MA2501 Numerical Methods
Spring 2016

Solutions to exercise set 5

Norwegian University of Science and Technology
Department of Mathematics

$\boxed{1}$ Assume that the *regula falsi* method (without modifications) is used for finding the solution of the equation $x^3 = 0$.

a) Show that during all the iterations one endpoint of the solution interval remains unchanged, unless the iteration finds the solution after the first step.

b) Denote by c_k the result of the method at the k-th step. Show that

$$\lim_{k\to\infty} \frac{|c_{k+1}|}{|c_k|} = 1. \tag{1}$$

Note: It can be proven that the method of false position converges to the root, so you can assume $\lim_{k\to\infty} c_k = 0$.

Remark: A sequence $(c_k)_{k\in\mathbb{N}}$ converging to 0 and satisfying (1) is said to converge *sublinearly*. Sublinear converge of a sequence means that the number of iterations it takes to come closer to the limit by even one digit will become arbitrarily large. Thus numerical methods that have such sequences as output should usually be avoided.

Possible solution:

a) First we note that the equation $x^3 = 0$ has the unique solution $x = 0$. Thus, unless one of the iterates becomes 0, we will always have the inequality $a < 0 < b$. Moreover we note that the result of the method given the input $[a, b]$ is

$$c = \frac{ab^3 - ba^3}{b^3 - a^3} = ab\frac{b^2 - a^2}{b^3 - a^3}.$$

Obviously $c = 0$ if and only if $a = -b$.

Assume now that $a \neq -b$. Then either of the inequalities $a^2 < b^2$ or $a^2 > b^2$ holds. Assume first that $a^2 < b^2$. Then, since $a < 0 < b$ and therefore $ab < 0$, we have

$$c = ab\frac{b^2 - a^2}{b^3 - a^3} < 0,$$

implying that in the next step a is replaced by c. Moreover this implies that also in the next step the inequality $a^2 < b^2$ holds, and thus, again, it is the left endpoint of the interval that is updated.

An analogous argumentation shows that in the case $a^2 > b^2$ it is always the point b that gets replaced by c.

b) Assume without loss of generality that at the start of the iteration we have $a^2 < b^2$ (the case where $a^2 > b^2$ can be treated similarly). Then the previous considerations show that the sequence c_k is defined by the iteration $c_0 = a_0$ and

$$c_{k+1} = c_k b \frac{b^2 - c_k^2}{b^3 - c_k^3}.$$

In particular,

$$\frac{c_{k+1}}{c_k} = b \frac{b^2 - c_k^2}{b^3 - c_k^3}.$$

From the assumption we may assume that $c_k \to 0$. Thus

$$\lim_{k\to\infty} \frac{|c_{k+1}|}{|c_k|} = \lim_{k\to\infty} b \frac{b^2 - c_k^2}{b^3 - c_k^3} = 1.$$

$\boxed{2}$ Apply fixed point iteration to the solution of the equation $\cos(x) = \frac{1}{2}\sin(x)$ using the mapping

$$\Phi(x) = x + \cos(x) - \frac{1}{2}\sin(x).$$

a) Show that the mapping Φ is a contraction on $[0, \pi/2]$ (note that you also have to show that $\Phi(x) \in [0, \pi/2]$ for all $x \in [0, \pi/2]$).

b) Compute the first five iterates of the fixed point iteration using the starting value $x^{(0)} = 0$.

c) Provide an estimate of the accuracy of the outcome of the method after the 5$^{\text{th}}$ iteration. How many iterations will be needed to obtain a result with an error smaller than 10^{-12}?

Possible solution:

a) Since $x \geq \sin(x)$ for $x \geq 0$ and $\cos(x) \geq 0$ on $[0, \pi/2]$, it follows that $\Phi(x) \geq 0$ for $x \in [0, \pi/2]$. Moreover $\sin(x) \geq 0$ on $[0, \pi/2]$, implying that $\Phi(x) \leq x + \cos(x)$ for all $x \in [0, \pi/2]$. Since the maximum of the function $x \mapsto x + \cos(x)$ on $[0, \pi/2]$ is $\pi/2$, this shows that, indeed, Φ maps $[0, \pi/2]$ to $[0, \pi/2]$.

Next note that the function Φ is differentiable with

$$\Phi'(x) = 1 - \sin(x) - \frac{1}{2}\cos(x).$$

Obviously we have $\Phi'(x) \geq 1 - 1 - 1/2 = -1/2$ for every $x \in \mathbb{R}$. In addition, the function Φ' is strictly convex on $[0, \pi/2]$ (since sin and cos are concave on $[0, \pi/2]$, and therefore the function Φ' does not have local maxima in $(0, \pi/2)$. As a consequence,

$$\Phi'(x) \leq \max\{\Phi'(0), \Phi'(\pi/2)\} = \max\{1/2, 0\} = \frac{1}{2}$$

for all $x \in [0, \pi/2]$. Thus we have shown that $|\Phi'(x)| \leq 1/2$ for all $x \in [0, \pi/2]$, showing that Φ is a contraction on $[0, \pi/2]$ with contraction factor $1/2$.

b) The first five iterates are

$$\begin{aligned} x^{(0)} &= 0, \\ x^{(1)} &= 1, \\ x^{(2)} &\approx 1.1196, \\ x^{(3)} &\approx 1.1057, \\ x^{(4)} &\approx 1.1073, \\ x^{(5)} &\approx 1.1071. \end{aligned}$$

c) For estimating the accuracy of the iterates, we use the inequality

$$\left| x^{(k)} - x^* \right| \le \frac{C}{1-C} \left| x^{(k)} - x^{(k-1)} \right|$$

with $k = 5$ and $C = 1/2$ (the contraction factor). Thus we obtain

$$\left| x^{(5)} - x^* \right| \le \left| x^{(5)} - x^{(4)} \right| \approx 2 \cdot 10^{-4}.$$

In order to estimate the number of iterations that are needed for obtaining an error smaller than 10^{-12}, we use the inequality

$$\left| x^{(k)} - x^* \right| \le \frac{C^{k-4}}{1-C} \left| x^{(5)} - x^{(4)} \right| = \frac{1}{2^{k-5}} \left| x^{(5)} - x^{(4)} \right| \le \frac{1}{2^{k-6}} 10^{-4}.$$

Since we want the error to be smaller than 10^{-12}, we obtain from this the condition

$$2^{k-6} \ge 10^8$$

or

$$k \ge 6 + \log_2(10^8) \approx 32.6.$$

Thus we would need 33 iterations.[1]

$\boxed{3}$ Find an approximation of a solution of the equation $x^3 - 2x - 5 = 0$ by applying three steps of:

a) the bisection method with starting interval $[2, 3]$,

b) the secant method with starting values $x^{(0)} = 3$ and $x^{(1)} = 3.5$,

c) the Newton method with starting value $x^{(0)} = 3$.

Possible solution:

a) For the bisection method we obtain the intervals given by

$$\begin{aligned} a &= 2, & b &= 3, \\ a &= 2, & b &= 2.5, \\ a &= 2, & b &= 2.25, \\ a &= 2, & b &= 2.125. \end{aligned}$$

[1] All these computations were made with the contraction factor 1/2. It is, however, possible to derive a much smaller one on a small interval around the solution. Thus, in fact, we are vastly overestimating the error and the required number of iterations.

b) For the secant method we obtain

$$\begin{aligned} x^{(0)} &= 3, \\ x^{(1)} &= 3.5, \\ x^{(2)} &\approx 2.4622, \\ x^{(3)} &\approx 2.2615, \\ x^{(4)} &\approx 2.1229. \end{aligned}$$

c) Newton's method yields

$$\begin{aligned} x^{(0)} &= 3, \\ x^{(1)} &= 2.36, \\ x^{(2)} &\approx 2.1271, \\ x^{(3)} &\approx 2.0951. \end{aligned}$$

The actual solution is

$$x^* \approx 2.09455.$$

$\boxed{4}$ Compute the first three steps of the Newton method for the solution of the system of equations

$$\begin{aligned} -5x + 2\sin(x) + \cos(y) &= 0, \\ 4\cos(x) + 2\sin(y) - 5y &= 0 \end{aligned}$$

with initial value $(x^{(0)}, y^{(0)}) = (0, 0)$ (you may want to use MATLAB for solving the linear systems).

Possible solution:

For the application of Newton's method we denote

$$F(x, y) := \begin{pmatrix} -5x + 2\sin(x) + \cos(y) \\ 4\cos(x) + 2\sin(y) - 5y \end{pmatrix}.$$

Then the Jacobian of F is

$$\mathbf{J_F}(x, y) = \begin{pmatrix} -5 + 2\cos(x) & -\sin(y) \\ -4\sin(x) & 2\cos(y) - 5 \end{pmatrix}.$$

Its inverse can be calculated analytically as

$$\mathbf{J_F}(x, y)^{-1} = \frac{1}{\det \mathbf{J_F}(x, y)} \begin{pmatrix} 2\cos(y) - 5 & \sin(y) \\ 4\sin(x) & -5 + 2\cos(x) \end{pmatrix}$$

with

$$\det \mathbf{J_F}(x, y) = 25 - 10\cos(x) - 10\cos(y) + 4\cos(x)\cos(y) - 4\sin(x)\sin(y).$$

Newton's method now defines iteratively

$$\begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \end{pmatrix} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \end{pmatrix} - \mathbf{J_F}(x^{(k)}, y^{(k)})^{-1} F(x^{(k)}, y^{(k)}).$$

Starting with $(x^{(0)}, y^{(0)}) = (0, 0)$, we thus obtain the iterates

$$\begin{aligned} \begin{pmatrix} x^{(1)} \\ y^{(1)} \end{pmatrix} &= \frac{1}{3} \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \\ \begin{pmatrix} x^{(2)} \\ y^{(2)} \end{pmatrix} &\approx \begin{pmatrix} 0.1302 \\ 1.1838 \end{pmatrix}, \\ \begin{pmatrix} x^{(3)} \\ y^{(3)} \end{pmatrix} &\approx \begin{pmatrix} 0.1330 \\ 1.1597 \end{pmatrix}. \end{aligned}$$

$\boxed{5}$ Implement in MATLAB Newton's method for the solution of an equation $f(x) = 0$ with $f\colon \mathbb{R} \to \mathbb{R}$. Your function should take as an input the function f, its derivative f', and a starting value $x^{(0)}$. Test your implementation on the functions in example $\boxed{3}$ of this exercise set.

MATLAB-comments

There are basically two methods to pass functions as arguments to another function. The first possibility is to pass the *name* of the function as a *string* (that is, the name written in single quotes). Then the command `feval` can be used for evaluating the function. Alternatively, you can pass a *function handle*, which is an argument of the form `@f`, where `f` is the function. Then you can evaluate the function either directly with `f(x)` or, again, using `feval`.

Consider for example the function

```
function y = myfun(f)
y = feval(f,1);
end
```

If you then call `myfun('sin')`, you obtain as result sin(1). The same result will be obtained using the function call `myfun(@sin)`.

Possible solution:

You find some example code on the webpage of the course.