Numerical Methods KOM2722- AVE3842

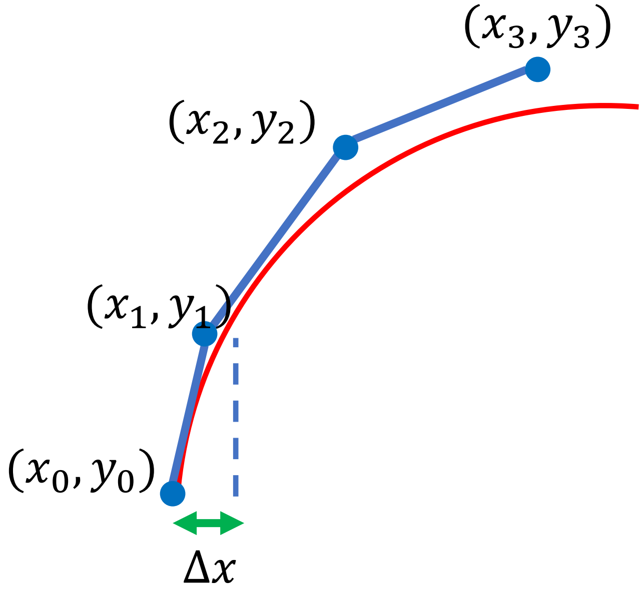
Week 4

Asst. Prof. Bilal Erol

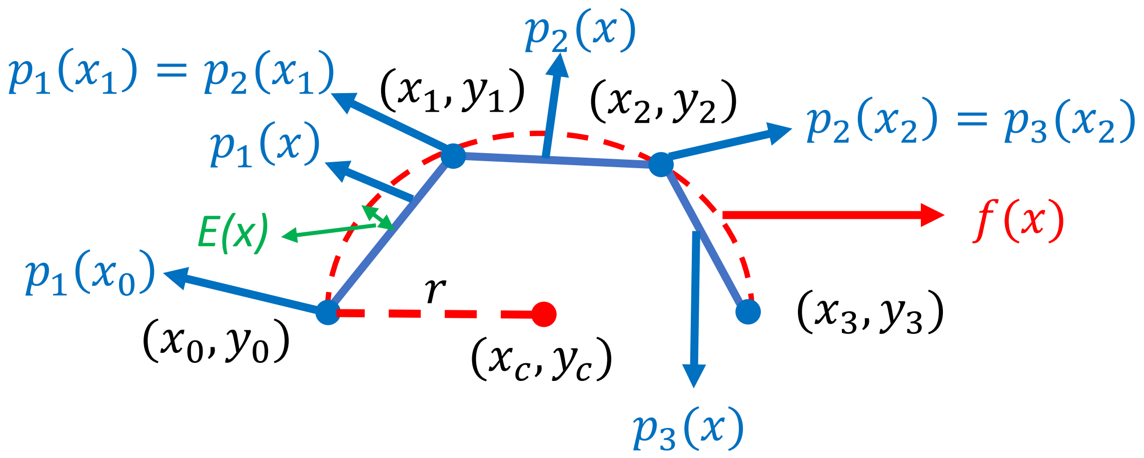
Euler Method vs Linear Interpolation

- $y' = f(x, y), \quad y(x_0) = y_0$
- $y_1 = y_0 + f(x_0, y_0)\Delta x$
- $x_1 = x_0 + \Delta x$
- $y_2 = y_1 + f(x_1, y_1)\Delta x$
- $x_2 = x_1 + \Delta x$

The number y_1, y_2 are the approximations of $y(x_1), y(x_2)$



- $y = f(x) = \sqrt{r^2 - (x - x_c)^2} + y_c$
- $p(x_k) = f(x_k) = y_k$ for all k
- $p_1(x) = \frac{x_1 - x}{x_1 - x_0} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1)$
- $E(x) = f(x) - p_1(x),$



Linear Interpolation Methods

Tells how approximating the function with straight lines for example, almost all graphs produced by computers are actually the results of piecewise linear interpolation in which the machine draws very large number of very small straight lines represent the curve.

Given a set of data points x_k: (called as nodes),

We say the function p interpolates the function f at these nodes,

if $\boldsymbol{p}(\boldsymbol{x_k}) = \boldsymbol{f}(\boldsymbol{x_k})$ for **all** $\boldsymbol{k}$.

We are most interested in the extent to $\boldsymbol{p} \approx \boldsymbol{f}$

Linear interpolation methods based on using straight line to approximate a given function.

$\Rightarrow$ 2 points to determine a straight line

x_0, x_1 and a function f $\Rightarrow$ we want to find the equation of a straight line that passes through these two points: $(x_0, f(x_0))$, $(x_1, f(x_1))$

$$p_1(x) = \frac{x_1 - x}{x_1 - x_0} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1)$$

Let us investigate how accurate the linear interpolation using Rolle's Theorem $\Rightarrow$ Special case of mean value theorem if $f(c) = f(b) \Rightarrow$ there exists ζ such that $f'(\zeta) = 0$.

Define,

$$E(x) = f(x) - p_1(x),$$

$$w(x) = (x - x_0)(x - x_1),$$

$$G(x) = E(x) - \frac{w(x)}{w(t)} E(t)$$

$t \rightarrow$ some fixed value in (x_0, x_1) where ($(x_0 < t < x_1)$)

$$G(x_0) = 0, G(x_1) = 0, G(t) = 0$$

Then the Rolle's Theorem states that;

There exists a point z_0 between x_0 and t such that $G'(z_0) = 0$

and a point z_1, between x_1 and t, such that $G'(z_1) = 0.$

Let's apply Rolle's Theorem to G' and assert that there exists

a point ζ between z_0 and z_1 such that $G''(\zeta) = 0.$

But,

$$G''(x) = F''(x) - \frac{2}{w(t)} E(t)$$

$$G''(\zeta) = 0 \quad => \quad F''(\zeta) - \frac{2}{w(t)} E(t) = 0$$

And we have $\underbrace{f(t) - p_1(t)}_{= E(t) \text{ because we defined } (E(x) = f(x) - p_1(x))} = \frac{1}{2}\overbrace{(t - x_0)(t - x_1)}^{w(t)} f''(\zeta)$

For any $t \in [x_0, x_1]$,

The error in the approximation will grow rapidly outside the interval $[x_0, x_1]$.

⇒ Let us take absolute values, investigate for the worst case for the second derivative term.

$$|f(x) - p_1(x)| \leq \frac{1}{2} |(x - x_0)(x - x_1)| \max_{x_0 \leq t \leq x_1} |f''(t)|$$
$$\leq \frac{1}{2} \left(\max_{x_0 \leq t \leq x_1} |(t - x_0)(t - x_1)| \right) \left(\max_{x_0 \leq t \leq x_1} |f''(t)| \right)$$

So the upper bound on the error depends on the maximum of the function:

$$g(x) = |(x - x_0)(x - x_1)| \;\; = (x_1 - x)(x - x_0)$$

$g(x_0) = g(x_1) = 0$, the Extreme Value Theorem says that the maximum value of g on the interval $[x_0, x_1]$ will be on critical point.

$$g'(x) = x_1 - x - x + x_0 = (x_1 + x_0) - 2x$$

$$x_c = \frac{1}{2}(x_0 + \; x_1) \qquad => g(x_c) = \frac{1}{4}(x_1 - \; x_0)^2$$

Finally our error is bounded

$$f(x) - p_1(x) \leq \frac{1}{8}(x_1 - x_0)^2 (\max_{x_0 \leq t \leq x_1} |f''(x)|)$$

<u>Theorem</u> (Linear Interpolation Error)

Let $f \in C^2([x_0, x_1])$ and let $p_1(x)$ be linear polynomials that interpolates f at x_0 and x_1, then for all $x \in [x_0, x_1]$,

$$\begin{aligned} |f(x) - p_1(x)| &\leq \frac{1}{2}|(x - x_0)(x - x_1)| \max_{x_0 \leq t \leq x_1} |f''(x)| \\ &\leq \frac{1}{8}(x_1 - x_0)^2 \max_{x_0 \leq t \leq x_1} |f''(x)| \end{aligned}$$

Note: $C^k([a, b])$ — The set of functions f such that f and its first k derivatives are all in C$([a, b])$

Example: Consider the problem of constructing a piecewise linear approximation to $f(x) = \log_2(x)$ using the nodes 1/4, 1/2, 1.

$$Q_1(x) = \left(\frac{\frac{1}{2} - \mathrm{x}}{\frac{1}{2} - \frac{1}{4}}\right) log_2\left(\frac{1}{4}\right) + \left(\frac{x - \frac{1}{4}}{\frac{1}{2} - \frac{1}{4}}\right) log_2\left(\frac{1}{2}\right) = 4x - 3$$

$$Q_2(x) = \left(\frac{1 - \mathrm{x}}{1 - \frac{1}{2}}\right) log_2\left(\frac{1}{2}\right) + \left(\frac{x - \frac{1}{2}}{1 - \frac{1}{2}}\right) log_2(1) = 2x - 2$$

$$q(x) = \begin{cases} 4x - 3, & \frac{1}{4} \leq x \leq \frac{1}{2} \\ 2x - 2, & \frac{1}{2} \leq x \leq 1 \end{cases}$$

The error

$$|log_2(x) - \ Q_1(x)\,| \le \frac{1}{8}\left(\frac{1}{2} - \frac{1}{4}\right)^2 \max_{t\in[\frac{1}{4},\frac{1}{2}]} |log_2(e)t^{-2}| \quad = 0.1803368801 \ \ for\ x \in [\frac{1}{4},\frac{1}{2}]$$

not always the same

$$|\ log_2(x) - \ Q_2(x)\,| \le \frac{1}{8}\left(1 - \frac{1}{2}\right)^2 \max_{t\in[\frac{1}{2},1]} \ |log_2(e)t^{-2}| = 0.1803368801 for\ x \in [\frac{1}{2},1]$$

$$\Rightarrow \ |\log_2(x) - q(x)| \le 0.1803368801$$

Application, The Trapezoid Rule

One of the most important application of the linear interpolation is the construction of the trapezoid rule for approximating definite integrals, define the integration of interest.

$$I(f) = \int_a^b f(x)dx$$

Let $p_1(x)$ be the linear polynomial that interpolates f at <u>x = a</u> and <u>x = b</u>

$$p_1(x) = \frac{x-a}{b-a} f(b) + \frac{b-x}{b-a} f(a)$$

The basic trapezoid rule is defined by exactly integrating $p_1(x)$;

$$T_1(f) = I(p_1) = \int_a^b p_1(x)dx = \frac{1}{2}(b-a)(f(b) + f(a))$$

$$\Rightarrow \frac{1}{b-a}\left[\left(\frac{x^2}{2}\Big|_a^b - a\cdot x\Big|_a^b\right)f(b) + \left(-\frac{x^2}{2}\Big|_a^b + b\cdot x\Big|_a^b\right)f(a)\right]$$

$$\frac{1}{b-a}\left[\left(\frac{b^2-a^2}{2} + a^2 - ab\right)f(b) + \left(\frac{a^2-b^2}{2} + b^2 - ab\right)f(a)\right]$$

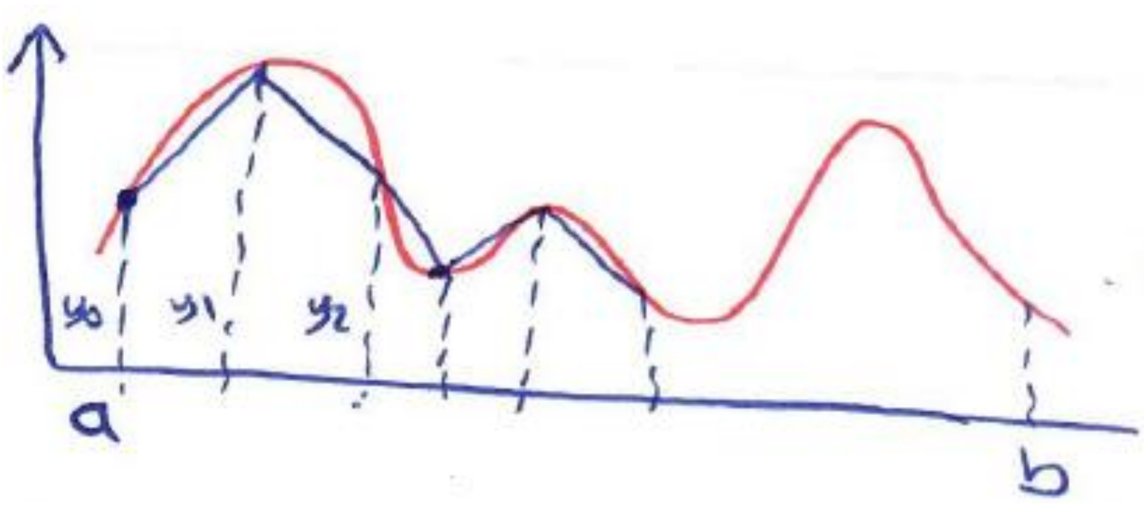
$$\frac{1}{b-a}\left[\left(\frac{a^2}{2} + \frac{b^2}{2} - ab\right)f(b) + \left(\frac{a^2}{2} + \frac{b^2}{2} - ab\right)f(a)\right]$$

$$\frac{1}{b-a}\left[\frac{1}{2}(a-b)^2 f(b) + \frac{1}{2}(a-b)^2 f(a)\right]$$

$$\frac{1}{b-a}\left[\frac{1}{2}(a-b)^2\big(f(b)+f(a)\big)\right]$$

$$\frac{1}{2}(b-a)\big(f(b)+f(a)\big)$$

*Note that here we want to make an approximation by replacing the "exact function" ($f(.)$) by a simpler function that approximates it ($p_1(.)$), and doing the desired calculation (integration) exactly on the simpler function.



$$I(f) = \int_a^b f(x)dx$$

$$\approx \left(\frac{b-a}{2n}\right)(y_0 + 2y_1 + \cdots + 2y_{n-1} + y_n)$$

How much accurate?

- The error = $I(f) - T_1(f)$
- $I(f) - T_1(f) = I(f) - I(p_1)$

$$I(f) - I(p_1) = \int_a^b f(x)dx - \int_a^b p_1(x)dx$$

$$= \int_a^b (f(x) - p_1(x))dx$$

$$= \frac{1}{2}\int_a^b (x-a)(x-b)f''(\zeta_x)dx$$

Using the interpolation error theory

$\zeta_x \in [a, b]$ and depends on x, since $(x-a)(x-b)$ does not change sign on $[a, b]$ we can now apply Integral Mean Value Theorem to set an error estimate.

[Remainder $\int_a^b g(t)f(t)dt = f(\zeta)\int_a^b g(t)dt$ $\zeta \in$ [a, b] $if\ g$ does not change sign on $[a, b]$]

$$\int_a^b (x-a)(x-b)f''(\zeta_x)dx = f''(z)\int_a^b (x-a)(x-b)dx$$

$$= -\frac{1}{6}(b-a)^3 f''(z)$$

$$(a^3 - b^3 - 3a^2b + 3ab^2)$$

$$(a-b)^3$$

$$\left[\frac{x^3}{3}\Big|_a^b - \frac{x^2}{2}\Big|_a^b \quad (a+b) + ab\,\Big|_a^b \right]$$

$$\frac{1}{6}(2b^3 - 2a^3 + 3a^3 - 3b^3 + 3a^2b + 3a^2b - 3ab^2 + 6ab^2 - 6a^2b)$$

We have:

$$I(f) - T_1(f) = -\frac{1}{12}(b-a)^3 f''(z) \qquad z \in [a,b]$$

Theorem: (Trapezoid Rule Error Estimate, (single subinterval))

Let $f \epsilon C^2([a,b])$ and let p_1 interpolate f at a and b. Define $T_1(f) = I(p_1)$. Then there exists $z \epsilon [a,b]$ such that:

$$I(f) - T_1(f) = -\frac{1}{12}(b-a)^3 f''(z)$$

** the error will be small if the length of the integration interval, $\underline{(b-a)}$, is <u>small</u>…

If we use more points in the approximation, we subdivide the interval [a, b] into n subintervals;

$$a = x_0 < x_1 < x_2 \ldots . < xn_{-1} < x_n = b$$

→ “the n-subinterval trapezoid rule” =>

$$I(f) = \sum_{i=1}^{n} \int_{x_i-1}^{x_i} f(x)dx \approx \sum_{i=1}^{n} \frac{1}{2}(x_i - x_{i-1})\big(f(x_i - 1) + f(x_i)\big) = T_n(f)$$

If we use uniform grid, meshpoints are equally spaced => $xi - x_{i-1} = h$

$$T_n(f) = \frac{h}{2}\big(f(x_0) + 2f(x_1) + 2f(x_2) + \cdots + 2f(x_{n-1}) + f(x_n)\big), \qquad where\ h = \frac{b-a}{n}$$

Theorem: (Trapezoid Rule Error Estimate, Uniform Grid)

Let $f \epsilon C^2([a,b])$ and let $T_n(f)$ be the n-subinterval trapezoid rule approximation to $I(f)$ using a uniform grid. Then there exists $\zeta_h \epsilon [a,b]$, depending on h, such that:

$$I(f) - Tn(f) = -\frac{b-a}{12} h^2 f''(\zeta_h)$$

Example: $f(x) = e^x$ and $[a,b] = [0,1]$ so that:

$$I(f) = \int_0^1 e^x \, dx = e - 1 = 1.71828 \ldots$$

Then the trapezoid rule using a single subinterval $T(f) = \frac{1}{2}(b-a)(f(b) + f(a))$

$$T_1(f) = \frac{1}{2} \cdot 1 \cdot (e^0 + e^1) = \frac{1+e}{2} = 1.859140$$

whereas the trapezoid rule using two subintervals is given by:

$$T_2(f) = \frac{1}{2}\left(\frac{b-a}{n}\right)(y_0 + 2y_1 + y_2) = \frac{1}{2}\frac{1}{2}\left(e^0 + 2e^{\frac{1}{2}} + e^1\right) = 1.75393 \ldots$$

Example: $I(f) = \int_0^1 e^{x^{-2}} dx$, how small does n have to be to guarantee that $|I(f) - T_n(f)| \leq 10^{-3}$?

Theorem: (Trapezoid Rule Error Estimate, Non-uniform Grid)

Let $f \epsilon C^2([a, b])$ and let $T_n(f)$ be the n-subinterval trapezoid rule approximation to $I(f)$ using a non-uniform grid defined by $a = x_0 < x_1 < x_2 \ldots . < x_{n-1} < x_n = b$.

with $h_i = xi - x_{i-1}$, $h = \max_i(h_i)$. Then,

$$|I(f) - T_n(f)| \leq \frac{b-a}{12} h^2 \max_{x \, \epsilon \, [a,b]} |f''|$$

ROOT FINDING

A common problem encountered in engineering analysis;

given a function $f(x)$, determine the values of x for which $f(x) = 0$.

-> The solutions are known as the roots of the equation $f(x) = 0$ or the zeros of the function $f(x)$.

- The equation $y = f(x)$ => contains three elements; an input value x, an output value y, and the rule “f” for computing y.
- The roots of equations maybe real or complex,
- The complex roots are seldom computed, because they rarely have physical significance
- In general, an equation may have any number of roots, or no roots at all.

For example: $sinx - x = 0 \Rightarrow$ has a single root $x = 0$

$\tan x - x = 0 \Rightarrow$ has infinite number of roots ($x = 0, \pm 4.493, \pm 7.725, \dots$)

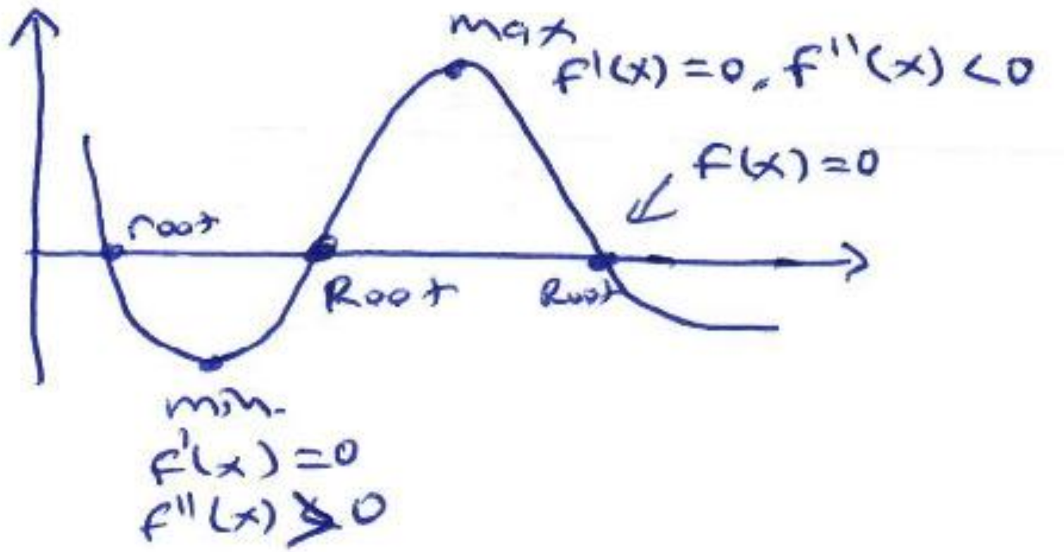
You have learned to use a quadrate formula

$x = -\frac{b \pm \sqrt{b^2 - 4ac}}{2a}$ to solve $f(x) = ax^2 + bx + c = 0$

Although the quadratic formula is handy for solving this equation, but there are many other functions for which the root cannot be determined so easily.

Before advent of digital computers, there were a number of ways to solve for the roots.

For complicated function, an approximate solution technique is needed. One method to obtain an approximate the solution is to plot the function and determine where it $\underbrace{\text{crosses the x axis}}_{\text{Represents x value } f(x) = 0}$.



Although graphical methods are useful for obtaining rough estimates of roots, they are limited because of their lack of precision.

An alternative approach is to use trial and error

-> based on guessing a value of x and evaluating whether $f(x)$ is zero

If not, another guess is made and again evaluate $f(x)$

-> it is repeated until a guess results in an $f(x)$ that is close to zero

** Besides roots, another feature of interest to engineers are a function's $\underbrace{\text{min. and max. values.}}_{f'(x)=0}$

- The determination of such optimal values is referred to as optimization

All methods of finding roots are iterative procedures that require a starting point -> an estimate of the root.

-> This estimate can be crucial a bad starting value may fail to convergence or it may converse to the wrong root.

There is NO universal recipe for estimating the value of a root a systematic numerical search for the roots is needed.

-> Bisection method

-> Newton’s method

-> The secant method

-> Fixed point iteration

-> Special Topics

1) Bisection Method

T h e bisection method is a variation of the incremental method in which the interval is always divided in half

- Halving the interval until it becomes sufficiently small

Interval halving method

If a function changes sign over an interval, the function value at the midpoint is evaluated

The location of root is determined within the subinterval where the sign change occurs.

The subinterval then becomes the interval for the next iteration.

The process is repeated until the root is known to the required precision

If there is a root in the interval (x_1, x_2) then

$f(x_1)f(x_2) < 0$ -> this means that f is negative at one point, and positive at the other point.

If we assume that f is continuous => (the Intermediate Value Theorem), there must be some value between a and b where $f(x) = 0$

In order to halve the interval, we compute $f(x_3)$ where $x_3 = \frac{1}{2}(x_1 + x_2)$, there are three possibilities;

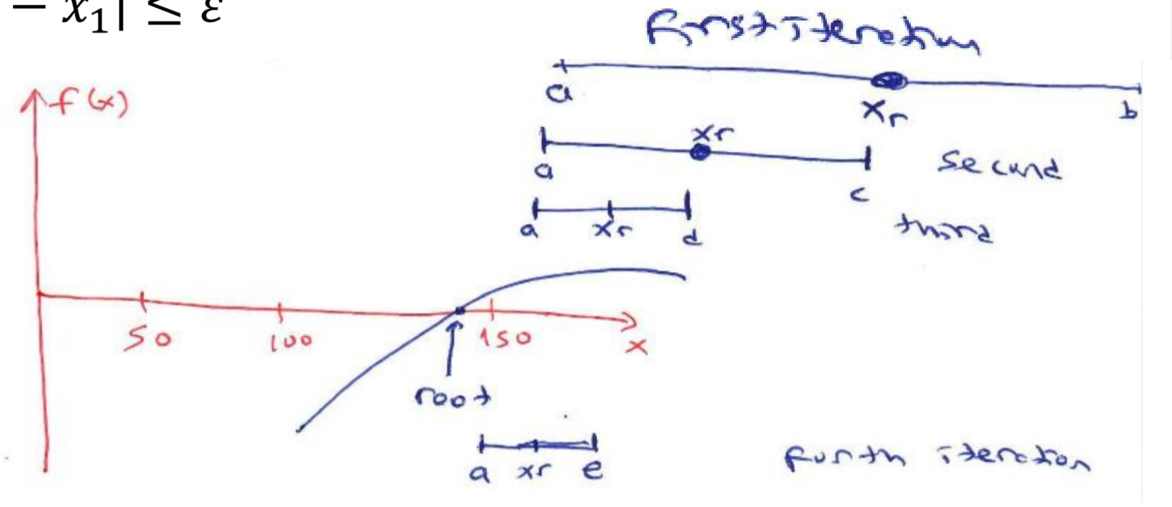
1-) $f(x_1)f(x_3) < 0 \Rightarrow$ This means that a root is between $x_1 \; and \; x_3$ (there might be more than one) $[x_1, x_3]$

2-) $f(x_1)f(x_3) = 0 \Rightarrow$ If we assume that we already know $f(x_1) \neq 0 \Rightarrow$ this means we have found the root x_3.

3-) $f(x_1)f(x_3) > 0 \;\Rightarrow$ This means that a root must lie in the other half of the interval $[x_3, x_2]$

The new interval is half the size of the original interval. The bisection is repeated until the interval has been reduced to a small value ε,

$$|x_2 - x_1| \leq \varepsilon$$



-> Use bisection to solve the problem graphically

$a = 50, b = 200, \qquad f(50)f(150) < 0$

$$x_r = \frac{50+200}{2} = 125$$

Note that exact value of the root is 142.7376

$f(50)f(125) > 0 \Rightarrow$ The root must lie in the other half of the interval

$f(125)f(200) < 0 \Rightarrow x_r = \frac{125+200}{2} = 162.5$

$f(125)f(162.5) < 0 \Rightarrow x_r = \frac{125+162.5}{2} = 143.75$

Relative error

$|\varepsilon_t| = \left|\frac{142.7376-125}{142.7376}\right| 100 = 12.43\%$ after first iteration, $|\varepsilon_t| = 13.85\%$ after second iteration,

$|\varepsilon_x| = 0.709\%$ -> The method can be repeated until the result is accurate enough.

Example: $f(x) = 2 - e^x, [a,b] = [0,1]$

$$f(a) = 1, f(b) = -0.7183 \Rightarrow c = \frac{0+1}{2} = \frac{1}{2} \Rightarrow f(c) = 0.3513 > 0$$

$$[a,b] \rightarrow [\frac{1}{2}, 1]$$

$$f(a) = 0.3513, f(b) = -0.7183 \Rightarrow c = \frac{1/2+1}{2} = \frac{3}{2} \Rightarrow f(c) = -0.117 < 0$$

$$[a,b] \rightarrow [\frac{1}{2}, \frac{3}{4}]$$

$$f(a) = 0.3513, f(b) = -0.117 \Rightarrow c = \frac{1/2+3/4}{2} = \frac{5}{8} \Rightarrow f(c) = 0.1318 > 0$$

$$[a,b] \rightarrow [\frac{5}{8}, \frac{3}{4}]$$

we have reduced interval of uncertainty from $[0,1]$ which has length 1 to $\left[\frac{5}{8}, \frac{3}{4}\right] \Rightarrow \frac{1}{8} = 0.125$

Algorithm (Bisection Method)

1) Given initial interval $[a_0, b_0] = [a, b]$, set $k = 0$
2) Compute $c_{k+1} = a_k + \frac{1}{2}[b_k - a_k]$
3) If $f(c_{k+1})f(a_k) < 0$, set $a_{k+1} = a_k$, $b_{k+1} = c_{k+1}$,
4) If $f(c_{k+1})f(b_k) < 0$, set $b_{k+1} = b_k$, $a_{k+1} = c_{k+1}$
5) Update k and go to Step 2

*For very large values of a and b, $(a + b)/2$ can lead to a computational overflow, whereas $a + \frac{1}{2}(b - a)$ will not