COURSE 10

5.1. One-step methods for solving nonlinear eq. in $\mathbb{R}$
(continuation)

5.1.2. Fixed point iteration (succesive approximation method)

Definition 1 *The number α is called* **a fixed point** *of the function g if $g(\alpha) = \alpha$.*

So, *a fixed point* of a function is a number for which the value of the function is not changing when the function is applied.

Example 2 *Find the fixed points of the function $g(x) = x^2 - 2x + 2$.*

Sol. A fixed point α of g has the property $\alpha = g(\alpha) = \alpha^2 - 2\alpha + 2$, so $0 = \alpha^2 - 3\alpha + 2 = (\alpha - 1)(\alpha - 2)$. Whence, the fixed points of g are $\alpha_1 = 1$ and $\alpha_2 = 2$.

Geometrically, the fixed points are the intersection points of the graph of the function g and the first bisection line ($y = x$).

Sufficient condition for the existence and uniqueness of a fixed point:

Theorem 3 *1. If $g \in C[a,b]$ and $g(x) \in [a,b]$ for any $x \in [a,b]$, then g has at least one fixed point in $[a,b]$. In fewer words, if $g : [a,b] \to [a,b]$ and $g \in C[a,b]$ then $\exists \alpha \in [a,b]$ fixed point.*

2. Moreover, if there exists $g'(x)$ in (a,b) and there exists $K < 1$ such that

$$|g'(x)| \leq K, \quad \forall x \in (a,b),$$

then the fixed point is unique in $[a,b]$.

Example 4 *Prove that $g(x) = (x^2 - 4)/5$ has a unique fixed point in $[-2,2]$.*

Sol. The minimum and maximum of $g(x)$ for $x \in [-2, 2]$ are the limits of the interval, or at the points where $g'(x) = 0$. We have $g'(x) = 2x/5$, g is continuous and there exists $g'(x)$ in $[-2, 2]$. So, the minimum and maximum of $g(x)$ on $[-2, 2]$ are at $x = -2$, $x = 0$ or $x = 2$. We have $g(-2) = 0$, $g(2) = 0$, $g(0) = -4/5$, so $x = -2$ and $x = 2$ are points of absolute maximum and $x = 0$ is a point of absolute minimum in $[-2, 2]$. Moreover,

$$|g'(x)| = \left|\frac{2x}{5}\right| \leq \left|\frac{4}{5}\right| < 1, \quad \forall x \in (-2, 2).$$

So, g satisfies the conditions of Theorem 3, so it follows that g has a unique fixed point in $[-2, 2]$.

Consider the equation

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

where $f : [a, b] \to \mathbb{R}$. Assume that $\alpha \in [a, b]$ is a zero of $f(x)$.

In order to compute α, we transform (1) algebraically into *fixed point form*,

$$x = F(x), \tag{2}$$

where F is chosen so that $F(x) = x \Leftrightarrow f(x) = 0$.

A simple way to do this is, for example, $x = x + f(x) =: F(x)$.

Finding a zero of $f(x)$ in $[a, b]$ is then equivalent to finding a fixed point $x = F(x)$ in [a, b].

The fixed point form suggests *the fixed point iteration*

$$x_0 - \text{initial guess}, x_{k+1} = F(x_k),\ \ k = 0, 1, 2, ...$$

The hope is that iteration will produce a convergent sequence $x_k \to \alpha$.

For example, consider

$$f(x) = xe^x - 1 = 0. \tag{3}$$

A first fixed point iterationis obtained rearranging and dividing (3) by e^x: $xe^x = 1 \Rightarrow x = e^{-x}$, so $x = F(x) = e^{-x}$ and

$$x_{k+1} = e^{-x_k}.$$

With the initial guess $x_0 = 0.5$ we obtain the iterates $x_1 = 0.6065306597, x_2 = 0.5452392119, ..., x_8 = 0.5664094527, x_9 = 0.5675596343, ..., x_{28} = 0.56714328, x_{29} = 0.56714329$

So x_k seems to converge to $\alpha = 0.5671432...$

A second fixed point form is obtained from $xe^x = 1$ by adding x on both sides: $xe^x + x = 1 + x \Rightarrow x(e^x + 1) = 1 + x \Rightarrow x = \frac{1+x}{e^x+1}$, we get

$$x = F(x) = \frac{1+x}{e^x + 1}.$$

This time the convergence is much faster (we need only three iterations to obtain a 10-digit approximation of α) : $x_0 = 0.5$, $x_1 = 0.5663110032$, $x_2 = 0.5671431650$, $x_3 = 0.5671432904$.

Another possibility for a fixed point iteration is $x = x + 1 - xe^x$. But this iteration function does not generate a convergent sequence.

Finally we could also consider the fixed point form $x = x + xe^x - 1$. Also this iteration function does not generate a convergent sequence.

The question is: when does the iteration converge?

Answer: when condition of Theorem 3.

For the example, we have two cases where $|F'(\alpha)| < 1$ and the algorithm converges and two where $|F'(s)| > 1$ and the algorithm diverges.

A more general statement for convergence is the theorem of Banach.

Definition 5 **A Banach space** *$\mathcal{B}$ is a complete normed vector space over some number field K such as $\mathbb{R}$ or $\mathbb{C}$. (*Complete *means that every Cauchy sequence converges in $\mathcal{B}$.)*

Definition 6 *Let $A \subset \mathcal{B}$ be a closed subset and $F : A \to A$. F is called* **Lipschitz continuous** *on A if there exists a constant $L \geq 0$*

such that $||F(x) - F(y)|| \leq L\,||x - y||\,,\ \forall x, y \in A$. *Furthermore,* F *is called a contraction if* L *can be chosen such that* $L < 1$.

Theorem 7 *(Banach Fixed Point Theorem) Let* A *be a closed subset of a Banach space* $\mathcal{B}$, *and let* F *be a contraction* $F : A \to A$. *Then:*

a) F *has a unique fixed point* α, *which is the unique solution of the equation* $x = F(x)$.

b) The sequence $x_{n+1} = F(x_n)$ *converges to* α *for every initial guess* $x_0 \in A$.

c) We have the estimate: $||\alpha - x_n|| \leq \frac{L^{n-l}}{1-L}||x_{l+1} - x_l||$, *for* $0 \leq l \leq n$ *(or* $||\alpha - x_n|| \leq \frac{L^n}{1-L}||x_1 - x_0||$)

For practical applications is useful the following estimation.

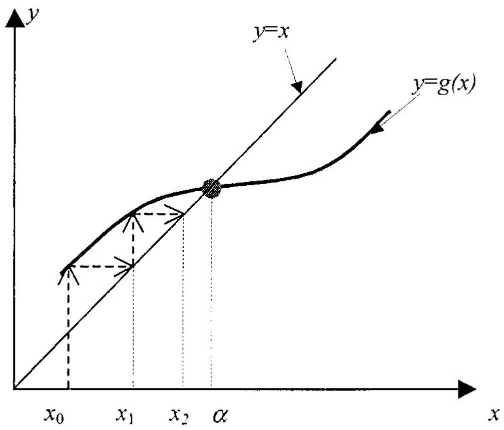
Lemma 8 *If* $\|F'(\alpha)\| < K$, $x \in V(\alpha)$ *then*

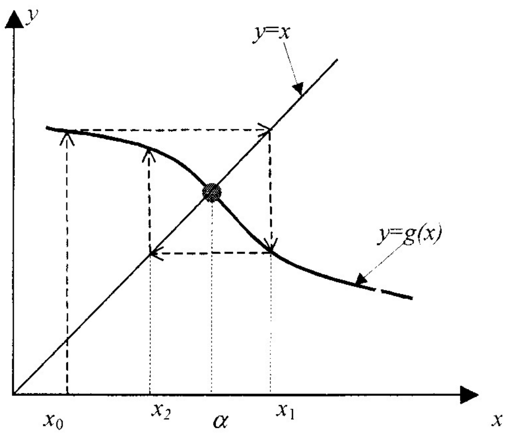
$$||\alpha - x_n|| \leq \frac{K}{1-K}||x_n - x_{n-1}||.$$

Geometric interpretation of the method: we plot $y = F(x)$ and $y = x$. The intersection points of the two functions are the solutions of $x = F(x)$. The computation of the sequence $\{x_k\}$ with x_0 chosen initial value, $x_{k+1} = F(x_k), k = 0, 1, 2, ...$ can be interpreted geometrically via sequences of lines parallel to the coordinate axes:

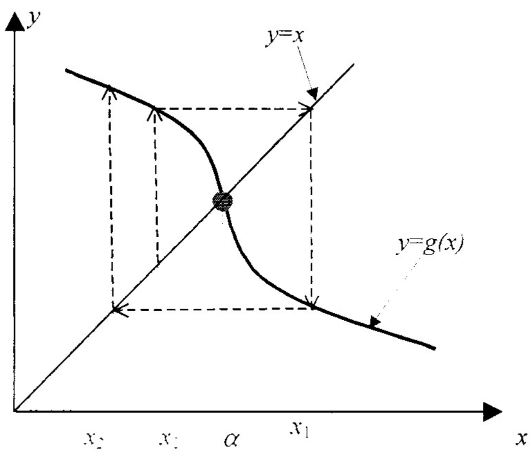
x_0	start with x_0 on the x-axis
$F(x_0)$	go parallel to the y-axis to the graph of F
$x_1 = F(x_0)$	move parallel to the x-axis to the graph $y = x$
$F(x_1)$	go parallel to the y-axis to the graph of F
$etc.$	

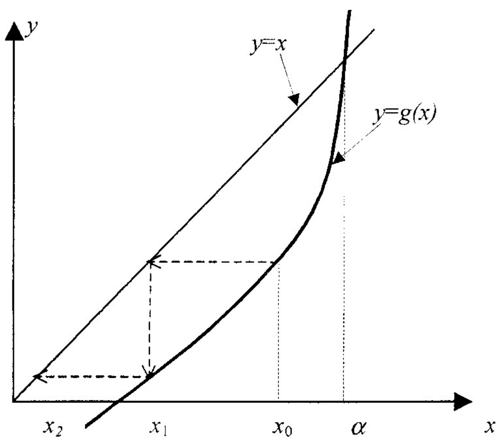
Case of convergence $|g'(x)| < 1$.





Case of divergence $|g'(x)| > 1$.





5.2. Multistep methods for solving nonlinear eq. in $\mathbb{R}$

Let $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}$. Consider the equation

$$f(x) = 0, \quad x \in \Omega, \tag{4}$$

We attach a mapping $F : D \to \Omega,\ D \subset \Omega^n$ to this equation.

Let $(x_0, ..., x_n) \in D$ be *the starting points*. We construct iteratively the sequence

$$x_0, x_1, ..., x_{n-1}, x_n, x_{n+1} \cdots \tag{5}$$

with

$$x_i = F(x_{i-n-1}, ..., x_{i-1}), \quad i = n+1, \ldots. \tag{6}$$

The problem consists in choosing F and $x_0, ..., x_n \in D$ such that the sequence (5) to be convergent to the solution of the equation (4).

In this case, the F-method is **a multistep method**.

It is based on interpolation methods with more than one interpolation node.

Let $\alpha \in \Omega$ be a solution of equation (4), let $(a, b) \subset \Omega$ be a neighborhood of α that isolates this solution and $x_0, ..., x_n \in (a, b)$, some given values.

Denote by g the inverse function of f, assuming it exists. Because $\alpha = g(0)$, the problem reduces to approximating g by an interpolation method with $n > 1$ nodes, for example Lagrange, Hermite, Birkhoff, etc...

Lagrange inverse interpolation

Let $y_k = f(x_k)$, $k = 0, ..., n$, hence $x_k = g(y_k)$. We attach the Lagrange interpolation formula to y_k and $g(y_k)$, $k = 0, ..., n$:

$$g = L_n g + R_n g, \tag{7}$$

where

$$(L_n g)(y) = \sum_{k=0}^{n} \frac{(y-y_0)...(y-y_{k-1})(y-y_{k+1})...(y-y_n)}{(y_k-y_0)...(y_k-y_{k-1})(y_k-y_{k+1})...(y_k-y_n)} g(y_k). \tag{8}$$

Taking

$$F_n^L(x_0, ..., x_n) = (L_n g)(0),$$

F_n^L is a $(n+1)-$ steps method defined by

$$\begin{aligned} F_n^L(x_0, ..., x_n) &= \sum_{k=0}^{n} \frac{y_0 \cdots y_{k-1} y_{k+1} \cdots y_n}{(y_k-y_0)...(y_k-y_{k-1})(y_k-y_{k+1})...(y_k-y_n)} (-1)^n g(y_k) \\ &= \sum_{k=0}^{n} \frac{y_0 \cdots y_{k-1} y_{k+1} \cdots y_n}{(y_k-y_0)...(y_k-y_{k-1})(y_k-y_{k+1})...(y_k-y_n)} (-1)^n x_k. \end{aligned}$$

Concerning the convergence of this method we state:

Theorem 9 *If $\alpha \in (a,b)$ is solution of equation (4), f' is bounded on (a,b), and the starting values satisfy $|\alpha - x_k| < 1/c, \quad k = 0, ..., n,$ with c =constant, then the sequence*

$$x_{i+1} = F_n^L(x_{n-i}, ..., x_i), \quad i = n, n+1, ...$$

converges to α.

Remark 10 *The order $ord(F_n^L)$ is the positive solution of the equation*

$$t^{n+1} - t^n - ... - t - 1 = 0.$$

Particular cases.

1) For $n = 1$, the nodes x_0, x_1, we get **the secant method**

$$F_1^L(x_0, x_1) = x_1 - \frac{(x_1 - x_0) f(x_1)}{f(x_1) - f(x_0)},$$

Thus,

$$x_{k+1} := F_1^L\left(x_{k-1}, x_k\right) = x_k - \frac{\left(x_k - x_{k-1}\right) f\left(x_k\right)}{f\left(x_k\right) - f\left(x_{k-1}\right)}, \quad k = 1, 2, ...$$

is the new approximation obtained using the previous approximations x_{k-1}, x_k.

The *order* of this method is the positive solution of equation:

$$t^2 - t - 1 = 0,$$

so $ord(F_1^L) = \frac{(1+\sqrt{5})}{2}$.

A modified form of the secant method: if we keep x_1 fixed and we change every time the same interpolation node, i.e.,

$$x_{k+1} = x_k - \frac{\left(x_k - x_1\right) f\left(x_k\right)}{f\left(x_k\right) - f\left(x_1\right)}, \quad k = 2, 3, \ldots .$$

2) For $n = 2$, the nodes x_0, x_1, x_2 and we get

$$F_2^L(x_0, x_1, x_2) = \frac{x_0 f(x_1) f(x_2)}{[f(x_0)-f(x_1)][f(x_0)-f(x_2)]} + \frac{x_1 f(x_0) f(x_2)}{[f(x_1)-f(x_0)][f(x_1)-f(x_2)]} + \frac{x_2 f(x_0) f(x_1)}{[f(x_2)-f(x_0)][f(x_2)-f(x_1)]}.$$

The *order* of this method is the positive solution of equation:

$$t^3 - t^2 - t - 1 = 0,$$

so $ord(F_2^L) = 1.8394$.

Comparing the Newton's method and secant method with respect to the time needed for finding a root with some given precision, we have:

-Newton's methods has more computation at one step: it is necessary to evaluete $f(x)$ and $f'(x)$. Secant method evaluate just $f(x)$ (supposing that $f(x_{previous})$ is stored.)

-The number of iterations for Newton's method is smaller (its order is $p_N = 2$). Secant method has order $p_S = 1.618$ and we have that three steps of this method are equivalent with two steps of Newton's method.

- It is proved that if the time for computing $f'(x)$ is greater than 0.44$\times$the time for computing $f(x)$, then the secant method is faster.

Remark 11 *The computation time is not the unique criterion in choosing the method! Newton's method is easier to apply. If $f(x)$ is explicitly known (for example, it is the solution the numerical integration of a differential equation), then its derivative it is calculated numerically. If we consider the following expression for the numerical calculation of derivative:*

$$f'(x) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} \tag{9}$$

then the Newton's method becomes the secant method.

Another way of obtaining secant method.

Based on approx. the function by a straight line connecting two points on the graph of f (not required f to have opposite signs at the initial points).

The first point, x_2, of the iteration is taken to be the point of intersection of the Ox-axis and the secant line connecting two starting points

$(x_0, f(x_0))$ and $(x_1, f(x_1))$. The next point, x_3, is generated by the intersection of the new secant line, joining $(x_1, f(x_1))$ and $(x_2, f(x_2))$ with the Ox-axis. The new point, x_3, together with x_2, is used to generate the next point, x_4, and so on.

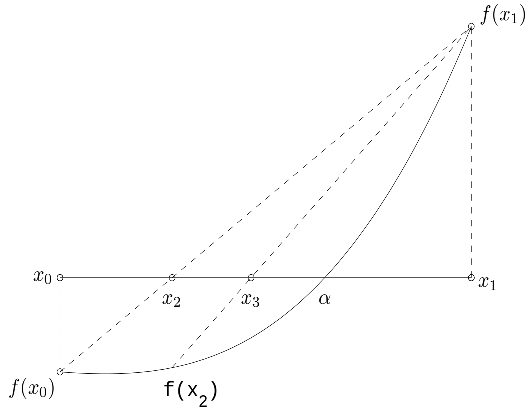
The formula for x_{n+1} is obtained by setting $x = x_{n+1}$ and $y = 0$ in the equation of the secant line from $(x_{n-1}, f(x_{n-1}))$ to $(x_n, f(x_n))$:

$$\frac{x - x_n}{x_{n-1} - x_n} = \frac{y - f(x_n)}{f(x_{n-1}) - f(x_n)} \Leftrightarrow x = x_n + \frac{(x_{n-1} - x_n)(y - f(x_n))}{f(x_{n-1}) - f(x_n)},$$

we get

$$x_{n+1} = x_n - f(x_n)\left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right]. \tag{10}$$

Note that x_{n+1} depends on the two previous elements of the sequence $\Rightarrow$ two initial guesses, x_0 and x_1,for generating $x_2, x_3, \ldots$.



The algorithm:

Let x_0 and x_1 be two initial approximations.

for $n = 1, 2, ..., ITMAX$

$$x_{n+1} \leftarrow x_n - f(x_n) \left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \right].$$

A suitable stopping criterion is

$$|f(x_n)| \leq \varepsilon \text{ or } \left| x_{n+1} - x_n \right| \leq \varepsilon \text{ or } \frac{\left| x_{n+1} - x_n \right|}{\left| x_{n+1} \right|} \leq \varepsilon,$$

where ε is a specified tolerance value.

Example 12 *Use the secant method with* $x_0 = 1$ *and* $x_1 = 2$ *to solve* $x^3 - x^2 - 1 = 0$, *with* $\varepsilon = 10^{-4}$.

Sol. *With* $x_0 = 1,\ f(x_0) = -1$ *and* $x_1 = 2,\ f(x_1) = 3$, *we have*

$$x_2 = 2 - \frac{(2-1)(3)}{3-(-1)} = 1.25$$

from which $f(x_2) = f(1.25) = -0.609375$. *The next iterate is*

$$x_3 = 1.25 - \frac{(1.25-2)(-0.609375)}{-0.609375-3} = 1.3766234.$$

Continuing in this manner the iterations lead to the approximation 1.4655713.

Examples of other multi-step methods

1. THE BISECTION METHOD

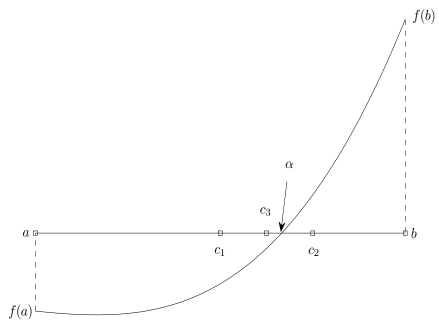
Let f be a given function, continuous on an interval $[a, b]$, such that

$$f(a)f(b) < 0. \tag{11}$$

By Mean Value Theorem, it follows that there exists at least one zero α of f in (a, b).

The bisection method is based on halving the interval $[a, b]$ to determine a smaller and smaller interval within α must lie.

First we give the midpoint of $[a, b]$, $c = (a + b)/2$ and then compute the product $f(c)f(b)$. If the product is negative, then the root is in the interval $[c, b]$ and we take $a_1 = c$, $b_1 = b$. If the product is positive, then the root is in the interval $[a, c]$ and we take $a_1 = a$, $b_1 = c$. Thus, a new interval containing α is obtained.



Bisection method

The algorithm:

Suppose $f(a)f(b) \le 0$. Let $a_o = a$ and $b_o = b$.

for $n = 0, 1, ...,$ITMAX

$c \leftarrow \frac{a_n + b_n}{2}$

if $f(a_n)f(c) \le 0$, set $a_{n+1} = a_n, b_{n+1} = c$

else, set $a_{n+1} = c, b_{n+1} = b_n$

The process of halving the new interval continues until the root is located as accurately as desired, namely

$$\frac{|a_n - b_n|}{|a_n|} < \varepsilon,$$

where a_n and b_n are the endpoints of the n-th interval $[a_n, b_n]$ and ε is a specified precision. The approximation of the solution will be $\frac{a_n+b_n}{2}$.

Some other stopping criterions: $|a_n - b_n| < \varepsilon$ or $|f(a_n)| < \varepsilon$.

Example 13 *The function $f(x) = x^3 - x^2 - 1$ has one zero in $[1, 2]$. Use the bisection algorithm to approximate the zero of f with precision 10^{-4}.*

Sol. *Since $f(1) = -1 < 0$ and $f(2) = 3 > 0$, then (11) is satisfied. Starting with $a_0 = 1$ and $b_0 = 2$, we compute*

$$c_0 = \frac{a_0 + b_0}{2} = \frac{1+2}{2} = 1.5 \text{ and } f(c_0) = 0.125.$$

Since $f(1.5)f(2) > 0$, the function changes sign on $[a_0, c_0] = [1, 1.5]$.

To continue, we set $a_1 = a_0$ *and* $b_1 = c_0$; *so*

$$c_1 = \frac{a_1 + b_1}{2} = \frac{1 + 1.5}{2} = 1.25 \text{ and } f(c_1) = -0.609375$$

Again, $f(1.25)f(1.5) < 0$ *so the function changes sign on* $[c_1, b_1] = [1.25,\ 1.5]$. *Next we set* $a_2 = c_1$ *and* $b_2 = b_1$. *Continuing in this manner we obtain a sequence* $(c_i)_{i>0}$ *which converges to* 1.465454, *the solution of the equation.*

2. THE METHOD OF FALSE POSITION

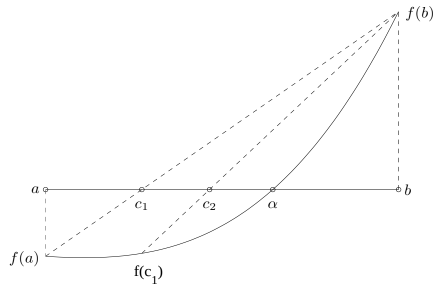
This method is also known as *regula falsi*, is similar to the Bisection method but has the advantage of being slightly faster than the latter. The function have to be continuous on $[a, b]$ with

$$f(a)f(b) < 0.$$

The point c is selected as point of intersection of the Ox-axis, and the straight line joining the points $(a,\ f(a))$ and $(b, f(b))$. From the equation of the secant line, it follows that

$$c = b - f(b)\frac{b-a}{f(b)-f(a)} = \frac{af(b)-bf(a)}{f(b)-f(a)} \tag{12}$$

Compute $f(c)$ and repeat the procedure between the values at which the function changes sign, that is, if $f(a)f(c) < 0$ set $b = c$, otherwise set $a = c$. At each step we get a new interval that contains a root of f and the generated sequence of points will eventually converge to the root.



Method of false position.

The algorithm:

Given a function f continuous on $[a_0, b_0]$, with $f(a_0)f(b_0) < 0$,

input: a_0, b_0

for $n = 0, 1, ..., ITMAX$

$c \leftarrow \frac{f(b_n)a_n - f(a_n)b_n}{f(b_n) - f(a_n)}$

if $f(a_n)f(c) < 0$, set $a_{n+1} = a_n, b_{n+1} = c$ **else** set $a_{n+1} = c, b_{n+1} = b_n$.

Stopping criterions: $|f(a_n)| \leq \varepsilon$ or $|a_n - a_{n-1}| \leq \varepsilon$, where ε is a specified tolerance value.

One of the main disadvantages of this method is that if the sequence of points generated by its algorithm is one-sided, the convergence of the method is slow.

Remark 14 *The bisection and the false position methods converge at a very low speed compared to the secant method.*

Example 15 *The function* $f(x) = x^3 - x^2 - 1$ *has one zero in* $[1, 2]$*. Use the method of false position to approximate the zero of* f *to within* 10^{-4}*.*

Sol. *A root lies in the interval* $[1, 2]$ *since* $f(1) = -1$ *and* $f(2) = 3$*. Starting with* $a_0 = 1$ *and* $b_0 = 2$*, we get using (12)*

$$c_0 = 2 - \frac{3(2-1)}{3-(-1)} = 1.25 \text{ and } f(c_0) = -0.609375.$$

Here, $f(c_0)$ has the same sign as $f(a_0)$ and so the root must lie on the interval $[c_0, b_0] = [1.25, 2]$. Next we set $a_1 = c_0$ and $b_1 = b_0$ to get the next approximation

$$c_1 = 2 - \frac{3 - (2 - 1.25)}{3 - (-0.609375)} = 1.37662337 \text{ and } f(c_1) = -0.2862640.$$

Now $f(x)$ change sign on $[c_1, b_1] = [1.37662337, 2]$. Thus we set $a_2 = c_1$ and $b_2 = b_{1.}$ Continuing in this manner the iterations lead to the approximation 1.465558.

Example 16 *Compare the false position method, the secant method and Newton's method for solving the ecuation $x = cosx$, having as starting points $x_0 = 0.5$ și $x_1 = \pi/4$, respectively $x_0 = \pi/4$.*

n	(a) x_n False position	(b) x_n Secant	(c) x_n Newton
0	0.5	0.5	0.785398163397
1	0.785398163397	0.785398163397	0.739536133515
2	0.736384138837	0.736384138837	0.739085178106
3	0.739058139214	0.739058139214	0.739085133215
4	0.739084863815	0.739085149337	0.739085133215
5	0.739085130527	0.739085133215	
6	0.739085133188	0.739085133215	
7	0.739085133215		

The extra condition from the false position method usually requires more computation than the secant method, and the simplifications from the secant method come with more iterations than in the case of Newton's method.