Motilal Nehru National Institute of Technology, Allahabad
NUMERICAl METHODS AND STATISTICAL TECHNIQUES
Unit: Solution Methods for algebraic and transcendental equations
Course Instructor: Dr. Prashanta Majee

1 Introduction

In scientific and engineering studies, a problem of grate importance is that of determining a root/zero of an equation of the form

$$f(x) = 0. \tag{1}$$

As a matter of fact, determination of any unknown appearing implicitly in scientific or engineering formulas, gives rise to a root finding problem. We consider one such simple application here. (state the ladder problemin mine shafts)

If $f(x)$ is a quadratic, cubic or a bi-quadratic equation, then algebraic formulas are available to express the roots in terms of the coefficients. For example, the roots for the quadratic equation $ax^2 + bx + c = 0$ can be expressed as

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

To solve cubic equations, one can use Cardon's method. To solve bi-quadratic equation, one can use Ferrari method.
Algebraic and transcendental equation: A polynomial equation of the form $f(x) = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n = 0$ is called algebraic equation.
An equation which contains polynomials, exponential functions, trigonometric functions, logarithmic functions or any combinations of them are called transcendental equations.
For examples, $x^2 + 5x + 7 = 0$, $x^7 - 5x^3 + x - 9 = 0$ etc. are algebraic equations, and, $xe^{2x} - 1 = 0$, $cosx - xe^x = 0$, $tanx = x$ etc. are transcendental equations.

If $f(x) = 0$ is a polynomial equation of degree more than four or $f(x)$ is transcendental equation, then no algebraic or analytic methods are available. One must have to apply some numerical methods to find the roots.

In this chapter, we discuss some methods for finding a real root of an algebraic or transcendental equations. We also discuss some methods to determine all real or complex roots of polynomial equation. Some method to find the solution of a system of nonlinear equations will also be discussed.

Root: A real or complex number α is called root of an equation $f(x) = 0$ if $f(\alpha) = 0$. Geometrically, the graph $y = f(x)$ cuts the x-axis at $x = \alpha$.

Now, the methods for finding the roots are classified as (i) direct methods, and (ii) iterative methods.
Direct Methods: These methods give the exact values of all the roots in a finite number

of steps (disregarding the round-off errors). Therefore, for any direct method, we can give the total number of operations (additions, subtractions, divisions and multiplications). This number is called the operational count of the method. For example, the roots of the quadratic equation $ax^2 + bx + c = 0$ can be expressed as

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

For this method, we can give the count of the total number of operations. There are direct methods for finding all the roots of cubic and bi-quadratic polynomial equations. However, these formulas are difficult to remember.. Direct methods for finding the roots of polynomial equations of degree greater than 4 or transcendental equations are not available in the literature.

Iterative Methods: These methods are based on the idea of successive approximations. We start with one or two initial approximations to the root and obtain a sequence of approximations $x_0, x_1, ..., x_k, ...$which in the limit as $k \to \infty$ converge to the exact root α. The iterative method for finding the root of the equation $f(x) = 0$ can be written as

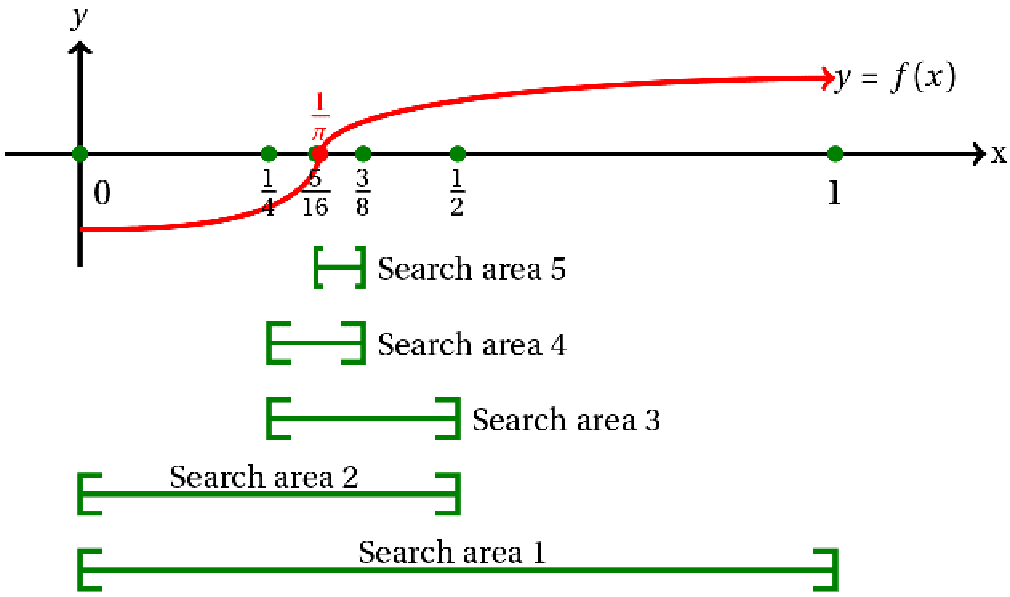
$$x_{k+1} = \phi(x_k),\ \ k = 0, 1, 2, ...$$

This method uses one initial approximation to the root. Suppose the initial approximation is x_0. Then the sequence of approximations are given by $x_1 = \phi(x_0),\ \ x_2 = \phi(x_1), x_3 = \phi(x_2)$ and so on.

In this unit, we shall study some iterative methods to find root of some algebraic and transcendental equations.

2 Bisection method:

This method is based on the repeated application of intermediate value theorem. Let the function $f(x)$ be continuous on $I_0 = [a, b]$. For definiteness, let $f(a)$ be $(-)ve$ and $f(b)$ be $(+)ve$. Then the first approximation to the root is $x_1 = \frac{a+b}{2}$. If $f(x_1) = 0$, then x_1 is a root of the equation $f(x) = 0$. Otherwise, the root lies between a and x_1 or x_1 and b according to $f(x_1)$ is $(+)ve$ or $(-)ve$. Then, we bisect this subinterval and the process is continued until the root is found to desired accuracy.



Basically, In each iteration, the length of the subinterval is reduced by half. Continuing the process, we obtain a sequence of nested subintervals $I_0 \supset I_1 \supset I_2 ... \supset I_n .....$ such that each of the subinterval contains the root. After repeating the bisection process n times, we obtain an interval I_n of length $\frac{b-a}{2^n}$, which contains the root.
Note that the method does not use the value $f(x)$, but only its sign. Hence, if an accuracy for the root is prescribed, the number of iteration can be determined in advance. For given accuracy ϵ, the number of iteration can be given by $\frac{b-a}{2^n} \le \epsilon \Rightarrow n \ge \frac{ln(b-a)-ln(\epsilon)}{log2}$. For example, if $[a, b] = [0, 1]$ and $\epsilon = 0.01$, then we get $n \ge 7$, that is, 7 iterations are required so that $|exact\ root - obtained\ root| < 0.01$.

The convergence of the bisection method is guaranteed. The main drawback to the bisection method is that this method is relatively to slow to converge as it may take many iterations before we can guarantee the desired accuracy.

Example 1 *Find a real root of the equation* $x^3 - 2x - 5 = 0$. *Perform four iterations of bisection methods.*

Answer: Let $f(x) = x^3 - 2x - 5$. Then $f(2) = -1$ and $f(3) = 16$. Hence, a root lies between 2 and 3 and we take

$$x_1 = \frac{2+3}{2} = 2.5$$

Since $f(x_1) = f(2.5) = 5.6250$, the root lies between 2 and 2.5
Hence,

$$x_2 = \frac{2+2.5}{2} = 2.25$$

Since $f(x_2) = f(2.25) = 1.8906$, the root lies between 2 and 2.25
Hence,

$$x_3 = \frac{2+2.25}{2} = 2.125$$

Since $f(x_3) = f(2.125) = 0.3457$, the root lies between 2 and 2.125

Hence,

$$x_4 = \frac{2 + 2.125}{2} = 2.0625$$

So, after four iterations, the approximate root is 2.0625

3 Regula-Falsi Method or Method of False Position

This is one of the oldest method for finding a real root of a nonlinear equation $f(x) = 0$ and closely resembles the bisection method. First we choose two points a and b such that $f(a)$ and $f(b)$ are of opposite signs. Hence, by the intermediate value theorem, a root must lie between a and b. Now, we find the equation of the chord joining the two points $(a, f(a))$ and $(b, f(b))$. The equation of the chord is

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

In this method, we replace the part of the curve between the points $(a, f(a))$ and $(b, f(b))$ by the chord joining these points. This chord will cut intersect x-axis at some point. We call that point our next approximate root. The point of intersect in the present case can be obtained by putting $y = 0$ in the equation of the chord. Thus, we obtain,

$$x_1 = a - \frac{b - a}{f(b) - f(a)} f(a), \tag{2}$$

which is the first approximate root of the equation $f(x) = 0$. Now, if $f(x_1)$ and $f(a)$ are of opposite sign, then we replace b by x_1 in (2), and obtain the next approximation. Otherwise, we replace a by x_1 (because in this case the root will be lying between a and x_1) and generate the next approximation. This process is repeated until the root is obtain to the desired accuracy. Graphical representation of the method is given bellow.

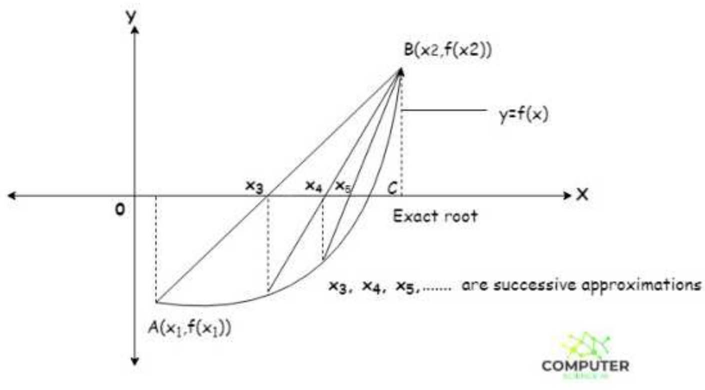


Fig. False Position / Regula-Falsi Method

- One can notice that in regula-falsi method, one of the end point of the initial interval (a, b) is always fixed and the other end point varies with n.
- Method require two initial approximations to the root.
- Convergence of the iterative method is guaranteed.

4 Secant Method

Secant method is similar to the regula-falsi method with some differences. In secant method, two initial approximation x_0 and x_1 to the root of the equation $f(x) = 0$ are chosen. Then the next iterative values are obtain by the following formula:

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

In regula-falsi method, after each iteration the position of the root was checked. But in secant method, once the two initial guesses x_0 and x_1 are chosen, we keep on applying the formula (3).

Remark 1 *Since, after each iteration, position of the root is not checked, for some initial guesses the method may converge and for some other initial guesses the method may diverge. Thus the convergence of the secant method is not guaranteed. But, if the secant method converges, the convergence of the secant method is better than that of regula-falsi method. Order of convergence for the regula-falsi method is 1, while the order of convergence of the secant method is 1.62.*

Example 2 *Use regula falsi method and secant method to obtain a root of the equation* $cosx - xe^x = 0$*. Take the initial approximations as* $x_0 = 0$ *and* $x_1 = 1$*. Perform three iterations in each method.*

Solution: **By Regula-Falsi Method:** Let $f(x) = cosx - xe^x$. then $f(0) = 1$ and $f(1) = cos1 - e = -2.177979$. Hence, a root lies between $x_0 = 0$ and $x_1 = 1$ and we take

$$x_2 = x_1 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_1) = 0.314665$$

Since $f(x_2) = f(0.314665) = 0.519871$, the root lies between x_2 and x_1.
Hence,

$$x_3 = x_2 - \frac{x_2 - x_1}{f(x_2) - f(x_1)} f(x_2) = 0.446728$$

Since $f(x_3) = f(0.446728) = 0.203544$, the root lies between x_3 and x_1.
Hence,

$$x_4 = x_3 - \frac{x_3 - x_1}{f(x_3) - f(x_1)} f(x_3) = 0.494015$$

So, after third iterations, the approximate root is 0.494015

By Secant Method: Let $f(x) = cosx - xe^x$. Then $f(0) = 1$ and $f(1) = cos1 - e = -2.177979$. Then

$$x_2 = x_1 - \frac{x_1 - x_0}{f(x_1) - f(x_0)} f(x_1) = 0.314665$$

and $f(x_2) = f(0.314665) = 0.519871$. So,

$$x_3 = x_2 - \frac{x_2 - x_1}{f(x_2) - f(x_1)} f(x_2) = 0.446728$$

and $f(x_3) = f(0.446728) = 0.203544$. So,

$$x_4 = x_3 - \frac{x_3 - x_2}{f(x_3) - f(x_2)} f(x_3) = 0.531705$$

So, after third iterations, the approximate root is 0.531705

5 Newton-Raphson Method:

Let x_0 be an initial approximation to the root of $f(x) = 0$. Then $(x_0, f(x_0))$ is a point on the curve $y = f(x)$. The equation of the tangent to the curve $y = f(x)$ at the point $(x_0, f(x_0))$ is

$$y - f(x_0) = f'(x_0)(x - x_0). \tag{4}$$

The point where it crosses x-axis is called the next approximate value x_1. Putting $y = 0$ in (4), we obtain

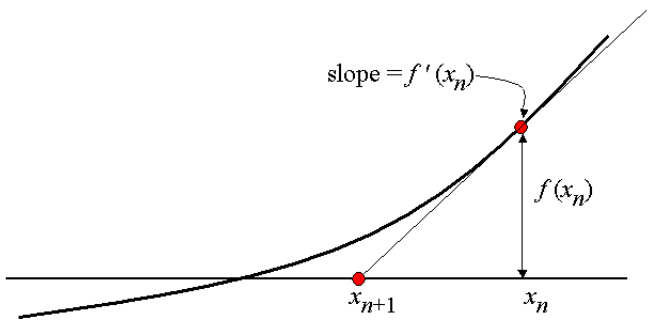
$$x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Thus,

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

In general, for n-th iteration,

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



Alternative proof: Let at n-th iteration, x_n be the approximate root of the equation $f(x) = 0$. Let h_n be the error in n-th iteration. Then $x_n + h_n$ is the exact root. So, $f(x_n + h_n) = 0$.

Then we have the followings:

$$
\begin{aligned}
& f(x_n + h_n) = 0 \\
\Longrightarrow\ & f(x_n) + h_n f^{'}(x_n) + \frac{{h_n}^2}{2} f^{''}(x_n) + ... = 0 \ \ (Using\ Taylor\ series\ expansion) \\
\Longrightarrow\ & f(x_n) + h_n f^{'}(x_n) \simeq 0 \ \ (Since\ h_n\ is\ small, neglecting\ higher\ order\ terms) \\
\Longrightarrow\ & h_n = -\frac{f(x_n)}{f^{'}(x_n)}.
\end{aligned}
$$

Since, higher order terms are discarded, we have not obtained the true value of h_n. Hence, $x_n + h_n$ is not the exact root, it is an approximate root, call it x_{n+1}.
Hence,

$$x_{n+1} = x_n - \frac{f(x_n)}{f^{'}(x_n)}, \ n = 0, 1, 2, 3, ...$$

Observation:
1. Newton-Raphson method requires one initial guess.
2. Cost of the method is one evaluation of $f(x)$ and one evaluation of $f^{'}(x)$ per iteration.
3. The method may fail if the initial guess if far from the exact root or at any iteration the derivative of the function is close to zero.
4. The order of convergence of the Newton-Raphson method is 2 (it will be discussed later).

Example 3 *Perform Newton-Raphson method to obtain the approximate value of* $(17)^{\frac{1}{3}}$ *with initial guess* $x_0 = 2$.

Solution: Let $x = (17)^{\frac{1}{3}}$. Then $x = (17)^{\frac{1}{3}} \Longrightarrow x^3 = 17 \Longrightarrow x^3 - 17 = 0$.
Take $f(x) = x^3 - 17$. Then $f^{'}(x) = 3x^2$.
Hence, by Newton-Raphson method, we obtain

$$x_{n+1} = x_n - \frac{{x_n}^3 - 17}{3{x_n}^2}, \ n = 0, 1, 2, 3, ...$$

Here, $x_1 = 2.75, \ x_2 = 2.582645, \ x_3 = 2.571332....$ etc.

6 General Iterative Method/ Successive Approximation Method:

We write the given equation $f(x) = 0$ as $x = \phi(x)$. We want to find a root of the equation in $I = [a, b]$. Choose $x_0 \in I$ and write an iterative method as

$$x_{n+1} = \phi(x_n) \ n = 0, 1, 2, 3, ...$$

Question:
(i) Does the sequence $\{x_n\}$ always converge to some root of $f(x) = 0$?
(ii) How to choose the function ϕ so that the sequence converge to a root of $f(x) = 0$?
The answers to the above questions are discussed bellow.
We can write the equation $f(x) = 0$ in the form $x = \phi(x)$ in several ways and can have several

iterative method of the form

$$x_{n+1} = \phi(x_n) \ n = 0, 1, 2, 3, ...$$

For example, the equation $x^3 - 5x + 1 = 0$, which has a root in $(0, 1)$. Consider the following representations:
(i) $x = \frac{x^3+1}{5}$ which will give successive approximation method

$$x_{n+1} = \frac{{x_n}^3 + 1}{5}, \ n = 0, 1, 2, 3, ...$$

Taking $x_0 = 1$, we get $x_1 = 0.4, \ x_2 = 0.2128, \ x_3 = 0.2019, ...$ which converges to the root in $(0, 1)$.

(ii) $x = (5x - 1)^{\frac{1}{3}}$ which will give successive approximation method

$$x_{n+1} = (5x_n - 1)^{\frac{1}{3}}, \ n = 0, 1, 2, 3, ...$$

Taking $x_0 = 1$, we get $x_1 = 1.5874, \ x_2 = 1.9072, \ x_3 = 2.0437, ...$ which is not converging to the root in $(0, 1)$.

(iii) $x = x^3 - 4x + 1$ which will give successive approximation method

$$x_{n+1} = {x_n}^3 - 4x_n + 1, \ n = 0, 1, 2, 3, ...$$

Taking $x_0 = 1$, we get $x_1 = -2, \ x_2 = 1, \ x_3 = -2, ...$ which is not converging to any root, as the sequence is oscillatory.

Hence, the successive approximation method depends on the suitable choice of the iteration function $\phi(x)$ and the initial approximation x_0. So the question is, how to choose $\phi(x)$ so that the iteration method converges to a root of $f(x) = 0$ for suitable initial value.
Condition of convergence: Let $f(x) = 0$ has a root in the interval $I = [a, b]$. We write $f(x) = 0$ as $x = \phi(x)$, where ϕ is a continuous function in the interval I. The iterative method is written as

$$x_{n+1} = \phi(x_n) \ n = 0, 1, 2, 3, ...$$

Let α be the exact root in I. Then $\alpha = \phi(\alpha)$.
Let the error in the n-th iteration be $\epsilon_n = \alpha - x_n$.

Hence,

$$
\begin{aligned}
&\alpha - x_{n+1} = \phi(\alpha) - \phi(x_n)\\
&\Rightarrow \alpha - x_{n+1} = (\alpha - x_n)\phi'(c_n),\ c_n \textit{ strictly lies in } \alpha \textit{ and } x_n \textit{ (by Lagrange mean value theorem)}\\
&\Rightarrow \alpha - x_{n+1} = (\alpha - x_{n-1})\phi'(c_n)\phi'(c_{n-1}),\ c_{n-1} \textit{ strictly lies in } \alpha \textit{ and } x_{n-1}\\
&\vdots\\
&\Rightarrow \alpha - x_{n+1} = (\alpha - x_0)\phi'(c_n)\phi'(c_{n-1})...\phi'(c_0),\ c_0 \textit{ strictly lies in } \alpha \textit{ and } x_0
\end{aligned}
$$

Let $|\phi'(c_i)| < K$ for $i = 0, 1, 2, ...n$. Thus,

$$|\alpha - x_{n+1}| = K^{n+1}|\alpha - x_0|.$$

The sequence $\{x_n\}$ will converge to the root α if the terms in the right hand side tends to zero as n tends to infinity. Since $|\alpha - x_0|$ is fixed quantity, right hand side tends to zero iff $K < 1$. Hence, the successive approximation method converges iff $|\phi'(x)| \le k < 1$ for all $x \in I$.

In the previous problem,
(i) for $\phi(x) = \frac{x^3+1}{5}$, we have $\phi'(x) = \frac{3x^2}{5} < 1$ for all $x \in [0, 1]$. Hence, the successive approximation method $x_{n+1} = \frac{{x_n}^3+1}{5}$ $n = 0, 1, 2, 3, ...$ converges to the root in $(0, 1)$.
(ii) for $\phi(x) = (5x-1)^{\frac{1}{3}}$, max value of $\phi'(x)$ in $[0, 1]$ is $\frac{5}{3}$. Hence, the successive approximation method $x_{n+1} = (5x_n - 1)^{\frac{1}{3}}$ $n = 0, 1, 2, 3, ...$ do not converge to the root in $(0, 1)$.
(iii) for $\phi = x^3 - 4x + 1$, max $\phi'(x)$ is greater than 1. Hence, the successive approximation method $x_{n+1} = {x_n}^3 - 4x_n + 1$ $n = 0, 1, 2, 3, ...$ do not converge to the root.

7 Order of convergence

An iterative method is said to have an order of convergence p if p is the largest positive real number for which there exists a finite positive real constant C such that

$$|\epsilon_{n+1}| \le C|\epsilon|^p,$$

or

$$|x_{n+1} - \alpha| \le C|x_n - \alpha|^p,$$

where α is the exact root of the equation $f(x) = 0$.
In order of convergence, we discuss how quickly the error is diminishing to zero. We consider the two scenarios:
Suppose in method I and method II, the errors in n-th iteration are bounded by, say, 0.001, i.e.,$|x_n - \alpha| \le 0.001$. In both the methods, we assume $C = 1$. Also, let us assume that method I has order of convergence 1, while method II has order of convergence 2.
We want to check the error bound in $(n + 1)$-th iteration.

In method I, error in $(n + 1)$-th iteration is bounded by $(0.001)^1 = 0.001$

In method I, error in $(n+2)$-th iteration is bounded by $(0.001)^1 = 0.001$

In method II, error in $(n+1)$-th iteration is bounded by $(0.001)^2 = 0.000001$
In method II, error in $(n+2)$-th iteration is bounded by $(0.000001)^2 = 10^{-12}$
Hence, the error in method II is diminishing to zero very quickly compare to method I.

7.1 Bisection Method

In bisection method, in each iteration the original interval is divided in to two sub intervals. If we take the mid-point of the successive intervals as the approximations of the root, then one half of the current interval is the upper bound to the error. So, in bisection method,

$$|x_{n+1} - \alpha| \leq \frac{1}{2}|x_n - \alpha|,$$

or

$$|\epsilon_{n+1}| \leq \frac{1}{2}|\epsilon_n|,$$

where ϵ_{n+1} and ϵ_n are the errors in the n-th and $(n+1)$-th approximations, respectively. Comparing with the definition of order of convergence

$$|\epsilon_{n+1}| \leq C|\epsilon_n|^p,$$

we obtain $C = 0.5$ and $p = 1$. Thus the bisection method has order of convergence 1.

7.2 Secant Method

Secant method for finding the root of $f(x) = 0$ is

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n), \quad n = 0, 1, 2, ... \tag{5}$$

Let α be the exact root i.e., $f(\alpha) = 0$ and ϵ_n be the error in the n-th approximation.
Hence, $x_{n-1} = \alpha + \epsilon_{n-1}$, $x_n = \alpha + \epsilon_n$ and $x_{n+1} = \alpha + \epsilon_{n+1}$. Putting these values in (5), we obtain,

$$\begin{aligned}\alpha + \epsilon_{n+1} &= \alpha + \epsilon_n - \frac{\alpha + \epsilon_n - (\alpha + \epsilon_{n-1})}{f(\alpha + \epsilon_n) - f(\alpha + \epsilon_{n-1})} f(\alpha + \epsilon_n) \\ \Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1}) f(\alpha + \epsilon_n)}{f(\alpha + \epsilon_n) - f(\alpha + \epsilon_{n-1})}\end{aligned}$$

Using the Taylor series expansion, we obtain

$$
\begin{aligned}
&\epsilon_{n+1} = \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1}) f(\alpha + \epsilon_n)}{f(\alpha + \epsilon_n) - f(\alpha + \epsilon_{n-1})} \\
&\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1})[f(\alpha) + \epsilon_n f'(\alpha) + \frac{\epsilon_n{}^2}{2} f''(\alpha) + ...]}{[f(\alpha) + \epsilon_n f'(\alpha) + \frac{\epsilon_n{}^2}{2} f''(\alpha) + ...)] - [f(\alpha) + \epsilon_{n-1} f'(\alpha) + \frac{\epsilon_{n-1}{}^2}{2} f''(\alpha) + ...)]} \\
&\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{(\epsilon_n - \epsilon_{n-1}) f'(\alpha)[\epsilon_n + \epsilon_n{}^2 \frac{f'(\alpha)}{2f''(\alpha)} + ...]}{f'(\alpha)[(\epsilon_n - \epsilon_{n-1}) + (\epsilon_n{}^2 - \epsilon_{n-1}{}^2) \frac{f''(\alpha)}{2f'(\alpha)} + ...)]} \\
&\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{(\epsilon_n + \epsilon_n{}^2 \frac{f'(\alpha)}{2f''(\alpha)} + ...)}{[1 + (\epsilon_n + \epsilon_{n-1}) \frac{f''(\alpha)}{2f'(\alpha)} + ...]} \\
&\Rightarrow \epsilon_{n+1} = \epsilon_n - \frac{(\epsilon_n + A\epsilon_n{}^2 + ...)}{[1 + A(\epsilon_n + \epsilon_{n-1}) + ...]}, \ \textit{where} \ \frac{f'(\alpha)}{2f''(\alpha)} \\
&\Rightarrow \epsilon_{n+1} = \epsilon_n - (\epsilon_n + A\epsilon_n{}^2 + ...)(1 - A(\epsilon_n + \epsilon_{n-1}) + ...) \\
&\Rightarrow \epsilon_{n+1} \simeq A\epsilon_n \epsilon_{n-1} \ (\textit{neglecting the higher order terms})
\end{aligned} \tag{6}
$$

Let $\epsilon_{n+1} = C\epsilon_n{}^p \Rightarrow \epsilon_n = C\epsilon_{n-1}{}^p \Rightarrow \epsilon_{n-1} = C^{-\frac{1}{p}} \epsilon_n{}^{\frac{1}{p}}$.
Hence, from (6), we get $C\epsilon_n{}^p = A\epsilon_n C^{-\frac{1}{p}} \epsilon_n{}^{\frac{1}{p}}$. Comparing the power of ϵ_n, we obtain $p = 1 + \frac{1}{p} \Rightarrow p^2 - p - 1 = 0$. Which gives, $p = \frac{1 \pm \sqrt{5}}{2}$. Since p is positive, hence $p = 1.618$. Thus, the order of convergence for the secant method is 1.618. (super linear order of convergence).

7.3 Regula-Falsi Method

We have earlier noticed that in regula-falsi method, one of the end point of the initial guesses x_0 and x_1 of the interval $[x_0, x_1]$ is always fixed and the other end point varies with n. With out loss of generality, let us assume that x_0 is fixed. Then the formula for regula-falsi method for finding the root of $f(x) = 0$ is

$$
x_{n+1} = x_n - \frac{x_n - x_0}{f(x_n) - f(x_0)} f(x_n), \quad n = 0, 1, 2, ... \tag{7}
$$

Let α be the exact root i.e., $f(\alpha) = 0$ and ϵ_n be the error in the n-th approximation.
Hence, $x_0 = \alpha + \epsilon_0$, $x_n = \alpha + \epsilon_n$ and $x_{n+1} = \alpha + \epsilon_{n+1}$. Putting these values in (5), we obtain,

$$
\begin{aligned}
\alpha + \epsilon_{n+1} &= \alpha + \epsilon_n - \frac{\alpha + \epsilon_n - (\alpha + \epsilon_0)}{f(\alpha + \epsilon_n) - f(\alpha + \epsilon_0)} f(\alpha + \epsilon_n) \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n - \epsilon_0) f(\alpha + \epsilon_n)}{f(\alpha + \epsilon_n) - f(\alpha + \epsilon_0)}
\end{aligned}
$$

Using the Taylor series expansion, we obtain

$$
\begin{aligned}
\epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n - \epsilon_0) f(\alpha + \epsilon_n)}{f(\alpha + \epsilon_n) - f(\alpha + \epsilon_0)} \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n - \epsilon_0)[f(\alpha) + \epsilon_n f^{'}(\alpha) + \frac{{\epsilon_n}^2}{2} f^{''}(\alpha) + ...]}{[f(\alpha) + \epsilon_n f^{'}(\alpha) + \frac{{\epsilon_n}^2}{2} f^{''}(\alpha) + ...)] - [f(\alpha) + \epsilon_0 f^{'}(\alpha) + \frac{{\epsilon_0}^2}{2} f^{''}(\alpha) + ...)]} \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n - \epsilon_0) f^{'}(\alpha)(\epsilon_n + {\epsilon_n}^2 \frac{f^{'}(\alpha)}{2f^{''}(\alpha)} + ...)}{f^{'}(\alpha)[(\epsilon_n - \epsilon_0) + ({\epsilon_n}^2 - {\epsilon_0}^2)\frac{f^{''}(\alpha)}{2f^{'}(\alpha)} + ...]} \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n + {\epsilon_n}^2 \frac{f^{'}(\alpha)}{2f^{''}(\alpha)} + ...)}{[1 + (\epsilon_n + \epsilon_0)\frac{f^{''}(\alpha)}{2f^{'}(\alpha)} + ...]} \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{(\epsilon_n + A{\epsilon_n}^2 + ...)}{[1 + A(\epsilon_n + \epsilon_0) + ...]} \ where \ \frac{f^{'}(\alpha)}{2f^{''}(\alpha)} \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - (\epsilon_n + A{\epsilon_n}^2 + ...)(1 - A(\epsilon_n + \epsilon_0) + ...) \\
\Rightarrow \epsilon_{n+1} &\simeq A\epsilon_n\epsilon_0 \ (neglecting\ the\ higher\ order\ terms) \\
\Rightarrow \epsilon_{n+1} &\simeq K\epsilon_n \ (where\ K = A\epsilon_0) \qquad (8)
\end{aligned}
$$

Hence, the order of convergence of the regula falsi method is 1.

7.4 Newton-Raphson Method

Newton-Raphson method for finding the root of $f(x) = 0$ is

$$x_{n+1} = x_n - \frac{f(x_n)}{f^{'}(x_n)}, \quad n = 0, 1, 2, ... \qquad (9)$$

Let α be the exact root and ϵ_n be the error in the n-th approximation.
Then, $x_n = \alpha + \epsilon_n$ and $x_{n+1} = \alpha + \epsilon_{n+1}$. Putting these values in (9), we obtain,

$$\begin{aligned}
\alpha + \epsilon_{n+1} &= \alpha + \epsilon_n - \frac{f(\alpha+\epsilon_n)}{f'(\alpha+\epsilon_n)} \\
\Rightarrow \epsilon_{n+1} &= \epsilon_n - \frac{f(\alpha+\epsilon_n)}{f'(\alpha+\epsilon_n)} \\
&= \epsilon_n - \frac{f(\alpha) + \epsilon_n f'(\alpha) + \frac{{\epsilon_n}^2}{2} f''(\alpha) + ...}{f'(\alpha) + \epsilon_n f''(\alpha) + \frac{{\epsilon_n}^2}{2} f'''(\alpha) + ...} \\
&= \epsilon_n - \frac{\epsilon_n + {\epsilon_n}^2 \frac{f''(\alpha)}{2f'(\alpha)} + ...}{1 + \epsilon_n \frac{f''(\alpha)}{f'(\alpha) + ...}} \\
&= \epsilon_n - \frac{\epsilon_n + A{\epsilon_n}^2 + ...}{1 + 2A\epsilon_n + ...}, \quad where\ A = \frac{f''(\alpha)}{2f'(\alpha)} \\
&= \epsilon_n - (\epsilon_n + A{\epsilon_n}^2 + ...)(1 + 2A\epsilon_n + ...)^{-1} \\
&= \epsilon_n - (\epsilon_n + A{\epsilon_n}^2 + ...)(1 - 2A\epsilon_n + ...) \\
&= \epsilon_n - \epsilon_n + 2A{\epsilon_n}^2 - A{\epsilon_n}^2 - 2A^2{\epsilon_n}^3 + ... \\
&\simeq A{\epsilon_n}^2 \ (neglecting\ the\ higher\ order\ terms)
\end{aligned}$$

Hence, the order of convergence of Newton-Raphson method is 2.

7.5 Successive Approximation Method

Suppose $f(x) = 0$ has a root in the interval $I = [a, b]$. We write $f(x) = 0$ as $x = \phi(x)$, where ϕ is a continuous function in the interval I. The iterative method is written as

$$x_{n+1} = \phi(x_n) \ n = 0, 1, 2, 3, ...$$

Let α be the exact root in I. Then $\alpha = \phi(\alpha)$.
Let the error in the n-th iteration be $\epsilon_n = \alpha - x_n$.
Hence,

$$\alpha - x_{n+1} = \phi(\alpha) - \phi(x_n)$$

$$\Rightarrow \alpha - x_{n+1} = (\alpha - x_n)\phi'(c_n),$$

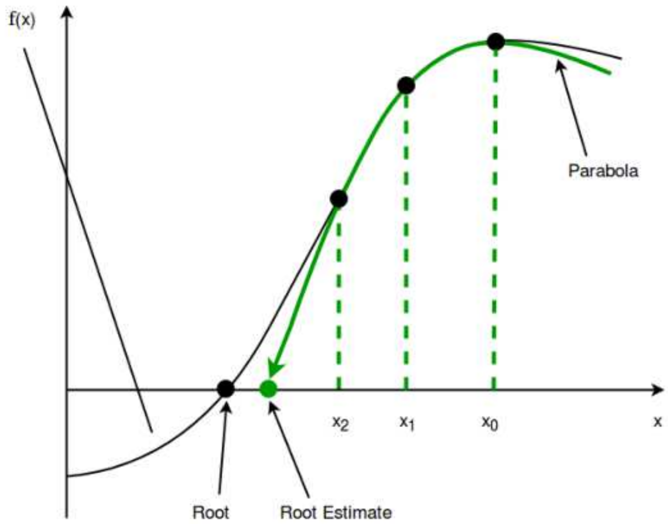
(by Lagrange's mean value theorem) c_n strictly lies in α and x_n. Thus, if $\phi'(c_n) \neq 0$, then

$$|\alpha - x_{n+1}| = |\alpha - x_n||\phi'(c_n)| \Rightarrow |\alpha - x_{n+1}| \leq |\alpha - x_n| \ (as|\phi'(c_n)| < 1).$$

Hence, if $\phi'(c_n) \neq 0$, the order of convergence for successive approximation method is 1.

8 Muller's Method

Most of the root finding methods (regula-falsi method, secant method, Newton-Raphson method) we have discussed so far have approximated the function $y = f(x)$ in the neighborhood of a root by a straight line.
Muller's method is based on approximating the function in the neighborhood of a root by a quadratic polynomial. This gives much better match to the actual curve than by a straight line. This method converges almost quadratically and can be used to obtain complex roots.



Let $x_{k-2},\ x_{k-1}$ and x_k be three approximations of the root α of $f(x) = 0$. Let $y_{k-2},\ y_{k-1}$ and y_k be the corresponding functional values of $y = f(x)$.
Assume that $P(x) = a(x - x_k)^2 + b(x - x_k) + c$ is the parabola passing through the points $(x_{k-2}, y_{k-2}), (x_{k-1}, y_{k-1})$ and (x_k, y_k). Then, we have

$$\begin{aligned}
y_k &= c \\
y_{k-1} &= a(x_{k-1} - x_k)^2 + b(x_{k-1} - x_k) + c \\
y_{k-2} &= a(x_{k-2} - x_k)^2 + b(x_{k-2} - x_k) + c.
\end{aligned}$$

Solving for a, b, c from the above equations, we obtain

$$\begin{aligned}
a &= \frac{(x_k - x_{k-2})(y_k - y_{k-1}) - (x_k - x_{k-1})(y_k - y_{k-2})}{(x_k - x_{k-1})(x_k - x_{k-2})(x_{k-1} - x_{k-2})} \\
b &= \frac{(x_k - x_{k-2})^2(y_k - y_{k-1}) - (x_k - x_{k-1})^2(y_k - y_{k-2})}{(x_k - x_{k-1})(x_k - x_{k-2})(x_{k-1} - x_{k-2})} \\
c &= y_k.
\end{aligned}$$

If $h_k = x_k - x_{k-1}$ and $\triangle y_k = y_k - y_{k-1}$, then after little manipulation a, b, c can be written as

$$\begin{aligned}
a &= \frac{(x_k - x_{k-2})(y_k - y_{k-1}) - (x_k - x_{k-1})(y_k - y_{k-2})}{(x_k - x_{k-1})(x_k - x_{k-2})(x_{k-1} - x_{k-2})} \\
&= \frac{(h_k + h_{k-1}) \triangle y_k - h_k(\triangle y_k + \triangle y_{k-1})}{(h_k + h_{k-1})h_k h_{k-1}} \\
&= \frac{h_{k-1} \triangle y_k - h_k \triangle y_{k-1}}{(h_k + h_{k-1})h_k h_{k-1}} \\
&= \frac{1}{(h_k + h_{k-1})}\Big(\frac{\triangle y_k}{h_k} - \frac{\triangle y_{k-1}}{h_{k-1}}\Big)
\end{aligned}$$

and

$$b = \frac{\triangle y_k}{h_k} + a h_k$$

and

$$c = y_k.$$

Any one of the above formulas can be used to evaluate a, b and c. Now, to find the next iterative value x_{k+1} we solve

$$a(x - x_k)^2 + b(x - x_k) + c = 0.$$

Solving, we obtain

$$x - x_k = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Thus,

$$x_{k+1} = x_k + \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

To avoid the computational error, we use

$$x_{k+1} = x_k + \frac{(-b \pm \sqrt{b^2 - 4ac})}{2a} * \frac{(-b \mp \sqrt{b^2 - 4ac})}{(-b \mp \sqrt{b^2 - 4ac})},$$

which gives

$$x_{k+1} = x_k - \frac{2c}{b \pm \sqrt{b^2 - 4ac}}.$$

The sign of the denominator should be chosen in such a way that the denominator is largest in magnitude.

Example 4 *Using Muller's method, find the root of the equation $x^3 - 2x - 5 = 0$, which lies* *between 2 and 3. Perform two iterations of Muller's method.*

Solution: Let $f(x) = x^3 - 2x - 5$
Iteration 1: Let $x_0 = 1$, $x_1 = 2$, $x_2 = 3$. Then $y_0 = -6$, $y_1 = -1$, $y_2 = 16$.
Let the parabola be $P(x) = a(x-x_k)^2 + b(x-x_k) + c$, which passes through the points $(1, -6)$, $(2, -1)$ and $(3, 16)$.
Then,

$$a = \frac{(x_2 - x_0)(y_2 - y_1) - (x_2 - x_1)(y_2 - y_0)}{(x_2 - x_1)(x_2 - x_0)(x_1 - x_0)} = 6$$

$$b = \frac{(x_2 - x_0)^2(y_2 - y_1) - (x_2 - x_1)^2(y_2 - y_0)}{(x_2 - x_1)(x_2 - x_0)(x_1 - x_0)} = 23$$

and

$$c = 16$$

Hence, the next iterative value is

$$x_3 = x_2 - \frac{2c}{b \pm \sqrt{b^2 - 4ac}} = 3 - \frac{2 * 16}{23 \pm \sqrt{23^2 - 4 * 6 * 16}} = 2.0868$$

Here, $f(x_3) = -0.0861$
Iteration 2: Let the parabola be $P(x) = a(x - x_k)^2 + b(x - x_k) + c$, which passes through the points $(2, -1)$, $(3, 16)$ and $(2.0868, -0.0861)$.
Then,

$$\begin{aligned} a &= \frac{(x_3 - x_1)(y_3 - y_2) - (x_3 - x_2)(y_3 - y_1)}{(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)} \\ &= \frac{(2.0868 - 2)(-0.0861 - 16) - (2.0868 - 3)(-0.0861 + 1)}{(2.0868 - 3)(2.0868 - 2)(3 - 2)} = 5.1024 \end{aligned}$$

$$\begin{aligned} b &= \frac{(x_3 - x_1)^2(y_3 - y_2) - (x_3 - x_2)^2(y_3 - y_1)}{(x_3 - x_2)(x_3 - x_1)(x_2 - x_1)} \\ &= \frac{(2.0868 - 2)^2(-0.0861 - 16) - (2.0868 - 3)^2(-0.0861 + 1)}{(2.0868 - 3)(2.0868 - 2)(3 - 2)} = 12.9556 \end{aligned}$$

and

$$c = -0.0861$$

Hence, the next iterative value is

$$\begin{aligned}
x_4 &= x_3 - \frac{2c}{b \pm \sqrt{b^2 - 4ac}} \\
&= 2.0868 - \frac{2 * (-0.0861)}{12.9556 \pm \sqrt{12.9556^2 - 4 * 5.1024 * (-0.0861)}} \\
&= 2.0970
\end{aligned}$$

Thus, after second iteration, the approximate root is 2.0970

9 Lin-Bairstow Method/Bairstow Method

Lin-Bairstow method is useful in determining a quadratic factor of a polynomial. Using this method, we can extract a quadratic factor of the form $x^2 + px + q$ from the polynomial $P_n(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0,\ (a_n \neq 0)$. Thus, $P_n(x)$ can be written as

$$P_n(x) = (x^2 + px + q) * Q_{n-2}(x),$$

where $Q_{n-2}(x)$ is a polynomial of degree $n-2$, say, $Q_{n-2}(x) = b_{n-1}x^{n-2} + b_{n-3}x^{n-3} + ... + b_1 x + b_0,\ (b_{n-2} \neq 0)$. So, if we want to find roots of the equation $P_n(x) = 0$, we have $x^2 + px + q = 0$ and $Q_{n-2}(x) = 0$. The equation $x^2 + px + q = 0$ can be easily solved. To solve $Q_{n-2}(x) = 0$, we will again extract a quadratic factor from $Q_{n-2}(x)$. We will keep on repeating the process until at the end we get a quadratic polynomial or a linear polynomial. So our main aim is to learn the method to extract quadratic factor from a given polynomial.

9.1 Method to extract a quadratic factor from a polynomial:

If we divide the polynomial $P_n(x)$ by a quadratic polynomial $x^2 + rx + s = 0$, then we obtain a quotient polynomial $Q_{n-2}(x)$ of degree $n - 2$ and reminder which is a polynomial of degree 1, that is,

$$P_n(x) = (x^2 + rx + s) * Q_{n-2}(x) + (Cx + D).$$

We need to find r and s in such a way that $C = 0$ and $D = 0$. Then $(x^2 + rx + s)$ will be a factor of $P_n(x)$.

Let us discuss the method of extracting quadratic factor from a cubic polynomial. Let the polynomial be

$$P_3(x) = A_3 x^3 + A_2 x^2 + A_1 x + A_0. \tag{10}$$

Suppose, $x^2 + Rx + S$ be the exact factor of $P_3(x)$ and $x^2 + rx + s$ be an approximate factor. Initially, r and s is obtained from the last three terms of the given polynomial as $r = \frac{A_1}{A_2}$ and $s = \frac{A_1}{A_2}$ (You can make any other guesses, there is no mathematical logic). So, we have

$$P_3(x) = (x^2 + rx + s)(B_1 x + B_0) + Cx + D. \tag{11}$$

Comparing (10) and (11), we obtain

$$\begin{aligned} A_3 &= B_1 \\ A_2 &= rB_1 + B_0 \\ A_1 &= C + rB_0 + sB_1 \\ A_0 &= D + sB_0 \end{aligned}$$

Using forward calculations, we need to find the values of C and D. Doing so, we obtain

$$\begin{aligned} C &= A_1 - r(A_2 - rA_3) - sA_3 \\ D &= A_0 - s(A_2 - rA_3). \end{aligned}$$

Since A_0, A_1, A_2 and A_3 are known, C and D are functions of r and s, i.e., $C = C(r, s)$ and $D = D(r, s)$.
Since $x^2 + Rx + S$ is the exact factor of the given polynomial, We have

$$C(R, S) = 0 \quad \text{and} \quad D(R, S) = 0. \tag{12}$$

We will create an iterative method and convert the approximate factor x^2+rx+s to the exact factor $x^2 + Rx + S$. Let $R = r + \triangle r$ and $S = s + \triangle s$. Putting these values in (12), we obtain

$$C(r + \triangle r, s + \triangle s) = 0 \quad \text{and} \quad D(r + \triangle r, s + \triangle s) = 0.$$

Now, expanding in Taylor series and discarding the higher terms, we get

$$\begin{aligned} C(r, s) + \triangle r\frac{\partial C}{\partial r} + \triangle s\frac{\partial C}{\partial s} &\simeq 0 \\ D(r, s) + \triangle r\frac{\partial D}{\partial r} + \triangle s\frac{\partial D}{\partial s} &\simeq 0, \end{aligned}$$

where the C, D and the derivatives are to be computed at the point (r, s). Solve the above equations to obtain the values of $\triangle r$ and $\triangle s$. Since, we have discarded the higher order terms, we will not get the exact values of $\triangle r$ and $\triangle s$, instead we will obtain some approximate values. So, $x^2 + (r + \triangle r)x + s + \triangle s$ will not be the exact factor, it will be an approximate factor for $P_3(x)$. The process will be repeated until successive values of R and S shows no significant changes.

Note: If we want to find a quadratic factor from a n-th degree polynomial

$$P_n(x) = A_nx^n + ... + A_2x^2 + A_1x + A_0,$$

consider

$$P_n(x) = (x^2 + rx + s)(B_{n-2}x^{n-2} + B_{n-3}x^{n-3} + ... + B_1x + B_0) + Cx + D,$$

and express C and D as functions of r and s. Then, we continue the calculations as above.

Example 5 *Find the quadratic factor of the polynomial given by $x^3 - 2x^2 + x - 2$.*

Solution: We have $A_3 = 1$, $A_2 = -2$, $A_1 = 1$ and $A_0 = -2$. Let the initial guesses be $r = \frac{-1}{2} = -0.5$ and $s = \frac{-2}{-2} = 1$. Let $x^2 + rx + s$ be an approximate factor of the given polynomial. Then,

$$x^3 - 2x^2 + x - 2 = (x^2 + rx + s)(B_1 x + B_0) + Cx + D.$$

Comparing the coefficient of the like powers, we get

$$\begin{aligned} &B_1 = 1, && B_0 = -2 - r \\ &C = 1 + 2r + r^2 - s, && D = -2 + 2s + rs. \end{aligned}$$

Here,

$$\frac{\partial C}{\partial r} = 2 + 2r, \quad \frac{\partial C}{\partial s} = -1, \quad \frac{\partial D}{\partial r} = s, \quad \frac{\partial D}{\partial s} = 2 + r.$$

Iteration 1: Let $x^2 + (r + \triangle r)x + (s + \triangle s)$ be the next approximate factor. To find $\triangle r$ and $\triangle s$, we need to solve

$$\begin{aligned} C(r,s) + \triangle r \frac{\partial C}{\partial r} + \triangle s \frac{\partial C}{\partial s} &= 0 \\ D(r,s) + \triangle r \frac{\partial D}{\partial r} + \triangle s \frac{\partial D}{\partial s} &= 0, \end{aligned}$$

Here, C, D and the derivatives are to be computed at the point $(-0.5, 1)$. So,

$$\begin{aligned} &C(r,s)\Big|_{(-0.5,1)} = -0.75, && D(r,s)\Big|_{(-0.5,1)} = -0.5, \\ &\frac{\partial C}{\partial r}\Big|_{(-0.5,1)} = (2+2r)\Big|_{(-0.5,1)} = 1, && \frac{\partial C}{\partial s}\Big|_{(-0.5,1)} = -1 \\ &\frac{\partial D}{\partial r}\Big|_{(-0.5,1)} = s\Big|_{(-0.5,1)} = 1, && \frac{\partial D}{\partial s}\Big|_{(-0.5,1)} = (2+r)\Big|_{(-0.5,1)} = 1.5 \end{aligned}$$

Hence, we get

$$\begin{aligned} \triangle r - \triangle s &= 0.75 \\ \triangle r + 1.5 \triangle s &= 0.5 \end{aligned}$$

Solving we get,

$$\triangle r = 0.65, \qquad \triangle s = -0.1$$

Hence, next approximate factor is
$x^2 + (r + \triangle r)x + (s + \triangle s) = x^2 + (-0.5 + 0.65)x + (1 - 0.1) = x^2 + 0.15x + 0.9$

Iteration 2: Let the approximate factor be $x^2 + 0.15x + 0.9$. Here, $r = 0.15$ and $s = 0.9$. So,

$$C(r,s)\Big|_{(0.15,0.9)} = 0.4225, \qquad D(r,s)\Big|_{(0.15,0.9)} = -0.065,$$
$$\frac{\partial C}{\partial r}\Big|_{(0.15,0.9)} = (2+2r)\Big|_{(0.15,0.9)} = 2.30, \qquad \frac{\partial C}{\partial s}\Big|_{(0.15,0.9)} = -1$$
$$\frac{\partial D}{\partial r}\Big|_{(0.15,0.9)} = s\Big|_{(0.15,0.9)} = 0.9, \qquad \frac{\partial D}{\partial s}\Big|_{(0.15,0.9)} = (2+r)\Big|_{(0.15,0.9)} = 2.15$$

Hence, we get

$$2.30\triangle r - \triangle s = -0.4225$$
$$.9\triangle r + 2.15\triangle s = -0.65$$

Solving we get,

$$\triangle r = -0.1665, \qquad \triangle s = 0.0394$$

Hence, next approximate factor is
$x^2 + (r + \triangle r)x + (s + \triangle s) = x^2 + (0.15 - 0.1665)x + (0.9 + 0.0394) = x^2 - 0.0165x + 0.9394$.
Actual factor is $x^2 + 1$. So, we see that after second iterations, our obtained answer is very close to the actual quadratic factor.