Solving Nonlinear Equations

In previous lectures, we learnt how to solve linear systems. Today's lecture, focuses on solving nonlinear equations (finding roots of a single variable function). Some examples include

$$x^2 - 6x + 9 = 0, \qquad x - \cos(x) = 0, \qquad e^x \ln(x^2) - x\cos(x) = 0$$

Numerical methods to consider include

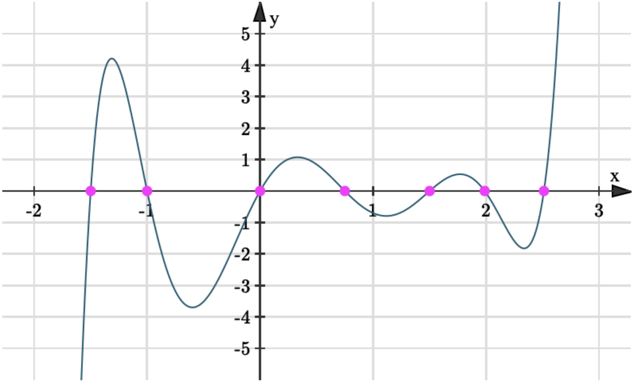
- Fixed Point Method
- Bisection Method
- Regula-Falsi Method
- Newton's Method
- Secant Method

Finding Roots

A root of a function $F : \mathbb{R} \to \mathbb{R}$ is a number x^* such that

$$F(x^*) = 0,$$

in other word, x^* is the point where F crosses the x axis.



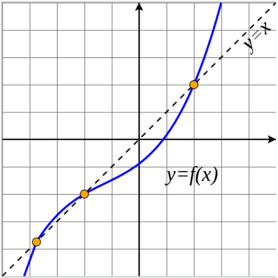
Necessity of Numerical Methods

- If direct methods are not available, numerical iterative techniques are used.
- **Iterative Method:** starts with an initial solution estimate x_0 and proceed by recursively computing improved estimates $x_1, x_2, \ldots, x_n$ until a certain stopping criterion is satisfied.
- Numerical methods typically give only an *approximation* to the exact solution.
- The approximation can be of a very good (predefined) accuracy, depending on the amount of computational effort one is willing to invest.
- One of the advantages of numerical methods is their simplicity: they can be concisely expressed in algorithmic form and can be easily implemented on a computer.

Fixed Point Method

- **Assumption:** $F : [a, b] \to \mathbb{R}$, F has a root in the interval (a, b)
- **Goal:** find $x^* \in (a, b)$ such that x^* solves $F(x) = 0$, that is $F(x^*) = 0$.

  Note that at x^*, we have $x^* = x^* - F(x^*)$.
- **Key Step:** Define
$$f(x) = x - F(x)$$
  and find the fixed point of $f(x)$, that is
$$x = f(x)$$



Fixed Point Iteration

Choose

$$x_0 \in (a, b)$$

and sequentially calculate $x_1, x_2, \ldots$ using the formula

$$x_{k+1} = f(x_k), \qquad k = 0, 1, 2, \ldots$$

Note: if $x_k \to x^*$ as $k \to \infty$ and if f is continuous, then

$$x^* = \lim_{k\to\infty} f(x_k) = f(\lim_{k\to\infty} x_k) = f(x^*)$$

Then

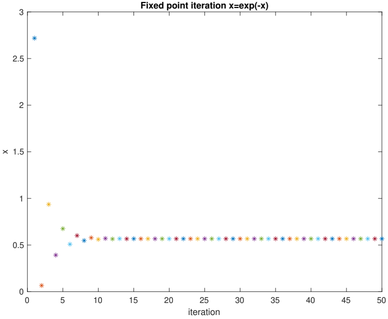
$$x^* = f(x^*) \quad \leftrightarrow \quad x^* - f(x^*) = 0 \quad \leftrightarrow \quad F(x^*) = 0$$

Example 1

Find the fixed point of $\exp(-x)$, using fixed point method starting with $x_0 = -1$.

$$x_{k+1} = \exp(-x_k)$$

$$x_1 = \exp(-x_0) = \exp(1) \approx 2.7182, \quad x_2 = 0.0660, \quad x_3 = 0.9361,$$
$$x_4 = 0.3921, \quad \cdots x_{50} = 0.5671$$



Convergence Criteria

Theorem

Assume that $f : (a, b) \to \mathbb{R}$ *is differentiable on* (a, b) *and there exists a constant* q *such that*

$$|f'(x)| \leq q \leq 1 \quad \forall x \in (a, b).$$

Then there exists a unique solution $x^* \in (a, b)$ *that* $x^* = f(x^*)$*, and* $x_{k+1} = f(x_k)$ *satisfies the following inequality:*

$$|x_k - x^*| \leq q^k |x_0 - x^*|, \quad k \geq 0$$

Proof. Let $x^* \in (a, b)$ be an arbitrary fixed point of f, and let $x_k \in (a, b)$. By the mean value theorem, there exists a c_k between x_k and x^* in (a, b) such that

$$f(x_k) - f(x^*) = f'(c_k)(x_k - x^*).$$

$$\begin{aligned} |x_k - x^*| &= |f(x_{k-1}) - f(x^*)| \\ &= |f'(c_{k-1})(x_{k-1} - x^*)| \\ &= |f'(c_{k-1})||x_{k-1} - x^*)| \\ &\leq q|x_{k-1} - x^*| \end{aligned}$$

Hence,

$$\begin{aligned} |x_k - x^*| &\leq q\,|x_{k-1} - x^*| \\ &\leq q^2\,|x_{k-2} - x^*| \\ &\vdots \\ &\leq q^k\,|x_0 - x^*| \end{aligned}$$

Note that since $0 \leq q < 1$, as k approaches to infinity

$$q^k|x_0 - x^*| \to 0$$

Bisection Method

- $f : [a, b] \to \mathbb{R}$ is continuous and $f(a)f(b) \leq 0$, then there is a $x^* \in [a, b]$ such that $f(x^*) = 0$. Let assume $f(a) > 0$ and $f(b) < 0$.
- **Process:** Start with $a_0 = a$ and $b_0 = b$, find the midpoint of a_0 and b_0 as follows

$$c_0 = \frac{a_0 + b_0}{2} \quad \text{and} \quad f(c_0).$$

  Assume that $f(c_0) < 0$. Since $f(a)f(c_0) < 0$, then $[a_0, c_0]$ is guaranteed to contain a root of f.
- We then set

$$a_1 = a_0 \qquad b_1 = c_0$$

  and find

$$c_1 = \frac{a_1 + b_1}{2} \quad \text{and} \quad f(c_1).$$

- If $f(c_1) > 0$ then it is guaranteed that $[c_1, b_1]$ to contain a root, hence

$$a_2 = c_1 \qquad b_2 = b_1$$

- We continue the search until the length of the interval $b_n - a_n < \epsilon$, and

$$c_n = \frac{a_n + b_n}{2}$$

is the output as an approximation of the root x^*.

- Alternative stopping criteria is $|f(c_n)| < \epsilon$.
- The bisection algorithm produces a set of centers $c_1, c_2, \ldots, c_k, \ldots$ such that

$$\lim_{k \to \infty} c_k = x^*$$

- Since c_n is the midpoint of the interval $[a_n, b_n]$, we have

$$|x^* - c_n| \leq \frac{1}{2}(b_n - a_n)$$

- Note that

$$b_n - a_n = \frac{1}{2}(b_{n-1} - a_{n-1}) = \frac{1}{2^2}(b_{n-2} - a_{n-2}) = \cdots = \frac{1}{2^n}(b_0 - a_0)$$

Hence

$$|x^* - c_n| \leq \frac{1}{2^{n+1}}(b_0 - a_0)$$

The inequality

$$|x^* - c_n| \leq \frac{1}{2^{n+1}}(b_0 - a_0)$$

can be used to determine the number of iterations required to achieve a given precision ϵ by solving the inequality

$$\frac{1}{2^{n+1}}(b_0 - a_0) < \epsilon$$

which is equivalent to

$$n > \log_2\left(\frac{b_0 - a_0}{\epsilon}\right) - 1 = \frac{\ln\left(\frac{b_0 - a_0}{\epsilon}\right)}{\ln 2} - 1$$

For example, if $[a_0, b_0] = [0, 1]$, and $\epsilon = 10^{-5}$, we have

$$n > \frac{5\ln(10)}{\ln 2} - 1 \approx 15.6 \quad \rightarrow \quad n = 16$$

Example 2

Approximate a root $x^* \in [0, 1]$ of

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

Example 2

Approximate a root $x^* \in [0, 1]$ of

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

Answer. Note that

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

hence f has a root on $[0, 1]$. In fact, f has exactly one root, because $f'(x) = 3x^2 + 1 > 0$ for all $x \in [0, 1]$.

Example 2

Approximate a root $x^* \in [0, 1]$ of

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

Answer. Note that

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

hence f has a root on $[0, 1]$. In fact, f has exactly one root, because $f'(x) = 3x^2 + 1 > 0$ for all $x \in [0, 1]$.

- $a_0 = 0$, $b_0 = 1$, $c_0 = \frac{1}{2}$, $f(c_0) = (0.5)^3 + 0.5 + 1 = -0.375 < 0$.

Example 2

Approximate a root $x^* \in [0,1]$ of

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

Answer. Note that

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

hence f has a root on $[0,1]$. In fact, f has exactly one root, because $f'(x) = 3x^2 + 1 > 0$ for all $x \in [0,1]$.

- $a_0 = 0$, $b_0 = 1$, $c_0 = \frac{1}{2}$, $f(c_0) = (0.5)^3 + 0.5 + 1 = -0.375 < 0$.
- $f(c_0)f(b) < 0 \rightarrow a_1 = 0.5, b_1 = 1$, $c_1 = (0.5+1)/2 = 0.75$, $f(c_1) = (0.75)^3 + 0.75 + 1 = 0.172 > 0$

Example 2

Approximate a root $x^* \in [0, 1]$ of

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

Answer. Note that

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

hence f has a root on $[0, 1]$. In fact, f has exactly one root, because $f'(x) = 3x^2 + 1 > 0$ for all $x \in [0, 1]$.

- $a_0 = 0$, $b_0 = 1$, $c_0 = \frac{1}{2}$, $f(c_0) = (0.5)^3 + 0.5 + 1 = -0.375 < 0.$
- $f(c_0)f(b) < 0 \to a_1 = 0.5, b_1 = 1$, $c_1 = (0.5 + 1)/2 = 0.75$, $f(c_1) = (0.75)^3 + 0.75 + 1 = 0.172 > 0$
- $f(a_1)f(c_1) < 0 \to [a_2, b_2] = [0.5, 0.75]$, $c_2 = (0.5 + 0.75)/2 = 0.625$, $f(c_2) = -0.131 < 0$

Example 2

Approximate a root $x^* \in [0,1]$ of

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

Answer. Note that

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

hence f has a root on $[0,1]$. In fact, f has exactly one root, because $f'(x) = 3x^2 + 1 > 0$ for all $x \in [0,1]$.

- $a_0 = 0$, $b_0 = 1$, $c_0 = \frac{1}{2}$, $f(c_0) = (0.5)^3 + 0.5 + 1 = -0.375 < 0$.
- $f(c_0)f(b) < 0 \to a_1 = 0.5, b_1 = 1$, $c_1 = (0.5+1)/2 = 0.75$, $f(c_1) = (0.75)^3 + 0.75 + 1 = 0.172 > 0$
- $f(a_1)f(c_1) < 0 \to [a_2, b_2] = [0.5, 0.75]$, $c_2 = (0.5 + 0.75)/2 = 0.625$, $f(c_2) = -0.131 < 0$
- $f(b_2)f(c_2) < 0 \to [a_3, b_3] = [0.625, 0.75]$, $c_3 = (0.625 + 0.75)/2 = 0.6875$, $f(c_3) = 0.01245 > 0$

Example 2

Approximate a root $x^* \in [0, 1]$ of

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

Answer. Note that

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

hence f has a root on $[0, 1]$. In fact, f has exactly one root, because $f'(x) = 3x^2 + 1 > 0$ for all $x \in [0, 1]$.

- $a_0 = 0,\ b_0 = 1,\ c_0 = \frac{1}{2},\ f(c_0) = (0.5)^3 + 0.5 + 1 = -0.375 < 0.$
- $f(c_0)f(b) < 0 \rightarrow a_1 = 0.5, b_1 = 1,\ c_1 = (0.5 + 1)/2 = 0.75,$
  $f(c_1) = (0.75)^3 + 0.75 + 1 = 0.172 > 0$
- $f(a_1)f(c_1) < 0 \rightarrow [a_2, b_2] = [0.5, 0.75],\ c_2 = (0.5 + 0.75)/2 = 0.625,$
  $f(c_2) = -0.131 < 0$
- $f(b_2)f(c_2) < 0 \rightarrow [a_3, b_3] = [0.625, 0.75],\ c_3 = (0.625 + 0.75)/2 = 0.6875,$
  $f(c_3) = 0.01245 > 0$
- Resume the process

Bisection method for solving $f(x) = 0$

Input: f, ϵ, a, b such that $f(a)f(b) < 0$

Output: $\bar{x}$

$a_0 = a$, $b_0 = b$, $c_0 = (a_0 + b_0)/2$

$n = \lceil \ln(\frac{b_0 - a_0}{\epsilon}) / \ln 2 \rceil$

for $k = 1 : \ldots n$ do

  if $f(a_{k-1})f(c_{k-1}) \leq 0$ then

    if $f(c_{k-1}) = 0$ then

      return $\bar{x} = c_{k-1}$

    end if

    $a_k = a_{k-1}$, $b_k = c_{k-1}$

  else

    $a_k = c_{k-1}$, $b_k = b_{k-1}$

  end if

  $c_k = (a_k + b_k)/2$

end for

Intervals with Multiple Roots

If $[a, b]$ is known to contain multiple roots of s continuous function f, then

- if $f(a)f(b) > 0$, then there is an even numbers of roots of f in $[a, b]$;
- if $f(a)f(b) < 0$, then there is an odd number of roots of f in $[a, b]$.

Intervals with Multiple Roots

If $[a, b]$ is known to contain multiple roots of s continuous function f, then

- if $f(a)f(b) > 0$, then there is an even numbers of roots of f in $[a, b]$;
- if $f(a)f(b) < 0$, then there is an odd number of roots of f in $[a, b]$.

Example. Consider $f(x) = x^6 + 4x^4 + x^2 - 6$, where $x \in [-2, 2]$.

$$f(-2) = f(2) = 126, \quad f(-2)f(2) > 0$$

Intervals with Multiple Roots

If $[a, b]$ is known to contain multiple roots of s continuous function f, then

- if $f(a)f(b) > 0$, then there is an even numbers of roots of f in $[a, b]$;
- if $f(a)f(b) < 0$, then there is an odd number of roots of f in $[a, b]$.

Example. Consider $f(x) = x^6 + 4x^4 + x^2 - 6$, where $x \in [-2, 2]$.

$$f(-2) = f(2) = 126, \quad f(-2)f(2) > 0$$

By the bisection method,

$$[a_0, b_0] = [-2, 2], c_0 = 0, f(0) = -6 < 0.$$

Intervals with Multiple Roots

If $[a, b]$ is known to contain multiple roots of s continuous function f, then

- if $f(a)f(b) > 0$, then there is an even numbers of roots of f in $[a, b]$;
- if $f(a)f(b) < 0$, then there is an odd number of roots of f in $[a, b]$.

Example. Consider $f(x) = x^6 + 4x^4 + x^2 - 6$, where $x \in [-2, 2]$.

$$f(-2) = f(2) = 126, \quad f(-2)f(2) > 0$$

By the bisection method,

$$[a_0, b_0] = [-2, 2], c_0 = 0, f(0) = -6 < 0.$$

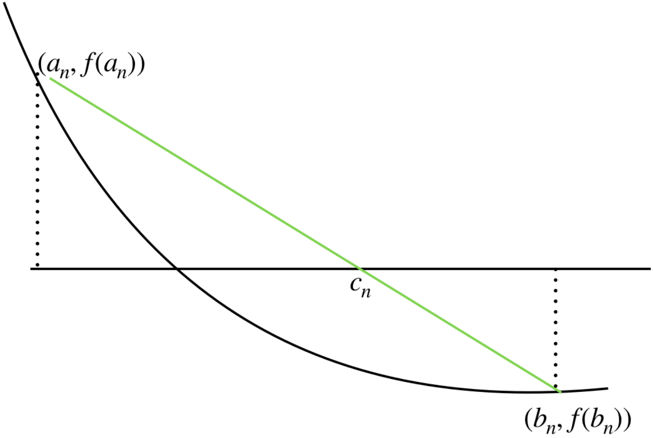
since

$$f(-2)f(0) < 0 \qquad f(0)f(2) < 0$$

then both intervals $[-2, 0]$ and $[0, 2]$ contain a root. Apply the bisection method for each of these intervals to find the corresponding roots.

Regula-falsi Method or False-position Method

Similar to bisection method, but instead of the mid-point at the kth iteration, we take the point c_k defined by an intersection of the line segment joining the points $(a_k, f(a_k))$ and $(b_k, f(b_k))$ with the x-axis.



The line passing through $(a, f(a))$ and $(b, f(b))$ is given by

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

The line passing through $(a, f(a))$ and $(b, f(b))$ is given by

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

so for $y = 0$,

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

implying

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

and then

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

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

The line passing through $(a, f(a))$ and $(b, f(b))$ is given by

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

so for $y = 0$,

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

implying

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

and then

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

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

Hence, we have the following expression for c_k

$$c_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}, \quad k = 1, 2, 3, \ldots$$

Regula-falsi method for solving $f(x) = 0$

Input: f, ϵ, a, b such that $f(a)f(b) < 0$
Output: $\bar{x}$ such that $|f(\bar{x})| < \epsilon$
$k = 0$, $a_0 = a$, $b_0 = b$, $c_0 = \frac{a_0 f(b_0) - b_0 f(a_0)}{f(b_0) - f(a_0)}$
repeat
    $k = k + 1$
    if $f(a_{k-1})f(c_{k-1}) \leq 0$ then
        if $f(c_{k-1}) = 0$
            return $\bar{x} = c_{k-1}$
        end if
        $a_k = a_{k-1}$, $b_k = c_{k-1}$
    else
        $a_k = c_{k-1}$, $b_k = b_{k-1}$
    end if

    $c_k = \frac{a_k f(b_k) - b_k f(a_k)}{f(b_k) - f(a_k)}$

until $|f(c_k)| < \epsilon$
return $\bar{x} = c_n$