Last updated: 21 July 2015
Disclaimer: Links are followed at own risk when viewing this document electronically.

Chapter 3

NONLINEAR EQUATIONS

3.1 Introduction

Generally speaking, nonlinear equations and, in particular, so-called *transcendental* equations, cannot be solved *analytically*, and so may *only* be solved numerically. We consider the solution of nonlinear equations in one variable, although we briefly describe a method for a system of two nonlinear equations at the end of this chapter.

3.2 The bisection method

Consider the equation $f(x) = 0$, where f is assumed to be continuous. By the intermediate value theorem, if an interval $[x_1, x_2]$ can be found on which f changes sign, i.e.

$$f(x_1)f(x_2) < 0,$$

then f *has at least one real root* on the interval.

Assume that f has only one root on $[x_1, x_2]$, denote this root as x_0. The interval is now halved by determining

$$x_3 = \frac{x_1 + x_2}{2}.$$

This value x_3 may be regarded as an *approximation* to x_0. The approximation may be improved by *iterating* (repeating) the halving process. At each iteration there are two possibilities:

(a) $f(x_1)f(x_3) < 0$. The root thus lies on $[x_1, x_3]$, and x_2 is replaced by x_3.

(b) $f(x_3)f(x_2) < 0$. The root lies on $[x_3, x_2]$, and x_1 is replaced by x_3.

The halving process is iterated until a specified *accuracy* ϵ *in the function value* is reached, that is

$$|f(x_3)| < \epsilon. \tag{3.1}$$

Convergence. Since any curve on a small enough interval may be approximated by a straight line, we have that $f(x_3)$ will converge to $f(x_0)$, in the vicinity of x_0, at the same rate that

$$\eta = |x_0 - x_3|$$

converges to zero. For the nth iteration we have

$$\eta_n \approx \frac{\eta_{n-1}}{2}.$$

The bisection method is thus said to be *linearly* convergent.

Example 3.1. Solve $f(x) = \sin x - 0.625x = 0$ to an accuracy of 8 decimal places in the function value.

We show the values for the first four iterations in the following table:

iteration	x_1	x_3	x_2	$f(x_1)$	$f(x_3)$	$f(x_2)$
1	1	1.5	2	+0.216	+0.060	−0.341
2	1.5	1.75	2	+0.060	−0.110	−0.341
3	1.5	1.625	1.75	+0.060	−0.171	−0.110
4	1.5	1.5625	1.625	+0.060	+0.023	−0.171

After 25 iterations we find

$$x = 1.59934789 \text{ rad } (91.635884^\circ)$$

3.3 Linear interpolation

Although the bisection method is reliable, it is also slow. This is because very little information about $f(x)$ is used—indeed, only the sign of f is used. In the linear interpolation method we also make use of the *numerical values* of $f(x)$.

Consider the equation $f(x) = 0$. Let (x_1, y_1) and (x_2, y_2) be two points on the curves $y = f(x)$ in the vicinity of the root $x = x_0$. We approximate the curve in this region by a *straight line* through the two points. The zero point x_3 of the straight line may be regarded as an approximation to the root x_0.

The value of x_3 is obtained from the equation for the straight line

$$y - y_1 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1).$$

The point $(x_3, 0)$ lies on this line so that

$$0 - y_1 = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x_3 - x_1)$$

and so

$$x_3 = \frac{x_1 y_2 - x_2 y_1}{y_2 - y_1}.$$

The interpolation process is now iterated by finding the straight line through (x_2, y_2) and (x_3, y_3) and hence the approximation x_4, and so on. In general, the $(i-1)$th iteration uses (x_{i-1}, y_{i-1}) and (x_i, y_i) to find x_{i+1}:

$$x_{i+1} = \frac{x_{i-1} y_i - x_i y_{i-1}}{y_i - y_{i-1}}. \tag{3.2}$$

For a given accuracy ϵ this process is repeated until the condition (3.1) is satisfied.

Convergence. Let the error in x after the $(i-1)$th iteration be denoted by η_i. In other words, $x_i = x_0 + \eta_i$. Then from equation (3.2) it follows that

$$x_0 + \eta_{i+1} = \frac{(x_0 + \eta_{i-1}) y_i - (x_0 + \eta_i) y_{i-1}}{y_i - y_{i-1}}$$

and so

$$\eta_{i+1} = \frac{\eta_{i-1} f(x_0 + \eta_i) - \eta_i f(x_0 + \eta_{i-1})}{f(x_0 + \eta_i) - f(x_0 + \eta_{i-1})}.$$

Each function may be expanded in a Taylor series about x_0:

$$\begin{aligned} f(x_0 + \eta_i) &= f(x_0) + \eta_i f'(x_0) + \frac{1}{2}\eta_i^2 f''(x_0) + \cdots \\ &= 0 + \eta_i f'(x_0) + \frac{1}{2}\eta_i^2 f''(x_0) + \cdots \end{aligned}$$

We assume that $f'(x_0) \neq 0$ and $f''(x_0) \neq 0$. Since η_i is small (by assumption) we have, to the *lowest order* in η_i

$$\begin{aligned}
\eta_{i+1} &\approx \frac{\eta_{i-1}\left[\eta_i f'(x_0) + \frac{1}{2}\eta_i^2 f''(x_0)\right] - \eta_i\left[\eta_{i-1} f'(x_0) + \frac{1}{2}\eta_{i-1}^2 f''(x_0)\right]}{\eta_i f'(x_0) - \eta_{i-1} f'(x_0)} \\
&= \frac{1}{2}\left(\frac{\eta_i\eta_{i-1}\left(\eta_i - \eta_{i-1}\right) f''(x_0)}{\left(\eta_i - \eta_{i-1}\right) f'(x_0)}\right) \\
&= \left(\frac{f''(x_0)}{2f'(x_0)}\right)\eta_i\eta_{i-1} \\
&\equiv A\eta_i\eta_{i-1}, \quad \text{where } A \text{ is a constant.}
\end{aligned} \tag{3.3}$$

We attempt to satisfy this relationship by assuming

$$\eta_i = K\eta_{i-1}^a \quad \Longleftrightarrow \quad \eta_{i+1} = K\eta_i^a$$

where K is some constant, so that, from (3.3), we have

$$\eta_{i+1} \approx A\eta_i\left(\frac{1}{K}\eta_i\right)^{\frac{1}{a}}.$$

But

$$\eta_{i+1} = K\eta_i^a$$

from our earlier assumption, and since the powers of η_i must be the same on both sides of the equation, it follows that

$$1 + \frac{1}{a} = a$$

which gives

$$a^2 - a - 1 = 0$$

which has the root

$$a = \frac{1+\sqrt{5}}{2} \approx 1.618$$

which is the also called the *golden mean number*. Hence

$$\eta_{i+1} \approx K\eta_i^{1.618}.$$

This rate of convergence is termed *superlinear*. (The other root $a \approx -0.62$ corresponds to divergence.)

Example 3.2. Solve $f(x) = \sin x - 0.625x = 0$ accurate to 8 decimal places in the function value.

The values for the first 3 iterations are shown in the following table.

i	x_i	x_{i+1}	y_i	y_{i+1}	x_{i+2}	y_{i+2}
1	1	2	0.2165	−0.3407	1.3885	0.1156
2	2	1.3885	−0.3407	0.1156	1.5434	0.0350
3	1.3885	1.5434	0.1156	0.0350	1.6106	−0.0074

After 6 iterations we obtain $x_0 = 1.59934789$ rad (compare with 25 iterations for the bisection method).

3.4 Newton's method

For this method we assume that f is at least twice differentiable. We now use both the values $f(x)$ and its first derivative to solve $f(x) = 0$. Let x_1 be an initial estimate of the root x_0, and draw a tangent line at the point $(x_1, f(x_1))$ to the curve of f. The x-intercept of the tangent line, denote it x_2, is presumably a better estimate of x_0 than x_1 was. The slope of the tangent line is $f'(x_1)$ and is determined by

$$f'(x_1) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

However, since x_2 is the x-intercept of the tangent line, we have $f(x_2) = 0$ and thus

$$f'(x_1) = \frac{f(x_1)}{x_1 - x_2}.$$

Solving this last equation for x_2 we obtain

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

This process can be repeated and we find, in general, after the ith iteration we have

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}. \tag{3.4}$$

Note: The above result could also have been obtained by calculating the error η for an initial estimate x_1 by solving

$$f(x_1 + \eta) = 0.$$

We expand the LHS in a Taylor series about $x = x_1$

$$f(x_1) + \eta f'(x_1) + \frac{\eta^2}{2!} f''(x_1) + \cdots = 0.$$

For small η, $f(x_1) + \eta f'(x_1) \approx 0$ and so

$$\eta \approx -\frac{f(x_1)}{f'(x_1)}.$$

Newton's method is thus equivalent to a first-order Taylor expansion of the function $f(x)$.

Convergence. Let the error in x after the ith iteration be denoted by η_i. In other words, $x_i = x_0 + \eta_i$. From equation (3.4)

$$x_0 + \eta_{i+1} = x_0 + \eta_i - \frac{f(x_0 + \eta_i)}{f'(x_0 + \eta_i)}.$$

A Taylor expansion about $x = x_0$ for both $f(x_0 + \eta_i)$ and $f'(x_0 + \eta_i)$ gives

$$\eta_{i+1} = \eta_i - \frac{f(x_0) + \eta_i f'(x_0) + \frac{1}{2}\eta_i^2 f''(x_0) + \cdots}{f'(x_0) + \eta_i f''(x_0) + \cdots}.$$

Expanding the denominator using

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

we find

$$\begin{aligned}
\frac{1}{f'(x_0) + \eta_i f''(x_0) + O(\eta_i^2)} &= \frac{1}{f'(x_0)\left(1 + \eta_i \frac{f''(x_0)}{f'(x_0)} + O(\eta_i^2)\right)} \\
&= \frac{1}{f'(x_0)} \, \frac{1}{1 + \eta_i \frac{f''(x_0)}{f'(x_0)} + O(\eta_i^2)} \\
&= \frac{1}{f'(x_0)} \left(1 - \eta_i \frac{f''(x_0)}{f'(x_0)} + O(\eta_i^2)\right)
\end{aligned}$$

To lowest order in η we have

$$\begin{aligned}\eta_{i+1} &\approx \eta_i - \frac{1}{f'(x_0)}\left(1 - \eta_i \frac{f''(x_0)}{f'(x_0)}\right)\left(0 + \eta_i f'(x_0) + \frac{\eta_i^2}{2} f''(x_0)\right) \\ &\approx \left(\frac{f''(x_0)}{2f'(x_0)}\right)\eta_i^2. \end{aligned} \tag{3.5}$$

The convergence is said to be *quadratic.*

It is important to note that (3.3) and (3.5) give estimates for the convergence rate *only if convergence actually occurs.* It is possible that both the linear interpolation method and Newton's method may diverge.

Example 3.3. Find the root of $f(x) = \sin x - 0.625x = 0$ correct to 8 decimal places in the function value.

Equation (3.4) is iterated using

$$f(x) = \sin x - 0.625x$$

and

$$f'(x) = \cos x - 0.625$$

with an initial estimate $x = 1.5$. The results of the process are shown in the following table:

i	x_i	$f(x_i)$	$f'(x_i)$	$\Delta x_i = -\frac{f(x_i)}{f'(x_i)}$
1	1.5	0.060	-0.554	0.108
2	1.608	-0.0059	-0.662	-0.00884
3	1.59941	-3.9×10^{-5}	-0.654	-6.0×10^{-5}
4	1.59934789	-2.1×10^{-9}		

After 4 iterations the root $x = 1.59934789$ rad is found.

Calculation hints

(a) Sometimes f' is a cumbersome analytical expression. In such cases f' may be estimated *numerically* by

$$f'(x_i) \approx \frac{f(x_i + h) - f(x_i)}{h}$$

where h is small. One should be aware of complications that may arise from numerical differentiation (see chapter 6 for better approximations. Linear interpolation is indeed Newton's method with simple numerical differentiation).

(b) When $f(x) = 0$ has more than one root, it is sometimes found that both linear interpolation and Newton's method converge to the same root, for any initial estimate. Say x_0 is such a root. Then we may write

$$f(x) = (x - x_0)g(x).$$

Once the root x_0 is found, it may be eliminated from the problem by subsequently solving

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

3.5 Fixed-point iteration

If an equation can be written in the form

$$x = g(x) \tag{3.6}$$

then a root of the equation may be regarded as a fixed point that is mapped to itself under the map $x = g(x)$. The iteration process

$$x_{i+1} = g(x_i) \tag{3.7}$$

may then be used, *possibly*, to find the root.

Example 3.4. Find the positive root of $x = 2\sin x$ correct to 8 decimal places.
The following table shows 6 iterations of (3.7) with $g(x) = 2\sin x$, using initial value $x = 2$:

i	x_i	$g(x_i)$
1	2	1.82
2	1.82	1.94
3	1.94	1.87
4	1.87	1.91
5	1.91	1.88
6	1.88	1.90

The required accuracy is actually achieved after 39 iterations and we find $x_0 = 1.89549427$.

A transcendental equation of the form (3.6) can often be written in this form in several ways. For example, $e^x = 3x^2$ may be written as

$$x = \ln 3x^2$$

or

$$x = \pm\sqrt{\frac{e^x}{3}}.$$

Convergence.

We now obtain an analytical condition for convergence. Consider the error in x after the ith iteration,

$$x_{i+1} - x_0 = g(x_i) - g(x_0)$$

where x_0 is the root of (3.6). From the mean-value theorem of differential calculus we have

$$\frac{g(x_i) - g(x_0)}{x_i - x_0} = g'(\xi_i)$$

where $x_0 < \xi_i < x_i$. We can thus rewrite the ith error as

$$x_{i+1} - x_0 = g'(\xi_i)(x_i - x_0).$$

Iterating this recursion relation i times gives

$$x_{i+1} - x_0 = g'(\xi_i)g(\xi_{i-1})\cdots g'(\xi_1)(x_1 - x_0).$$

Now let

$$m = \max\left(|g'(\xi_i)|\right).$$

Then we have

$$|x_{i+1} - x_0| \leq m^i\,|x_1 - x_0|\,.$$

For convergence we must have that $|x_{i+1} - x_0| \to 0$ as $i \to \infty$, which will only be true if $m < 1$.

Condition for convergence: For the iteration scheme (3.7) to converge, we require

$$|g'(x)| < 1 \tag{3.8}$$

in the neighbourhood of the root.

Example 3.5. We investigate the convergence for the case in example 3.4.

For $x = g(x) = 2\sin x$ we have $g'(x) = 2\cos x$. Hence, for convergence we require

$$|\cos x| < \frac{1}{2}$$

and so we obtain an interval of convergence

$$1.047 < x < 2.094$$

The root $x_0 = 1.895$ indeed lies on this interval.

Note: Condition (3.8) may be applied to any iterative method that can be written in the form (3.6). Newton's method may be written as $x = g(x)$ with

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

and so

$$g'(x) = 1 - \frac{\big(f'(x)\big)^2 - f(x)f''(x)}{\big(f'(x)\big)^2}.$$

Then we must have

$$\left|\frac{f(x)f''(x)}{\big(f'(x)\big)^2}\right| < 1$$

in the neighbourhood of the root.

3.6 Systems of nonlinear equations

We describe Newton's method for two simultaneous nonlinear equations. The extension to systems of more than two equations can be found in most standard texts.

Consider two equations in unknowns x and y:

$$\begin{aligned} f_1(x,y) &= 0 \\ f_2(x,y) &= 0 \end{aligned} \tag{3.9}$$

As for the one-dimensional Newton's method, we try to find the corrections (η_x, η_y) to an initial guess (x_1, y_1) such that (3.9) is satisfied:

$$\begin{aligned} f_1(x_1+\eta_x, y_1+\eta_y) &= 0 \\ f_2(x_1+\eta_x, y_1+\eta_y) &= 0 \end{aligned}$$

A first-order Taylor expansion yields

$$\begin{aligned} f_1(x_1,y_1) + \eta_x\left(\left.\frac{\partial f_1}{\partial x}\right|_{(x_1,y_1)}\right) + \eta_y\left(\left.\frac{\partial f_1}{\partial y}\right|_{(x_1,y_1)}\right) &\approx 0 \\ f_2(x_1,y_1) + \eta_x\left(\left.\frac{\partial f_2}{\partial x}\right|_{(x_1,y_1)}\right) + \eta_y\left(\left.\frac{\partial f_2}{\partial y}\right|_{(x_1,y_1)}\right) &\approx 0 \end{aligned}$$

which may be written as

$$\begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}_{(x_1,y_1)} \begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix} \approx \begin{bmatrix} -f_1(x_1,y_1) \\ -f_2(x_1,y_1) \end{bmatrix}$$

so that

$$\begin{bmatrix} \eta_x \\ \eta_y \end{bmatrix} \approx \left(\begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix}_{(x_1,y_1)} \right)^{-1} \begin{bmatrix} -f_1(x_1,y_1) \\ -f_2(x_1,y_1) \end{bmatrix}$$

The square matrix here is known as a *Jacobian matrix*, and is evaluated at the point (x_1, y_1).