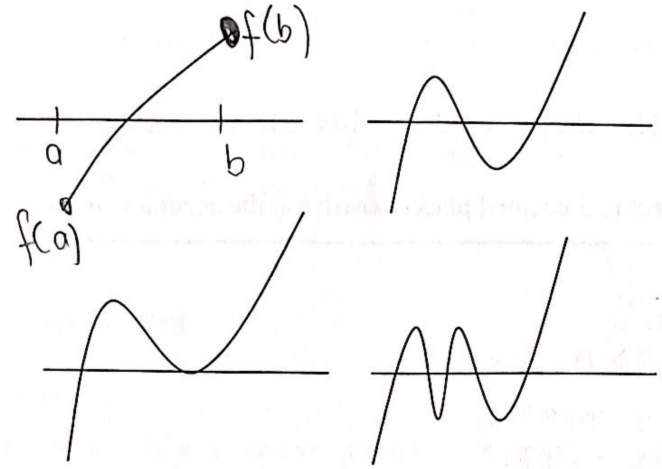
Result 1 : Given a polynomial $\mathrm{P}(x)$ of degree n, the equation $\mathrm{P}(x) = 0$ has *at most n* real roots.

Remark : Given a function $\mathrm{f}(x)$, a value x that satisfies $\mathrm{f}(x) = 0$ is called a zero of the function $\mathrm{f}(x)$.

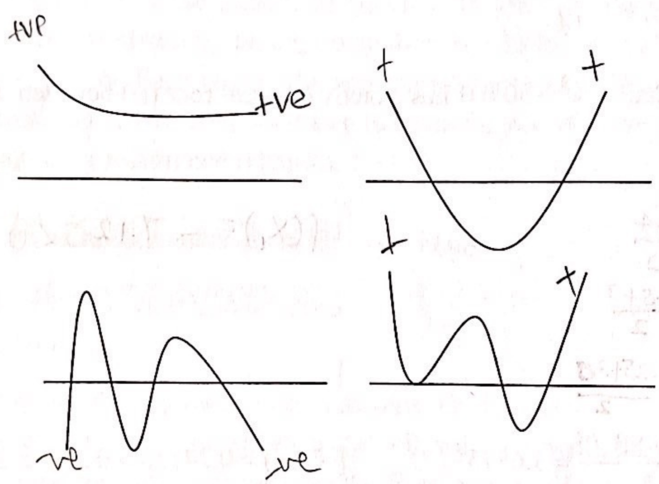
Discussion 2

Consider a *continuous function* f(x) on the interval (a, b). In the diagrams below, the endpoints of the graphs are located at $x = a$ and $x = b$.

(i) $f(a)\,f(b) < 0$



(ii) $f(a)\,f(b) > 0$



Result 2: If f is a *continuous function* in the interval (a, b) such that $f(a)\,f(b) < 0$, then the equation $f(x) = 0$ has an odd number of real roots between a and b.

Result 3: If f is a *continuous function* in the interval (a, b) such that $f(a)\,f(b) > 0$, then the equation $f(x) = 0$ has either no real roots or an even number of real roots between a and b.

Question: Is it necessary for the function to be *continuous* ?

2.2 BISECTION

Suppose that $f(x)$ is continuous on an interval $a \le x \le b$ and that $f(a)f(b) < 0$.
Then $f(x)$ changes sign on $[a,b]$, and $f(x) = 0$ has at least one root on the interval.

The simplest numerical procedure for finding a root is to repeatedly halve the interval $[a,b]$, keeping the half on which $f(x)$ changes sign. This procedure is called the bisection.

Example 2

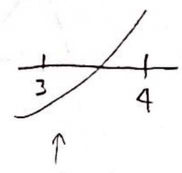
Show that the cubic equation $x^3 - 50 = 0$ has only one real root and that the root lies between $x = 3$ and $x = 4$.
Find the root correct to 3 decimal places, justifying the accuracy of your answer.

Solution

Let $f(x) = x^3 - 50$.

$\Rightarrow f'(x) = 3x^2 > 0 \ \forall x \in \mathbb{R}$

$\therefore f$ is increasing



We note that $f(3) = -23 < 0$
$f(4) = 14 > 0$

Therefore the equation $x^3 - 50 = 0$ has exactly one real root (α) between $x = 3$ and $x = 4$.

By bisection,

$x_1 = \frac{3+4}{2}$ [$f(x_1) = -7.125 < 0 \therefore \alpha \in ($]

$x_2 = \frac{3.5+3}{2}$ [ ]

$x_3 = \frac{3.25+3}{2}$ [ ]

$x_4 = \frac{3.625+3.75}{2} = 3.6875$ [$f(x_4) = 0.14136 > 0$. $\therefore \ \alpha \in (x_3, x_4)$]

$x_5 = \frac{3.625+3.6875}{2} = 3.65625$ [$f(x_5) = -1.12265 < 0$. $\therefore \ \alpha \in (x_4, x_5)$]

$x_6 = \frac{3.65625+3.6875}{2} = 3.67188$ [$f(x_6) = -0.49333 < 0$. $\therefore \ \alpha \in (x_4, x_6)$]

$x_7 = \frac{3.67188+3.6875}{2} = 3.67969$ [$f(x_7) = -0.17666 < 0$. $\therefore \ \alpha \in (x_4, x_7)$]

$x_8 = \frac{3.67969+3.6875}{2} = 3.68360$ [$f(x_8) = -0.01777 < 0$. $\therefore \ \alpha \in (x_4, x_8)$]

$x_9 = \frac{3.68360+3.6875}{2} = 3.68555$ [$f(x_9) = 0.06185 > 0$. $\therefore \ \alpha \in (x_8, x_9)$]

$x_{10} = \frac{3.68360+3.68555}{2} = 3.68458$ [$f(x_{10}) = 0.02213 > 0$. $\therefore \ \alpha \in (x_8, x_{10})$]

Checking,

$f(3.6835) = -0.02164 < 0$

$f(3.6845) = 0.01908 > 0$

So, $3.6835 < \alpha < 3.6845$

Hence, $\alpha = 3.684$ (3dp)

There are several advantages to the bisection. The principal one is that the method is guaranteed to converge. The principal disadvantage of the bisection is that it generally converges more slowly than most other methods.

3 ITERATIVE METHODS

3.1 Iterative Process

An iterative process is a repetitive procedure designed to produce a sequence of approximations $\{x_n\}$ to some numerical quantity. In general, this process is based on an algorithm or rule for obtaining an approximation x_{n+1} based on earlier approximations x_n, x_{n-1}, …, x_2, x_1 and x_0. Each step in the approximation process is called an iteration. Very often, each iteration can be expressed as a formula for x_{n+1} in terms of x_n. This is called an iterative formula (or recurrence relation).

The general iterative formula is given by $x_{n+1} = \mathrm{F}(x_n)$.

3.2 Fixed Point Iteration

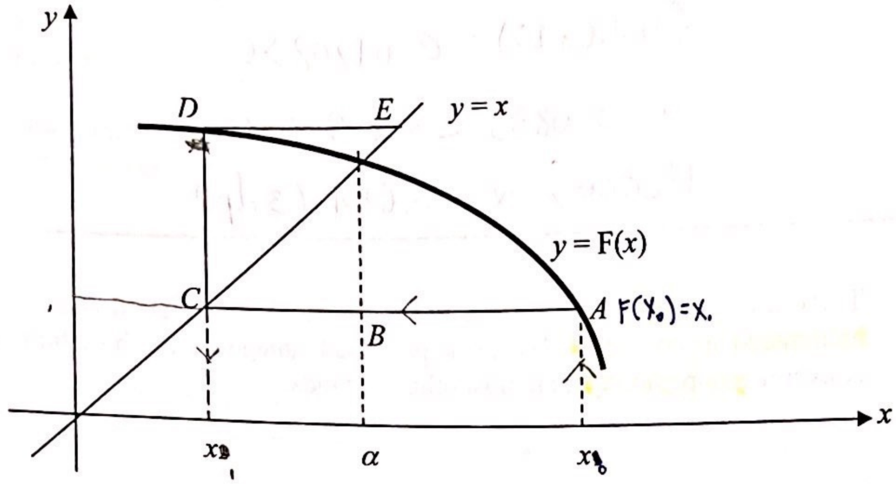
A point $x = \alpha$ is a **fixed point** of the function $\mathrm{F}(x)$ if $\mathrm{F}(\alpha) = \alpha$.
Hence, if we want to solve the equation $\mathrm{f}(x) = 0$, we can rewrite the equation into the form $x = \mathrm{F}(x)$ for some function $\mathrm{F}(x)$ and look for fixed points of $\mathrm{F}(x)$.

For example, to solve the equation $\mathrm{f}(x) = x^2 - 3 = 0$, we can rewrite it into:

(a) $x = x^2 + x - 3 = \mathrm{F}_1(x)$, or

(b) $x = \dfrac{3}{x} = \mathrm{F}_2(x)$, or

(c) $x = \dfrac{1}{2}\left(x + \dfrac{3}{x}\right) = \mathrm{F}_3(x)$.



Refer to the diagram above :

To start using the fixed point iterative formula $x_{n+1} = \mathrm{F}(x_n)$, we pick an initial value x_0 and compute $x_1 = \mathrm{F}(x_0)$ as the first approximation of α.

We then continue to compute $x_2 = \mathrm{F}(x_1)$ as the next approximation of α and so on until the required accuracy is achieved.

In general, $x_{n+1} = \mathrm{F}(x_n)$.

This whole procedure is called fixed point iteration.

As mentioned earlier, when solving $\mathrm{f}(x) = 0$ by using the concept of fixed point, we need to rewrite the equation into the form $x = \mathrm{F}(x)$. If the sequence $\{x_n\}$ converges, then the iterative formula $x_{n+1} = \mathrm{F}(x_n)$ will produce a sequence of approximations that converges to the root of $\mathrm{f}(x) = 0$. We will see that there are many choices for $\mathrm{F}(x)$. However, not all choices of $\mathrm{F}(x)$ will produce sequences that converge. Unfortunately, there is no easy way to predict which choices of $\mathrm{F}(x)$ is appropriate. But for 'A' levels, $\mathrm{F}(x)$ is usually given in the question. We just need to know how to apply the iterative formula and relate it to solving the equation $\mathrm{f}(x) = 0$.

Example 3

Show that the cubic equation $x^3+2x-11=0$ has only one real root and that the r
$x=1$ and $x=2$.

Show that two possible iterative formulas for finding the root are

(i) $x_{n+1}=\frac{1}{2}\left(11-x_n^{\,3}\right)$ and **(ii)** $x_{n+1}=\left(11-2x_n\right)^{\frac{1}{3}}$.

Show that only one of these iterative formulas converges from an initial estimate of $x=2$ and hence find the root correct to 3 decimal places, justifying the accuracy of your answer.

Solution

Let $\mathrm{f}(x)=x^3+2x-11$.

$\Rightarrow f'(x)=3x^2+2>0 \quad \forall x\in\mathbb{R}$ ~~for~~

$\therefore$ f is increasing

We note that $\mathrm{f}(1)=-8<0$

$\mathrm{f}(2)=1>0$

Therefore the equation $x^3+2x-11=0$ has exactly one real root between $x=1$ and $x=2$.

(i) Given $x_{n+1}=\frac{1}{2}\left(11-x_n^{\,3}\right)$.

Suppose the sequence converges to *l*, then *l* satisfies

$l=\frac{1}{2}(11-t^3)$

$\Rightarrow 2l=11-l^3$

$\Rightarrow l^3+2l-11=0$

l is a root of $x^3+2x-11=0$.

(ii) Given $x_{n+1}=\left(11-2x_n\right)^{\frac{1}{3}}$. $n\to\infty,\ x_n\to l,\ x_{n+1}\to l$

Suppose the sequence converges to *l*, then *l* satisfies

$l=\frac{1}{2}(11-l^3)$

$2l=11-l^3$

$l^3+2l-11=0$

l is a root of $x^3+2x-11=0$.

Using $x_{n+1}=\frac{1}{2}\left(11-x_n^{\,3}\right)$.

From G.C. :

x_1	x_2	x_3	x_4	x_5
2	1.5	3.8125	-22.20764	5481.67508

G.C. screenshot :

n	u(n)	
1	2	
2	1.5	
3	3.8125	
4	-22.21	
5	5481.7	
6	-8E10	
7	2.8E32	

n=7

The sequence diverges

Using $x_{n+1}=(11-2x_n)^{\frac{1}{3}}$.
From G.C. :

x_1	x_2	x_3	x_4	x_5	x_6
2	1.91293	1.92866	1.92584	1.92635	1.92626

The sequence Converges to 1.926

G.C. screenshot :

n	u(n)	
1	2	
2	1.9129	
3	1.9287	
4	1.9258	
5	1.9263	
6	1.9263	
7	1.9263	

n=7

Checking : $f(1.9255)=-0.0101<0$

$f(1.9265)=0.0030>0$

$\therefore$ the root is 1.926 (3dp)

Example 4 [9233/2003/01/Q9(i)]

Show that the equation $x^3+2x^2-2=0$ has exactly one positive root. This root is denoted by α and is to be found using an iterative method.

Show that α is a root of the equation $x=\sqrt{\dfrac{2}{x+2}}$ and use the iterative formula $x_{n+1}=\sqrt{\dfrac{2}{x_n+2}}$ with $x_1=1$, to find α correct to 3 significant figures.

Solution

Let $f(x)=x^3+2x^2-2 \Rightarrow f'(x)=3x^2+4x$

$f'(x)=3x^2+4x>0$ when $x>0$

$f(0)=-2<0$ and $f(1)=1>0$

$\therefore x^3+2x^2-2=0$ has exactly one +ve root

Given $x_{n+1} = \sqrt{\dfrac{2}{x_n + 2}}$.

$n \to \infty,\ x_n \to \alpha,\ x_{n+1} \to \alpha$

Suppose the sequence converges to α, then α satisfies

$$\alpha = \sqrt{\frac{2}{\alpha+2}}$$

$$\alpha^2 = \frac{2}{\alpha+2}$$

$$\alpha^2(\alpha+2 = 2$$

$$\alpha^3+2\alpha^2-2 = 0$$

Hence α is a root of $x = \sqrt{\dfrac{2}{x+2}}$.

Using $x_{n+1} = \sqrt{\dfrac{2}{x_n + 2}}$.

From G.C. :

x_1	x_2	x_3	x_4	x_5	x_6
1	0.81650	0.84268	0.83879	0.83936	0.83928

G.C. screenshot :

n	u(n)
1	1
2	.8165
3	.84268
4	.83879
5	.83936
6	.83928
7	.83929

n=1

The sequence converges to 0.839 (3 s.f.).

Checking $f(0.8385) = -0.0043010 < 0$

and $f(0.8395) = 0.0011667 > 0 \Rightarrow \alpha \in (0.8385, 0.8395)$

Hence $\alpha = 0.839$ (3sf)

3.3 Cases Where the Fixed Point Iteration May Fail

As we have seen earlier in Example 4, there are situations where the fixed point iteration may fail. Let us illustrate some of the possible cases here.

Consider solving the equation $x^2 - 5 = 0$ for the root $\alpha = 2.2361$.
Some possible iterative formulae for finding α are

(i) $x_{n+1} = 5 + x_n - (x_n)^2$

(ii) $x_{n+1} = \dfrac{5}{x_n}$

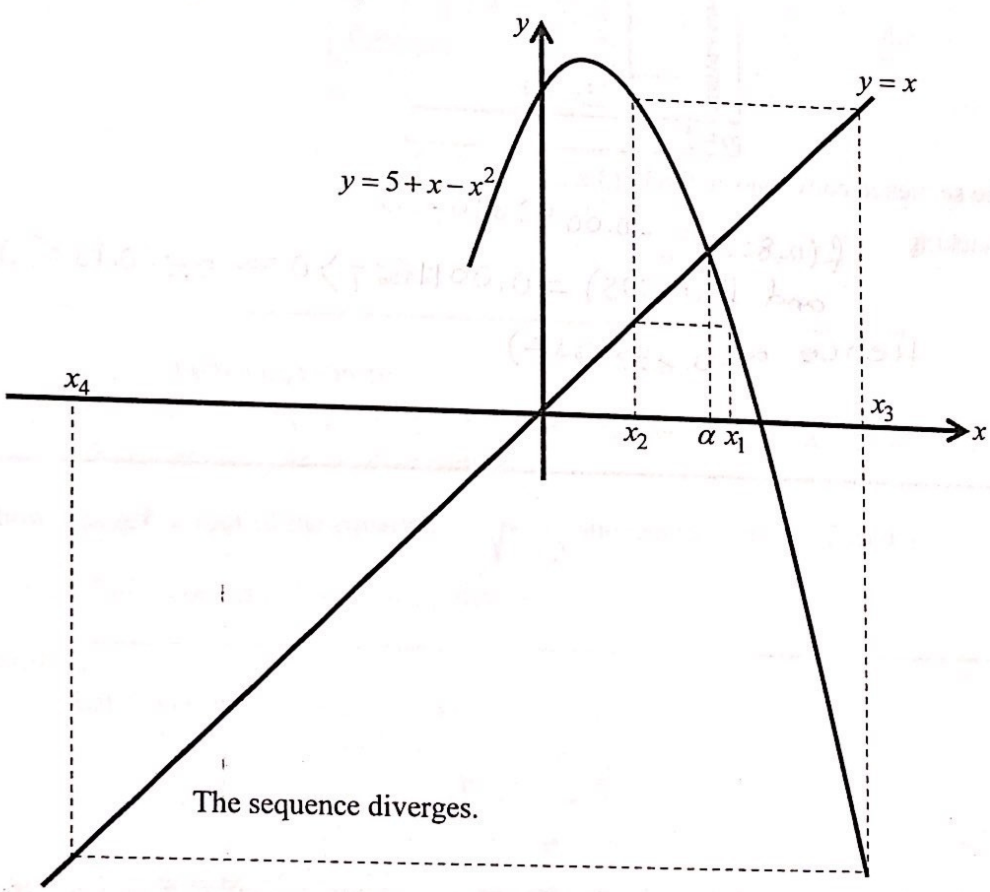
(iii) $x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{5}{x_n}\right)$

All three iterations have the property that if the sequence $\{x_n : n \geq 0\}$ has a limit l, then l is a root of $x^2 - 5 = 0$.

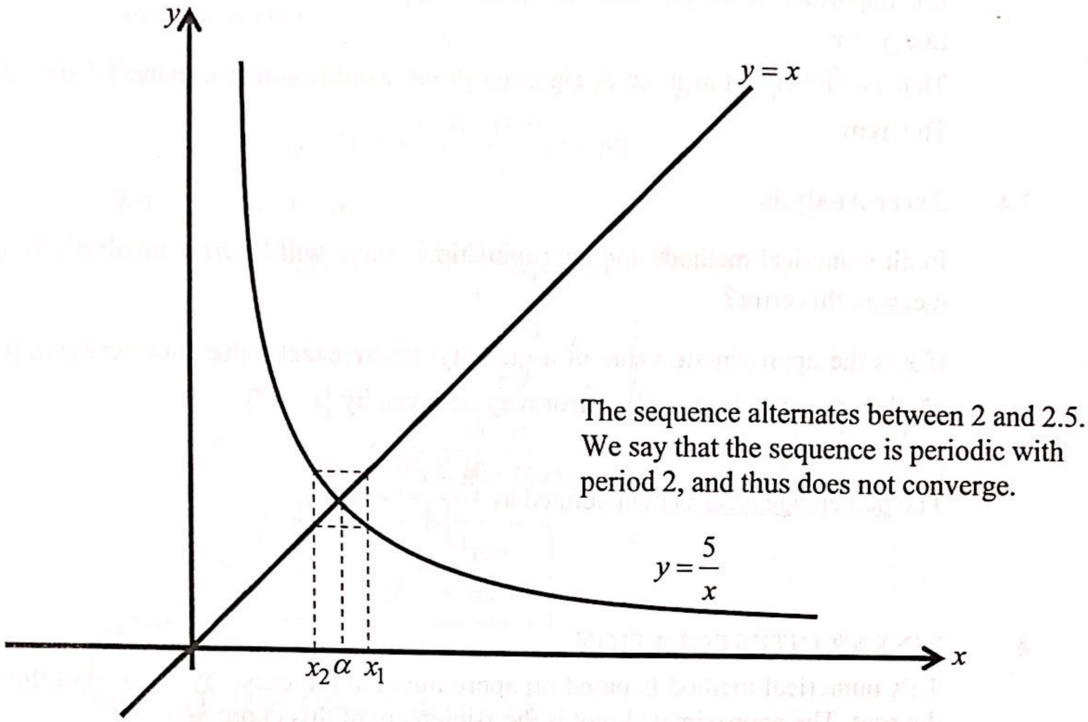
The table below gives the sequence of $\{x_n\}$ for these 3 iteration methods.

n	$\{x_n\}$: (i)	$\{x_n\}$: (ii)	$\{x_n\}$: (iii)
1	2.5	2.5	2.5
2	1.25	2.0	2.25
3	4.6875	2.5	2.2361
4	−12.2852	2.0	2.2361

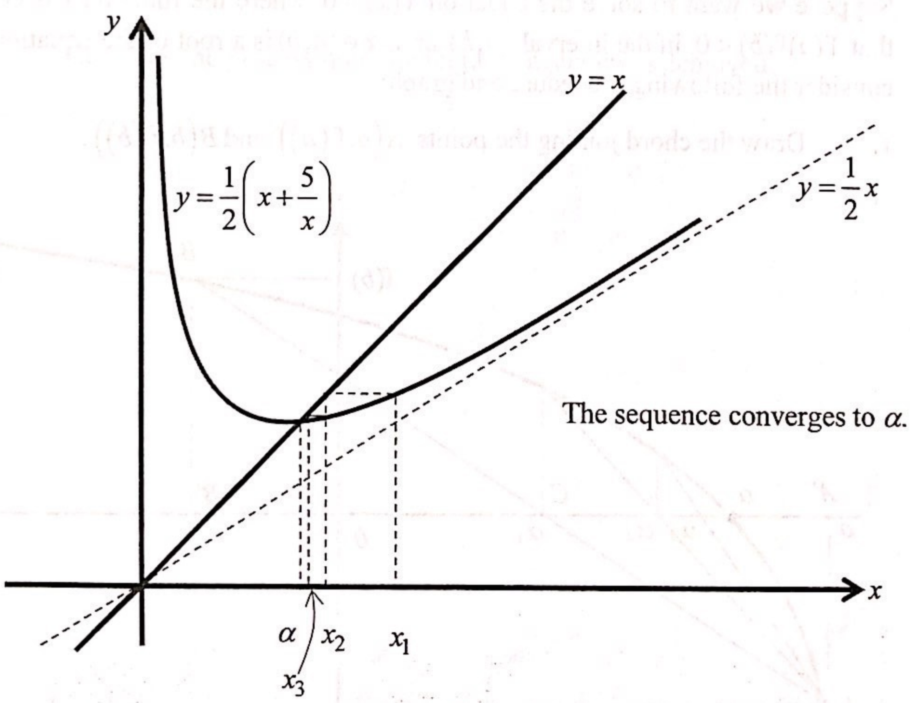
Method 1 : $x_{n+1} = 5 + x_n - (x_n)^2$



Method 2 : $x_{n+1} = \dfrac{5}{x_n}$



Method 3 : $x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{5}{x_n}\right)$



Consider the iterative formula $x_{n+1} = \mathrm{F}(x_n)$ used to approximate $\alpha \in (a,b)$.
It is important (in the region of α) for the slope of $y = \mathrm{F}(x)$ to be less steep than that of the line $y = x$.

That is, $|\mathrm{F}'(x)| < 1$ near α. A rigorous proof would require the use of the Mean Value Theorem.

3.4 Error Analysis

In all numerical methods and approximations, there will be error involved. So how do we measure this error?

If x is the approximate value of a quantity whose exact value (not necessarily known) is x^*, then the magnitude of the error may be given by $|x - x^*|$.

The percentage error is then defined as $\dfrac{|x - x^*|}{|x|} \times 100\%$.

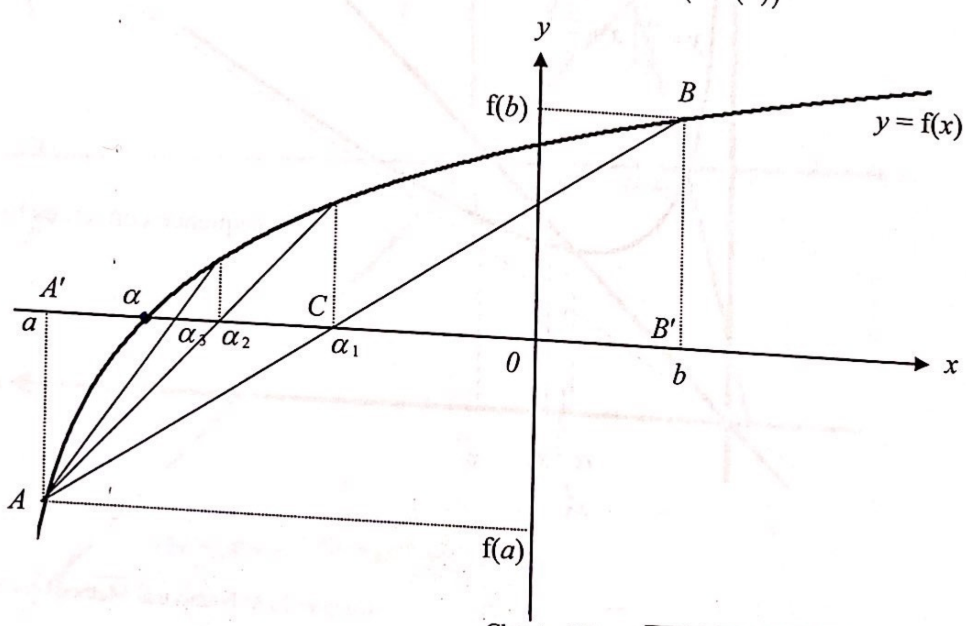
4 LINEAR INTERPOLATION

This numerical method is based on approximating the curve by a chord in the vicinity of the root. The approximated root is the x-intercept of this chord.

4.1 Derivation of Linear Interpolation

Suppose we want to solve the equation $\mathrm{f}(x) = 0$ where the function f is continuous such that $\mathrm{f}(a)\mathrm{f}(b) < 0$ in the interval (a,b) and $\alpha \in (a,b)$ is a root of the equation $\mathrm{f}(x) = 0$, we consider the following procedure and graph:

1. Draw the chord joining the points $A(a, \mathrm{f}(a))$ and $B(b, \mathrm{f}(b))$.



Obtain α_1, the first approximation of α, given by the x-intercept of the chord AB.

Gradient of chord $AB = \dfrac{f(b)-f(a)}{b-a}$.

Equation of the line passing through A and B, l_{AB}, is

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

When $y=0,\ x=\alpha_1$:

$$-f(b)=\frac{f(b)-f(a)}{b-a}(\alpha_1-b)$$

$$(\alpha_1-b)=-f(b)\left(\frac{b-a}{f(b)-f(a)}\right)$$

$$\Rightarrow \quad \alpha_1=-f(b)\left(\frac{b-a}{f(b)-f(a)}\right)+b=\frac{-f(b)(b-a)}{f(b)-f(a)}+\frac{b\left(f(b)-f(a)\right)}{f(b)-f(a)}$$

$$\Rightarrow \quad \boxed{\alpha_1=\frac{af(b)-bf(a)}{f(b)-f(a)}}$$

2. Now $f(a)f(\alpha_1)<0 \Rightarrow \alpha\in(a,\alpha_1)$.
   Repeat the steps above to obtain α_2, the second approximation of α will be
   $$\alpha_2=\frac{af(\alpha_1)-\alpha_1 f(a)}{f(\alpha_1)-f(a)}.$$

3. Continue the process until the required accuracy is achieved.

Example 5

Obtain, to three decimal places, the root of the equation $x = \cos x$ by performing linear interpolation on the interval $(0.5, 1.5)$.

Solution

Let $f(x) = x - \cos x$.

x	0.5	1.5
$f(x)$	−0.37758	1.42926

> The number of d.p. used in the intermediate calculation must be at least **two** more than what is required in the final answer. In this eg, since the required accuracy is 3 d.p., all intermediate calculations are given to 5 d.p.

By linear interpolation on (0.5, 1.5), we have

$$\alpha_1 = \frac{0.5f(1.5) - 1.5f(0.5)}{f(1.5) - f(0.5)}$$

$$= 0.70897$$

> When using a calculator, each intermediate step can be stored in the memory for use in the next iteration.

Since f(0.70897) = −0.05006 < 0 and f(1.5) = 1.42926 > 0

$\therefore \alpha$ lies in (0.70897, 1.5).

x	0.70897	1.5
$f(x)$	−0.05006	1.42926

$$\alpha_2 = \frac{0.70897 \times 1.4926 - 1.5(-0.05006)}{1.4226 - (-0.05006)} = 0.73573$$

Since f(0.73573) = −0.00559 < 0 and f(1.5) = 1.42926 > 0
α lies in (0.73574, 1.5).

x	0.73573	1.5
$f(x)$	−0.00559	1.42926

$$\alpha_3 = \frac{0.73573 \times 1.42926 - 1.5(-0.00559)}{1.42926 - (-0.00559)} = 0.73871$$

Since f(0.73871) = −0.00061, we suspect that $\alpha_3 = 0.739$.

Checking,
f(0.7385) = $-0.00097 < 0$ and
f(0.7395) = $0.00069 > 0$

Hence, $\alpha = 0.739$. (3 d.p.)

When α is approximated to 0.739 (i.e $\alpha \approx 0.739$), the error of the approximation is ± 0.0005.

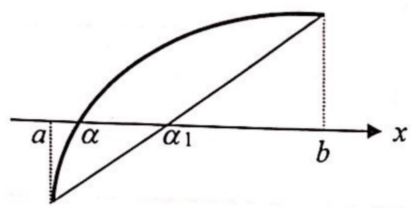
Hence α lies in the interval
$0.739 - 0.0005 \le \alpha < 0.739 + 0.0005$
ie. $0.7385 \le \alpha < 0.7395$

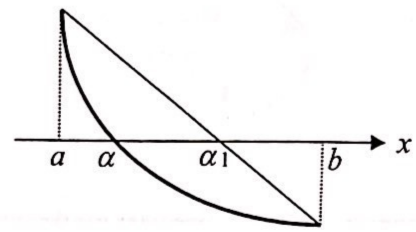
Always check the accuracy of the result.

4.2 Overestimation and Underestimation in Linear Interpolation

After obtaining a reasonable approximation (α_n) of α, the next question that one may ask is whether the approximation is overestimating or underestimating α. This depends on the shape of the graph.
Consider the four cases below.

Overestimation





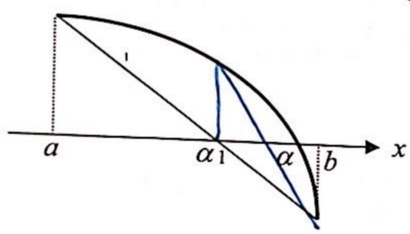
$f'(x) > 0$ and $f''(x) < 0 \Rightarrow \alpha_1 > \alpha$

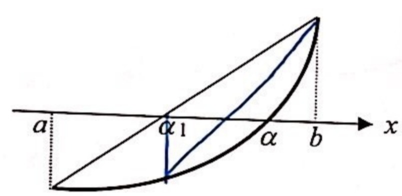
$f'(x) < 0$ and $f''(x) > 0 \Rightarrow \alpha_1 > \alpha$

⇕ A curve with $f''(x) < 0$ is said to be **concave downward.**

A curve with $f''(x) > 0$ is said to be **concave upward.** ⇕

Underestimation





$f'(x) < 0$ and $f''(x) < 0 \Rightarrow \alpha_1 < \alpha$

$f'(x) > 0$ and $f''(x) > 0 \Rightarrow \alpha_1 < \alpha$

Example 6 [9205/June 1991/01/Q13 (modified)]

Find the coordinates of the stationary points on the graph of $y = x^3 + x^2$. Sketch the graph and hence write down the set of values for the constant k for which the equation $x^3 + x^2 = k$ has three distinct roots. The positive root of the equation $x^3 + x^2 = 0.1$ is α.

(i) Find the integer N such that $N < \alpha < N+1$.

(ii) Use linear interpolation once, on the interval $[N, N+1]$, to find an approximation to α.

(iii) Explain by means of a sketch whether the first approximation overestimates or underestimates α.

(iv) With the aid of a suitable figure, indicate why, in this case, linear interpolation does not give a good approximation to α.

(v) Find an alternative first approximation to α by using the fact that if x is small, then x^3 is negligible, as compared to x^2.

Solution

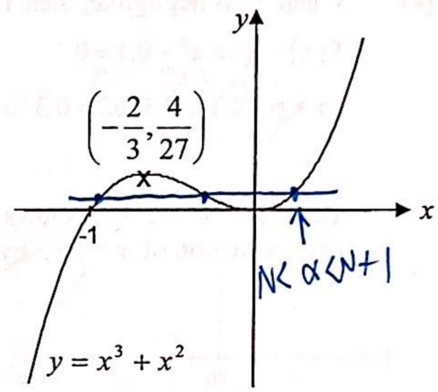
Now, $y = x^3 + x^2$

$\Rightarrow \dfrac{dy}{dx} = 3x^2 + 2x = 0 \Rightarrow x = 0$ or $x = -\dfrac{2}{3}$.

When $x = 0, y = 0$

When $x = -\dfrac{2}{3}, y = \dfrac{4}{27}$.

Hence, the coordinates of the stationary points are (0, 0) and $\left(-\dfrac{2}{3}, \dfrac{4}{27}\right)$.



From the sketch, it is clear that $0 < k < \dfrac{4}{27}$.

(i) Note that $0 < 0.1 < \dfrac{4}{27}$.

Let $f(x) = x^3 + x^2 - 0.1$.

$f(0) = 0 + 0 - 0.1 < 0$ and $f(1) = 1 + 1 - 0.1 = 1.9 > 0$

$\therefore 0 < \alpha < 1$ and $N = 0$

(ii) By linear interpolation on $[0,1]$, we have

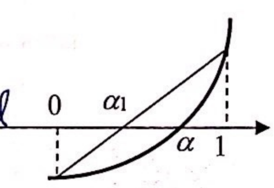
$$\alpha_1 = \frac{(0)f(1) - (1)f(0)}{f(1) - f(0)} = \frac{0.1}{1.9 + 0.1} = 0.05$$

(iii) Now, $f'(x) = 3x^2 + 2x > 0$ on $[0,1]$.

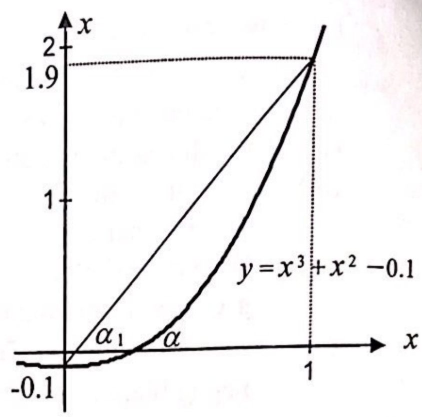
Also, $f''(x) = 6x + 2 > 0$ on $[0,1]$.

The curve is increasing & concave upward

$\therefore \alpha_1$ underestimates α



(iv) From the graph, we can see that the change in gradient of the curve in $[0,1]$ is too rapid. [In fact, $f'(0)=0$ and $f'(1)=5$]. The gradient increases slowly initially at $x=0$ and rapidly towards $x=1$. Hence, the chord with fixed gradient will intersect the x-axis far away from α, and thus linear interpolation on [0, 1] does not give a good approximation to α. #

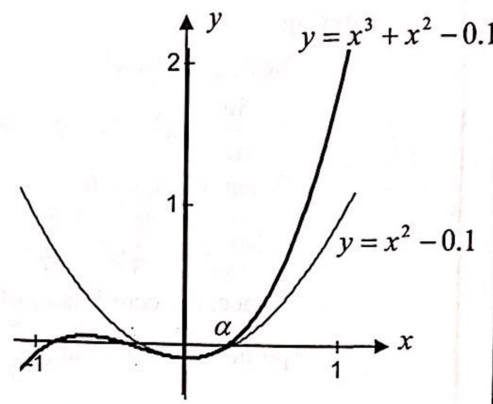


(v) When x^3 is negligible, then $f(x) \approx x^2 - 0.1$. Hence,

$f(x)=0 \Rightarrow x^2 - 0.1 \approx 0$

$\Rightarrow x \approx \sqrt{0.1} \approx 0.3162 = 0.316$ (3 s.f.)

The graph on the right shows the approximation of $y = f(x)$ by $y = x^2 - 0.1$.



5 NEWTON-RAPHSON METHOD (OR NEWTON'S METHOD)

When finding an approximate root of an equation by linear interpolation, we use a chord as a replacement for the curve near the root. In Newton-Raphson Method, we use the x-intercept of the tangent to the curve as an approximation to the root.

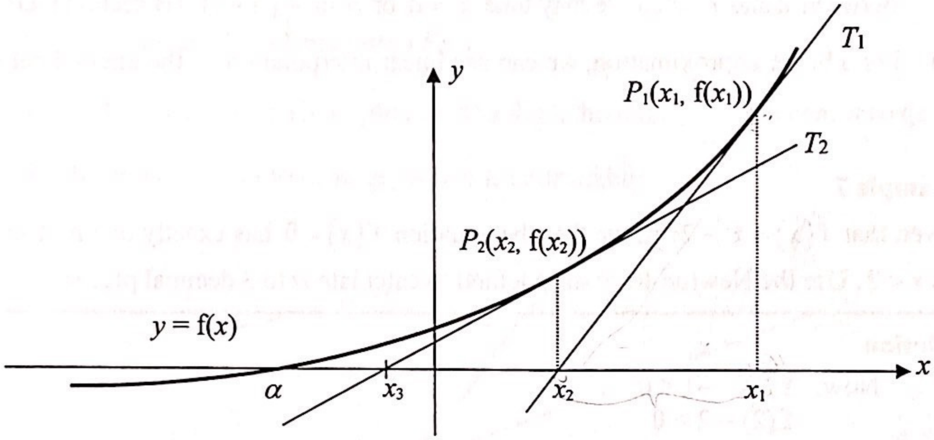
5.1 Derivation of the Newton-Raphson Method

Suppose we want to solve the equation $\mathrm{f}(x)=0$ where f is a continuous function.

We consider the following procedure and graph:

1. Pick a starting value x_1. This serves as the first approximation to the exact root α of the equation $\mathrm{f}(x)=0$.

2. Construct a tangent T_1 to the curve at the point $P_1(x_1, \mathrm{f}(x_1))$.



The gradient of the tangent T_1 at $P_1(x_1, \mathrm{f}(x_1))$ is given by $\mathrm{f}'(x_1)$.
The tangent T_1 intersects the x-axis at $(x_2, 0)$.

Obtain x_2, the second approximation of α, given by the x-intercept of T_1.
The equation of T_1 is given by

$$\frac{\mathrm{f}(x_1)-0}{x_1-x_2}=\mathrm{f}'(x_1) \Rightarrow x_2=x_1-\frac{\mathrm{f}(x_1)}{\mathrm{f}'(x_1)}.$$

3. Repeat the steps above to obtain x_3, given by $x_3=x_2-\dfrac{\mathrm{f}(x_2)}{\mathrm{f}'(x_2)}$.

4. Continue the process until the required accuracy is reached.

In general, the iterative formula is given by

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

In the above explanation, $x_1, x_2, x_3, \ldots$ is a sequence of approximations which converges to the root α of the equation $f(x) = 0$.

Question:
In order to apply Newton-Raphson method, we need to find an initial approximation to the root? How can this be done?

(a) Obtain a and b such that $f(a)$ and $f(b)$ are of opposite signs. Then we know that a root is between a and b, then we may take $x_1 = a$ or b or $\frac{1}{2}(a+b)$. (Bisection method)

(b) For a better approximation, we can use linear interpolation on the interval between a and b.

Example 7

Given that $f(x) = x^2 - 2$, prove that the equation $f(x) = 0$ has exactly one root α, in the interval $1 < x < 2$. Use the Newton-Raphson Method to calculate α to 3 decimal places.

Solution

Now, $f(1) = -1 < 0$
$f(2) = 2 > 0.$
And $f'(x) = 2x > 0$ on (1, 2).
f is continuous and strictly increasing over (1, 2).
Hence, the equation $f(x) = 0$ has exactly one root in (1, 2).

By Newton-Raphson Method, we have

$$X_{n+1} = X_n - \frac{f(X_n)}{f'(X_n)} = X_n - \frac{(X_n)^2 - 2}{2X_n}$$

Take the initial approximation to be $x_1 = 1.5$,

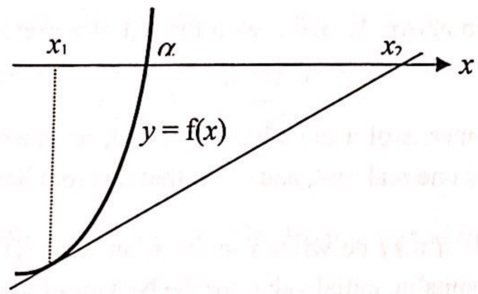
$\Rightarrow x_2 = 1.41667$
$\Rightarrow x_3 = 1.41422$
$\Rightarrow x_4 = 1.41421.$

Now, since $f(1.4135) = -0.00201775 < 0$ and
$f(1.4145) = 0.00081025 > 0$
Hence, $\alpha = 1.414$ (3dp)

5.2 Cases Where the Newton-Raphson Method May Fail

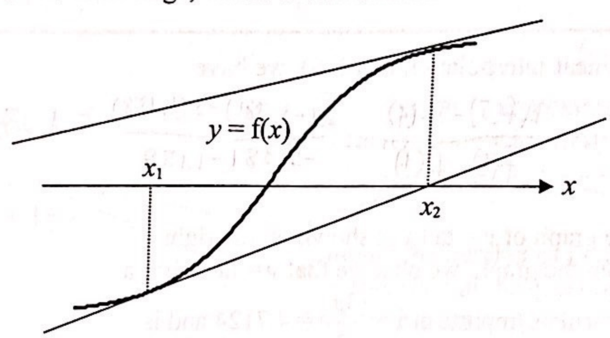
From Eg. 6(iv), we saw that linear interpolation may not give a good approximation to the root α of the equation $f(x) = 0$. There are also cases where the Newton-Raphson method may fail. Let us illustrate some of the common ones here.

Case 1: $f''(x_1)$ is too large. (Change in gradient of tangent is too rapid)



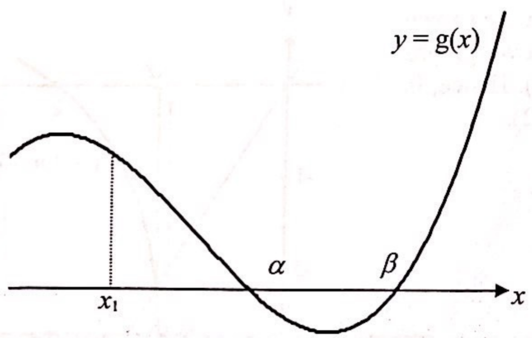
Case 2: $f'(x_1)$ is too small as compared to $f(x_1)$.

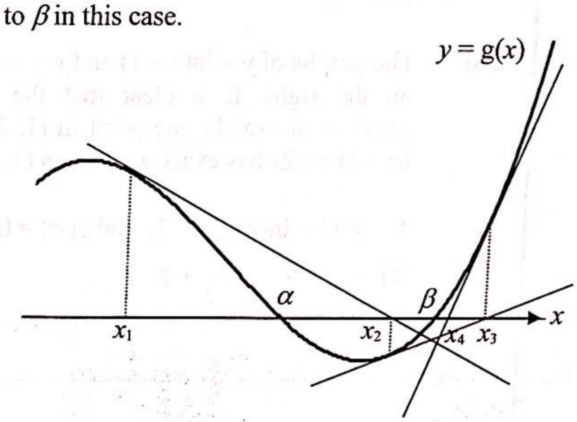
When $f'(x_1)$ is too small as compared to $f(x_1)$, the ratio $\dfrac{f(x_1)}{f'(x_1)}$ becomes large, causing the difference $x_2 - x_1$ to be large, which is undesirable.



Case 3: $|x_1 - \alpha|$ is too large. (Inappropriate choice of starting value)

The graph $y = g(x)$ below shows that if the first approximation x_1 is far away from α, then Newton-Raphson Method fails in the sense that the sequence $\{x_n\}$ does not converge to α. Instead, it converges to β in this case.





Example 8 [9205/1995/02/Q13]

(a) A function f is such that $f(4) = 1.158$ and $f(5) = -3.381$, correct to three decimal places in each case. Assuming that there exists a value of x between 4 and 5 for which $f(x) = 0$, use linear interpolation to estimate this value.

For the case when $f(x) = \tan x$, and x is measured in radians, the values of f(4) and f(5) are as given above. Explain, with the aid of a sketch, why linear interpolation using these values does not give an approximation to a solution of the equation $\tan x = 0$.

(b) Show, by means of a graphical argument, or otherwise, that the equation $-2x = \ln(x-1)$ has exactly one real root, and show that this root lies between 1 and 2.

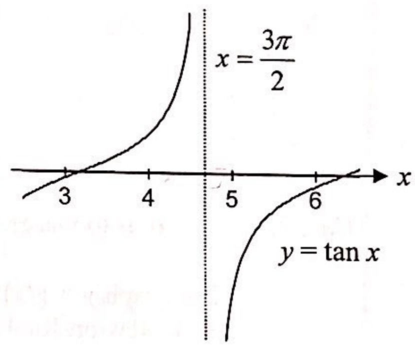
The equation may be written in the form $\ln(x-1) + 2x = 0$. Show that neither $x = 1$ nor $x = 2$ is a suitable initial value for the Newton-Raphson method.

The equation may also be written in the form $x - 1 - e^{-2x} = 0$. For this form, use two applications of the Newton-Raphson method, starting with $x_1 = 1$, to obtain an approximation to the root, giving three decimal places in your answer.

Solution

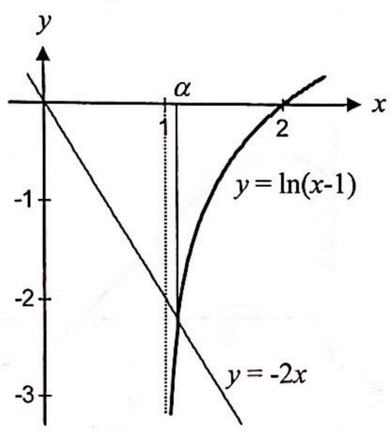
(a) By linear interpolation on (4, 5), we have

$$\alpha_1 = \frac{4f(5) - 5f(4)}{f(5) - f(4)} = \frac{4(-3.381) - 5(1.158)}{-3.381 - 1.158} = 4.255 \text{ (3dp)}$$



The graph of $y = \tan x$ is shown on the right. From the graph, we observe that $y = \tan x$ has a vertical asymptote at $x = \frac{3\pi}{2} = 4.7124$ and is therefore not continuous on (4, 5). In fact, there are no real roots in (4, 5). Hence, using linear interpolation on (4, 5) does not give an approximation to the solution of the equation $\tan x = 0$.

(b) The graphs of $y = \ln(x-1)$ and $y = -2x$ are shown on the right. It is clear that the two graphs intersect at exactly one point in (1, 2). Hence, $\ln(x-1) = -2x$ has exactly 1 root in (1, 2).



Let $g(x) = \ln(x-1) + 2x$ and $g(\alpha) = 0$.

Then $g'(x) = \frac{1}{x-1} + 2$.

By Newton-Raphson Method, we have

$$x_2 = x_1 - \frac{g(x_1)}{g'(x_1)} = 1 - \frac{g(1)}{g'(1)}$$

Since g(1) is not defined, $x_1 = 1$ is not a suitable initial value for the Newton-Raphson Method.

Using an initial value of $x_1 = 2$, we have

$$x_2 = 2 - \frac{g(2)}{g'(2)} = \frac{2}{3}$$

$$x_3 = \frac{2}{3} - \frac{g(\frac{2}{3})}{g'(\frac{2}{3})}$$

However, $g\left(\frac{2}{3}\right)$ is not defined. Hence, $x_1 = 2$ is also not a good initial value for Newton-Raphson Method.

Let $h(x) = x - 1 - e^{-2x}$. Then $h'(x) = 1 + 2e^{-2x}$.
By Newton-Raphson Method, we have

$$x_{n+1} = x_n - \frac{h(x_n)}{h'(x_n)} = x_n - \frac{x_n - 1 - e^{-2x}}{1 + 2e^{-2x}}$$

Initial approximation $x_1 = 1$,
$\Rightarrow \quad x_2 = 1.10650.$
$\Rightarrow \quad x_3 = 1.10885.$
Hence, $\alpha = 1.109$. (3 d.p.)

If the initial approximation is given, follow it strictly.

Should the number of iterations be specified in the question, follow it strictly. Also, no additional checking of accuracy is required.

6 Comparing Methods of Approximating Roots

We shall approximate the positive root for the equation $x^2-2=0$ using the various methods learnt.

Let $f(x)=x^2-2$.
Now, $f(1)=-1<0$
$f(2)=2>0$.
And $f'(x)=2x>0$ on $(1, 2)$.
f is continuous and strictly increasing over $(1, 2)$.
Hence, the equation $f(x)=0$ has exactly one root, α, in $(1, 2)$.

Let x_n denotes the n^{th} approximation of the root, α.

Method 1 : Applying bisection (BM) on the interval (1,2).
Method 2 : Applying linear interpolation (LI) on the interval (1,2).
Method 3 : Applying Newton-Raphson Method (NR) with $x_1=1.5$

n	BM (x_n)		LI (x_n)		NR (x_n)	
	x_n	% Error	x_n	% Error	x_n	% Error
1	1.5	6.07	1.3333333	5.72	1.5	6.07
2	1.25	11.6	1.4000000	1.01	1.4166667	1.73×10^{-1}
3	1.375	2.77	1.4117647	1.73×10^{-1}	1.4142157	1.51×10^{-4}
4	1.4375	1.65	1.4137931	2.97×10^{-2}	1.4142136	2.66×10^{-6}
5	1.40625	5.63×10^{-1}	1.4141414	5.10×10^{-3}		
6	1.421875	5.42×10^{-1}	1.4142012	8.74×10^{-4}		
7	1.4140625	1.07×10^{-2}	1.4142114	1.53×10^{-4}		
8	1.4179688	2.66×10^{-1}	1.4142132	2.56×10^{-5}		
9	1.4160157	1.27×10^{-1}				
10	1.4150391	5.84×10^{-2}				
11	1.4145508	2.38×10^{-2}				
12	1.4143067	6.59×10^{-3}				
13	1.4141846	2.05×10^{-3}				
14	1.4142457	2.27×10^{-3}				
15	1.4142152	1.16×10^{-4}				
16	1.4141999	9.66×10^{-4}				
17	1.4142076	4.22×10^{-4}				
18	1.4142114	1.53×10^{-4}				
19	1.4142133	1.86×10^{-5}				

In general, it can be observed that the percentage error when using Newton-Raphson Method decreases very fast as n increases.

SUMMARY