Rootfinding for Nonlinear Equations

Calculating the roots of an equation

$$f(x) = 0 \tag{7.1}$$

is a common problem in applied mathematics.

We will

- explore some simple numerical methods for solving this equation, and also will
- consider some possible difficulties

The function $f(x)$ of the equation (7.1)

- will usually have at least one continuous derivative, and often
- we will have some estimate of the root that is being sought.

By using this information, most numerical methods for (7.1) compute

a sequence of increasingly accurate estimates of the root.

These methods are called **iteration methods**.

We will study three different methods

1. the bisection method
2. Newton's method
3. secant method

and give a general theory for one-point iteration methods.

In this chapter we assume that $f : \mathbb{R} \to \mathbb{R}$ i.e.,
$f(x)$ is a function that is real valued and that x is a real variable.
Suppose that

- $f(x)$ is continuous on an interval $[a, b]$, and
- $$f(a)f(b) < 0 \qquad (7.2)$$

Then $f(x)$ changes sign on $[a, b]$, and $f(x) = 0$ has at least one root on the interval.

Definition

The simplest numerical procedure for finding a root is to repeatedly halve the interval $[a, b]$, keeping the half for which $f(x)$ changes sign. This procedure is called the **bisection method**, and is guaranteed to converge to a root, denoted here by α.

Suppose that we are given an interval $[a, b]$ satisfying (7.2) and an error tolerance $\varepsilon > 0$.
The bisection method consists of the following steps:

B1 Define $c = \frac{a+b}{2}$.

B2 If $b - c \leq \varepsilon$, then accept c as the root and stop.

B3 If $\text{sign}[f(b)] \cdot \text{sign}[f(c)] \leq 0$, then set $a = c$.
Otherwise, set $b = c$. Return to step B1.

The interval $[a, b]$ is halved with each loop through steps B1 to B3.
The test B2 will be satisfied eventually, and with it the condition $|\alpha - c| \leq \varepsilon$ will be satisfied.

<u>Notice</u> that in the step B3 we test the sign of $\text{sign}[f(b)] \cdot \text{sign}[f(c)]$ in order to avoid the possibility of underflow or overflow in the multiplication of $f(b)$ and $f(c)$.

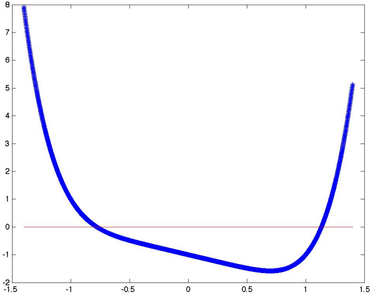
Example

Find the largest root of

$$f(x) \equiv x^6 - x - 1 = 0 \tag{7.3}$$

accurate to within $\varepsilon = 0.001$.

With a graph, it is easy to check that $1 < \alpha < 2$



We choose $a = 1, b = 2$; then $f(a) = -1, f(b) = 61$, and (7.2) is satisfied.

Use bisect.m

The results of the algorithm B1 to B3:

n	a	b	c	$b-c$	$f(c)$
1	1.0000	2.0000	1.5000	0.5000	8.8906
2	1.0000	1.5000	1.2500	0.2500	1.5647
3	1.0000	1.2500	1.1250	0.1250	-0.0977
4	1.1250	1.2500	1.1875	0.0625	0.6167
5	1.1250	1.1875	1.1562	0.0312	0.2333
6	1.1250	1.1562	1.1406	0.0156	0.0616
7	1.1250	1.1406	1.1328	0.0078	-0.0196
8	1.1328	1.1406	1.1367	0.0039	0.0206
9	1.1328	1.1367	1.1348	0.0020	0.0004
10	1.1328	1.1348	1.1338	0.00098	-0.0096

Table: Bisection Method for (7.3)

The entry n indicates that the associated row corresponds to iteration number n of steps B1 to B3.

Let a_n, b_n and c_n denote the n^{th} computed values of a, b and c:

$$b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n), \qquad n \geq 1$$

and

$$b_n - a_n = \frac{1}{2^{n-1}}(b - a) \tag{7.4}$$

where $b - a$ denotes the length of the original interval with which we started.
Since the root $\alpha \in [a_n, c_n]$ or $\alpha \in [c_n, b_n]$, we know that

$$|\alpha - c_n| \leq c_n - a_n = b_n - c_n = \frac{1}{2}(b_n - a_n) \tag{7.5}$$

This is the error bound for c_n that is used in step B2.
Combining it with (7.4), we obtain the further bound

$$\boldsymbol{|\alpha - c_n| \leq \frac{1}{2^n}(b - a).}$$

This shows that the iterates $\boldsymbol{c_n \to \alpha}$ as $n \to \infty$.

To see how many iterations will be necessary, suppose we want to have

$$|\alpha - c_n| \leq \varepsilon$$

This will be satisfied if

$$\frac{1}{2^n}(b-a) \leq \varepsilon$$

Taking logarithms of both sides, we can solve this to give

$$n \geq \frac{\log\left(\frac{b-a}{\varepsilon}\right)}{\log 2}$$

For the previous example (7.3), this results in

$$n \geq \frac{\log\left(\frac{1}{0.001}\right)}{\log 2} \doteq 9.97$$

i.e., we need $n = 10$ iterates, exactly the number computed.

There are several advantages to the bisection method

- It is guaranteed to converge.
- The error bound (7.5) is guaranteed to decrease by one-half with each iteration

Many other numerical methods have variable rates of decrease for the error, and these may be worse than the bisection method for some equations.

The principal disadvantage of the bisection method is that

- generally converges more slowly than most other methods.

For functions $f(x)$ that have a continuous derivative, other methods are usually faster. These methods may not always converge; when they do converge, however, they are almost always much faster than the bisection method.

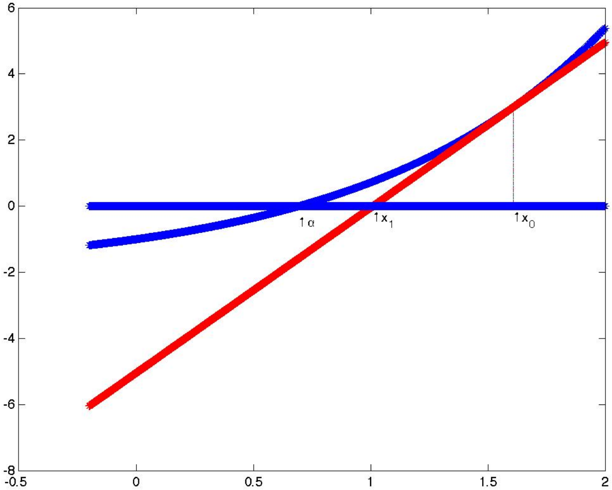


Figure: The schematic for Newton's method

There is usually an estimate of the root α, denoted x_0.
To improve it, consider the tangent to the graph at the point $(x_0, f(x_0))$.
If x_0 is near α, then the tangent line $\approx$ the graph of $y = f(x)$ for points about α.
Then the root of the tangent line should nearly equal α, denoted x_1.

The line tangent to the graph of $y = f(x)$ at $(x_0, f(x_0))$ is the graph of the linear Taylor polynomial:

$$p_1(x) = f(x_0) + f'(x_0)(x - x_0)$$

The root of $p_1(x)$ is x_1:

$$f(x_0) + f'(x_0)(x_1 - x_0) = 0$$

i.e.,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Since x_1 is expected to be an improvement over x_0 as an estimate of α, we repeat the procedure with x_1 as initial guess:

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

Repeating this process, we obtain a sequence of numbers, **iterates**, $x_1, x_2, x_3, \ldots$ hopefully approaching the root α.

The **iteration formula**

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n = 0, 1, 2, \ldots \tag{7.6}$$

is referred to as the **Newton's method**, or **Newton-Raphson**, for solving $f(x) = 0$.

Example

Using Newton's method, solve (7.3) used earlier for the bisection method.

Here

$$f(x) = x^6 - x - 1, \qquad f'(x) = 6x^5 - 1$$

and the iteration

$$x_{n+1} = x_n - \frac{x_n^6 - x_n - 1}{6x_n^5 - 1}, \qquad n \geq 0 \tag{7.7}$$

The true root is $\alpha \doteq 1.134724138$, and $x_6 \doteq \alpha$ to **nine** significant digits.

Newton's method may converge slowly at first. However, as the iterates come closer to the root, the speed of convergence increases.

Use `newton.m`

n	x_n	$f(x_n)$	$x_n - x_{n-1}$	$\alpha - x_{n-1}$
0	1.5	8.89E+1		
1	1.30049088	2.54E+1	-2.00E-1	-3.65E-1
2	1.18148042	5.38E−1	-1.19E-1	-1.66E-1
3	1.13945559	4.92E−2	-4.20E-2	-4.68E-2
4	1.13477763	5.50E−4	-4.68E-3	-4.73E-3
5	1.13472415	7.11E−8	-5.35E-5	-5.35E-5
6	1.13472414	1.55E−15	-6.91E-9	-6.91E-9
	1.134724138			

Table: Newton's Method for $x^6 - x - 1 = 0$

Compare these results with the results for the bisection method.

Example

One way to compute $\frac{a}{b}$ on early computers (that had hardware arithmetic for addition, subtraction and multiplication) was by multiplying a and $\frac{1}{b}$, with $\frac{1}{b}$ approximated by Newton's method.

$$f(x) \equiv b - \frac{1}{x} = 0$$

where we assume $b > 0$. The root is $\alpha = \frac{1}{b}$, the derivative is

$$f'(x) = \frac{1}{x^2}$$

and Newton's method is given by

$$x_{n+1} = x_n - \frac{b - \frac{1}{x_n}}{\frac{1}{x_n^2}},$$

i.e.,

$$x_{n+1} = x_n(2 - bx_n), \quad n \geq 0 \tag{7.8}$$

This involves only multiplication and subtraction.
The initial guess should be chosen $x_0 > 0$.
For the error it can be shown

$$Rel(x_{n+1}) = [Rel(x_n)]^2, \qquad n \geq 0 \tag{7.9}$$

where

$$Rel(x_n) = \frac{\alpha - x_n}{\alpha}$$

the relative error when considering x_n as an approximation to $\alpha = 1/b$. From (7.9) we must have

$$|Rel(x_0)| < 1$$

Otherwise, the error in x_n will not decrease to zero as n increases.
This contradiction means

$$-1 < \frac{\frac{1}{b} - x_0}{\frac{1}{b}} < 1$$

equivalently

$$0 < x_0 < \frac{2}{b} \tag{7.10}$$

The iteration (7.8), $x_{n+1} = x_n(2 - bx_n), n \geq 0$, converges to $\alpha = \frac{1}{b}$ if and only if the initial guess x_0 satisfies

$$0 < x_0 < \frac{2}{b}$$

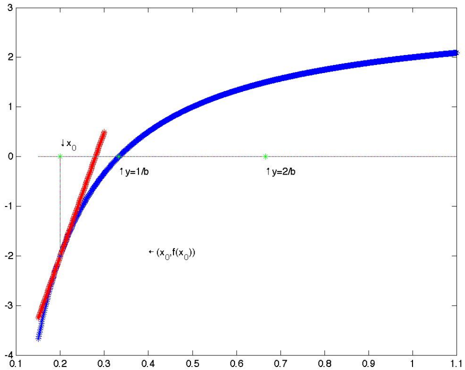


Figure: The iterative solution of $b - \frac{1}{x} = 0$

If the condition on the initial guess is violated, the calculated value of x_1 and all further iterates would be negative.

The result (7.9) shows that **the convergence is very rapid**, once we have a somewhat accurate initial guess.

For example, suppose $|Rel(x_0)| = 0.1$, which corresponds to a 10% error in x_0. Then from (7.9)

$$\begin{aligned} Rel(x_1) &= 10^{-2}, \quad Rel(x_2) = 10^{-4} \\ Rel(x_3) &= 10^{-8}, \quad Rel(x_4) = 10^{-16} \end{aligned} \tag{7.11}$$

Thus, x_3 or x_4 should be sufficiently accurate for most purposes.

Error analysis

Assume that $f \in C^2$ in some interval about the root α, and

$$f'(\alpha) \neq 0, \tag{7.12}$$

i.e., the graph $y = f(x)$ is not tangent to the x-axis when the graph intersects it at $x = \alpha$. The case in which $f'(\alpha) = 0$ is treated in Section 3.5. Note that combining (7.12) with the continuity of $f'(x)$ implies that $f'(x) \neq 0$ for all x near α.
By Taylor's theorem

$$f(\alpha) = f(x_n) + (\alpha - x_n)f'(x_n) + \frac{1}{2}(\alpha - x_n)^2 f''(c_n)$$

with c_n an unknown point between α and x_n.
Note that $f(\alpha) = 0$ by assumption, and then divide by $f'(x_n)$ to obtain

$$0 = \frac{f(x_n)}{f'(x_n)} + \alpha - x_n + (\alpha - x_n)^2 \frac{f''(c_n)}{2f'(x_n)}.$$

Quadratic convergence of Newton's method

Solving for $\alpha - x_{n+1}$, we have

$$\alpha - x_{n+1} = (\alpha - x_n)^2 \left[\frac{-f''(c_n)}{2f'(x_n)}\right] \tag{7.13}$$

This formula says that the error in x_{n+1} is nearly proportional to the square of the error in x_n.
When the initial error is sufficiently small, this shows that the error in the succeeding iterates will decrease very rapidly, just as in (7.11).

Formula (7.13) can also be used to give a formal mathematical proof of the convergence of Newton's method.

Example

For the earlier iteration (7.7), i.e., $x_{n+1} = x_n - \frac{x_n^6 - x_n - 1}{6x_n^5 - 1}, n \geq 0$, we have $f''(x) = 30x^4$. If we are near the root α, then

$$\frac{-f''(c_n)}{2f'(c_n)} \approx \frac{-f''(\alpha)}{2f'(\alpha)} = \frac{-30\alpha^4}{2(6\alpha^5 - 1)} \doteq -2.42$$

Thus for the error in (7.7),

$$\alpha - x_{n+1} \approx -2.42(\alpha - x_n)^2 \tag{7.14}$$

This explains the rapid convergence of the final iterates in table.
For example, consider the case of $n = 3$, with $\alpha - x_3 \doteq -.73E - 3$. Then (7.14) predicts

$$\alpha - x_4 \doteq 2.42(4.73E - 3)^3 \doteq -5.42E - 5$$

which compares well to the actual error of $\alpha - x_4 \doteq 5.35E - 5$.

If we assume that the iterate x_n is near the root α, the multiplier on the RHS of (7.13), i.e., $\alpha - x_{n+1} = (\alpha - x_n)^2 \left[\frac{-f''(c_n)}{2f'(x_n)}\right]$ can be written as

$$\frac{-f''(c_n)}{2f'(x_n)} \approx \frac{-2f''(\alpha)}{2f'(\alpha)} \equiv M. \tag{7.15}$$

Thus,

$$\alpha - x_{n+1} \approx M(\alpha - x_n)^2, \qquad n \geq 0$$

Multiply both sides by M to get

$$M(\alpha - x_{n+1}) \approx [M(\alpha - x_n)]^2$$

Assuming that all of the iterates are near α, then inductively we can show that

$$M(\alpha - x_n) \approx [M(\alpha - x_0)]^{2^n}, \qquad n \geq 0$$

Since we want $\alpha - x_n$ to converge to zero, this says that we must have

$$|M(\alpha - x_0)| < 1$$

$$|\alpha - x_0| < \frac{1}{|M|} = \left|\frac{2f'(\alpha)}{f''(\alpha)}\right| \tag{7.16}$$

If the quantity $|M|$ is very large, then x_0 will have to be chosen very close to α to obtain convergence. In such situation, the bisection method is probably an easier method to use.
The choice of x_0 can be very important in determining whether Newton's method will converge.

Unfortunately, there is *no single strategy that is always effective* in choosing x_0.

- In most instances, a choice of x_0 arises from physical situation that led to the rootfinding problem.
- In other instances, graphing $y = f(x)$ will probably be needed, possibly combined with the bisection method for a few iterates.

We are computing sequence of iterates x_n, and we would like to estimate their accuracy to **know when to stop the iteration**.
To estimate $\alpha - x_n$, note that, since $f(\alpha) = 0$, we have

$$f(x_n) = f(x_n) - f(\alpha) = f'(\xi_n)(x_n - \alpha)$$

for some ξ_n between x_n and α, by the mean-value theorem.
Solving for the error, we obtain

$$\alpha - x_n = \frac{-f(x_n)}{f'(\xi_n)} \approx \frac{-f(x_n)}{f'(x_n)}$$

provided that x_n is so close to α that $f'(x_n) \doteq f'(\xi_n)$. From the Newton-Raphson method (7.6), i.e., $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$, this becomes

$$\alpha - x_n \approx x_{n+1} - x_n \tag{7.17}$$

This is the standard error estimation formula for Newton's method, and it is usually fairly accurate.

However, this formula is not valid if $f'(\alpha) = 0$, a case that is discussed in Section 3.5.

Example

Consider the error in the entry x_3 of the previous table.

$$\begin{aligned}\alpha - x_3 &\doteq -4.73E-3\\ x_4 - x_3 &\doteq -4.68E-3\end{aligned}$$

This illustrates the accuracy of (7.17) for that case.

Linear convergence of Newton's method

Example

Use Newton's Method to find a root of $f(x) = x^2$.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2}{2x_n} = \frac{x_n}{2}.$$

So the method converges to the root $\alpha = 0$, but the convergence is only linear

$$e_{n+1} = \frac{e_n}{2}.$$

Example

Use Newton's Method to find a root of $f(x) = x^m$.

$$x_{n+1} = x_n - \frac{x_n^m}{m x_{m-1}} = \frac{m-1}{m} x_n.$$

The method converges to the root $\alpha = 0$, again with linear convergence

$$e_{n+1} = \frac{m-1}{m} e_n.$$

Linear convergence of Newton's method

Theorem

Assume $f \in C^{m+1}[a,b]$ and has a multiplicity m root α. Then Newton's Method is locally convergent to α, and the absolute error e_n satisfies

$$\lim_{n\to\infty} \frac{e_{n+1}}{e_n} = \frac{m-1}{m}. \tag{7.18}$$

Linear convergence of Newton's method

Example

Find the multiplicity of the root $\alpha = 0$ of $f(x) = \sin x + x^2\cos x - x^2 - x$, and estimate the number of steps in NM for convergence to 6 correct decimal places (use $x_0 = 1$).

$$\begin{aligned}
f(x) &= \sin x + x^2\cos x - x^2 - x & \Rightarrow f(0) = 0\\
f'(x) &= \cos x + 2x\cos x - x^2\sin x - 2x - 1 & \Rightarrow f'(0) = 0\\
f''(x) &= -\sin x + 2\cos x - 4x\sin x - x^2\cos x - 2 & \Rightarrow f''(0) = 0\\
f'''(x) &= -\cos x - 6\sin x - 6x\cos x + x^2\sin x & \Rightarrow f'''(0) = -1
\end{aligned}$$

Hence $\alpha = 0$ is a triple root, $m = 3$; so $e_{n+1} \approx \frac{2}{3}e_n$.
Since $e_0 = 1$, we need to solve

$$\left(\frac{2}{3}\right)^n < 0.5 \times 10^{-6}, \quad n > \frac{\log_{10} .5 - 6}{\log_{10} 2/3} \approx 35.78.$$

Modified Newton's Method

If the multiplicity of a root is known in advance, convergence of Newton's Method can be improved.

Theorem

Assume $f \in C^{m+1}[a,b]$ which contains a root α of multiplicity $m > 1$. Then Modified Newton's Method

$$x_{n+1} = x_n - m\frac{f(x_n)}{f'(x_n)} \tag{7.19}$$

converges locally and quadratically to α.

Proof. MNM: $mf(x_n) = (x_n - x_{n+1})f'(x_n)$.
Taylor's formula:

$$\begin{aligned} 0 &= \frac{x_n - x_{n+1}}{m} f'(x_n) + (\alpha - x_n) f'(x_n) + f''(c)\frac{(\alpha - x_n)^2}{2!} \\ &= \frac{\alpha - x_{n+1}}{m} f'(x_n) + (\alpha - x_n) f'(x_n)\left(1 - \frac{1}{m}\right) + f''(c)\frac{(\alpha - x_n)^2}{2!} \\ &= \frac{\alpha - x_{n+1}}{m} f'(x_n) + (\alpha - x_n)^2\left(1 - \frac{1}{m}\right) f''(\xi) + (\alpha - x_n)^2 \frac{f''(c)}{2!} \end{aligned}$$

Failure of Newton's Method

Example

Apply Newton's Method to $f(x) = -x^4 + 3x^2 + 2$ with starting guess $x_0 = 1$.

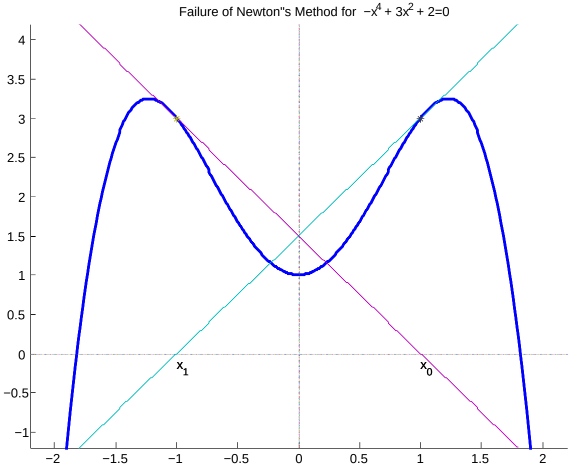
The Newton formula is

$$x_{n+1} = x_n - \frac{-x^4+3x_n^2+2}{-4x_n^3+6x_n},$$

which gives

$$x_1 = -1, \quad x_2 = 1, \ldots$$

Failure of Newton's Method



The Newton method is based on approximating the graph of $y = f(x)$ with a tangent line and on then using a root of this straight line as an approximation to the root α of $f(x)$.
From this perspective,
other straight-line approximation to $y = f(x)$ would also lead to methods of approximating a root of $f(x)$. One such straight-line approximation leads to the **secant method**.

Assume that two initial guesses to α are known and denote them by

$$\boldsymbol{x_0} \text{ and } \boldsymbol{x_1}.$$

They may occur

- on opposite sides of α, or
- on the same side of α.

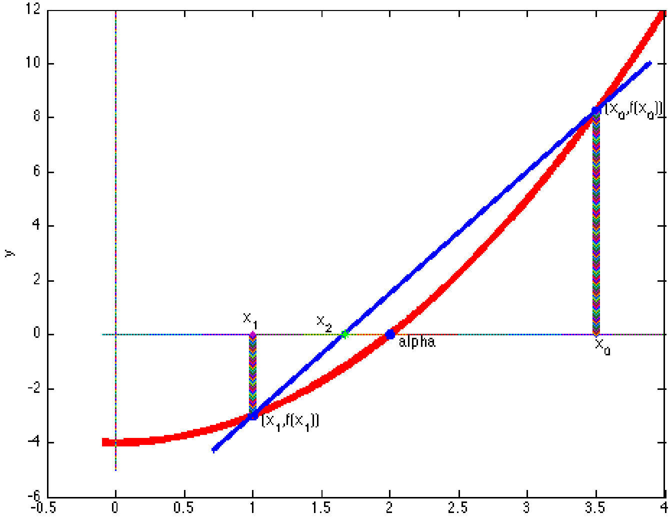


Figure: A schematic of the secant method: $x_1 < \alpha < x_0$

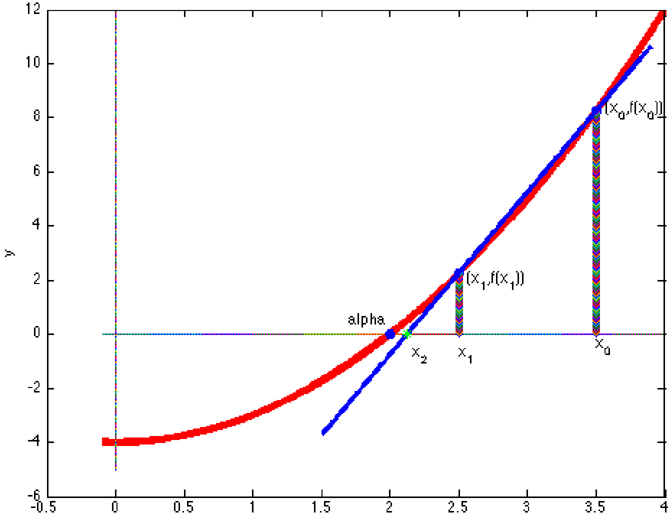


Figure: A schematic of the secant method: $\alpha < x_1 < x_0$

To derive a formula for x_2, we proceed in a manner similar to that used to derive Newton's method:

Find the equation of the line and then find its root x_2.

The equation of the line is given by

$$y = p(x) \equiv f(x_1) + (x - x_1) \cdot \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

Solving $p(x_2) = 0$, we obtain

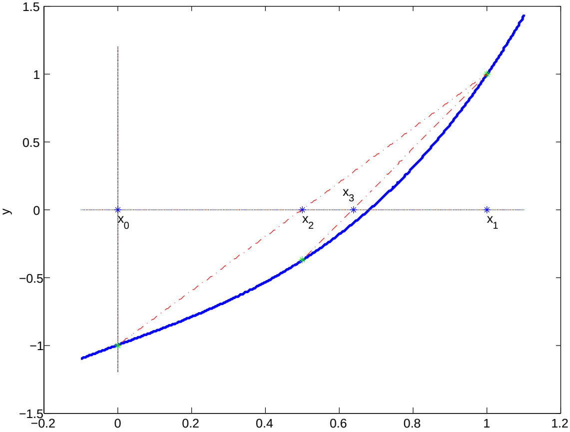
$$x_2 = x_1 - f(x_1) \cdot \frac{x_1 - x_0}{f(x_1) - f(x_0)}.$$

Having found x_2, we can drop x_0 and use x_1, x_2 as a new set of approximate values for α. This leads to an improved values x_3; and this can be continued indefinitely. Doing so, we obtain the general formula for the **secant method**

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}, \qquad n \geq 1. \tag{7.20}$$

It is called a *two-point method*, since two approximate values are needed to obtain an improved value. The bisection method is also a two-point method, but the secant method will almost always converge faster than bisection.

Two steps of the secant method for $f(x) = x^3 + x - 1$, $x_0 = 0, x_1 = 1$



Use `secant.m`

Example

We solve the equation $f(x) \equiv x^6 - x - 1 = 0$.

n	x_n	$f(x_n)$	$x_n - x_{n-1}$	$\alpha - x_{n-1}$
0	2.0	61.0		
1	1.0	-1.0	-1.0	
2	1.01612903	-9.15E-1	1.61E-2	1.35E-1
3	1.19057777	6.57E-1	1.74E-1	1.19E-1
4	1.11765583	-1.68E-1	-7.29E-2	-5.59E-2
5	.113253155	-2.24E-2	-2.24E-2	1.71E-2
6	1.13481681	9.54E-4	2.29E-3	2.19E-3
7	1.13472365	-5.07E-6	-9.32E-5	-9.27E-5
8	1.13472414	-1.13E-9	4.92E-7	4.92E-7

The iterate x_8 equals α rounded to nine significant digits.

As with the Newton method (7.7) for this equation, the initial iterates do not converge rapidly. But as the iterates become closer to α, the speed of convergence increases.

By using techniques from calculus and some algebraic manipulation, it is possible to show that the iterates x_n of (7.20) satisfy

$$\boldsymbol{\alpha - x_{n+1} = (\alpha - x_n)(\alpha - x_{n-1})\frac{-f''(\xi_n)}{2f'(\zeta_n)}}. \tag{7.21}$$

The unknown number ζ_n is between x_n and x_{n-1}, and the unknown number ξ_n is between the largest and the smallest of the numbers α, x_n and x_{n-1}. The error formula closely resembles the Newton error formula (7.13). This should be expected, since the secant method can be considered as an approximation of Newton's method, based on using

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Check that the use of this in the Newton formula (7.6) will yield (7.20).

The formula (7.21) can be used to obtain the further error result that if x_0 and x_1 are chosen sufficiently close to α, then we have convergence and

$$\lim_{n\to\infty} \frac{|\alpha - x_{n+1}|}{|\alpha - x_n|^r} = \left|\frac{f''(\alpha)}{2f'(\alpha)}\right|^{r-1} \equiv c$$

where $r = \frac{\sqrt{5}+1}{2} \doteq 1.62$. Thus,

$$|\boldsymbol{\alpha} - \boldsymbol{x_{n+1}}| \approx \boldsymbol{c}|\boldsymbol{\alpha} - \boldsymbol{x_n}|^{\mathbf{1.62}} \tag{7.22}$$

as x_n approaches α. Compare this with the Newton estimate (7.15), in which the exponent is 2 rather then 1.62. Thus, Newton's method converges more rapidly than the secant method. Also, the constant c in (7.22) plays the same role as M in (7.15), and they are related by

$$c = |M|^{r-1}.$$

The restriction (7.16) on the initial guess for Newton's method can be replaced by a similar one for the secant iterates, but we omit it.

Finally, the result (7.22) can be used to justify the error estimate

$$\alpha - x_{n-1} \approx x_n - x_{n-1}$$

for iterates x_n that are sufficiently close to the root.

Example

For the iterate x_5 in the previous Table

$$\begin{aligned} \alpha - x_5 &\doteq 2.19E-3 \\ x_6 - x_5 &\doteq 2.29E-3 \end{aligned} \tag{7.23}$$

- From the foregoing discussion, Newton's method converges more rapidly than the secant method. Thus,
  Newton's method should require fewer iterations to attain a given error tolerance.
- However, Newton's method requires two function evaluations per iteration, that of $f(x_n)$ and $f'(x_n)$. And the secant method requires only one evaluation, $f(x_n)$, if it is programed carefully to retain the value of $f(x_{n-1})$ from the preceding iteration. Thus,
  the **secant method will require less time per iteration** than the Newton method.

The decision as to which method should be used will depend on the factors just discussed, including the difficulty or expense of evaluating $f'(x_n)$; and it will depend on intangible human factors, such as convenience of use. Newton's method is very simple to program and to understand; but for many problems with a complicated $f'(x)$, the secant method will probably be faster in actual running time on a computer.

General remarks

The derivation of both the Newton and secant methods illustrate a general principle of numerical analysis.

When trying to solve a problem for which there is no direct or simple method of solution, approximate it by another problem that you can solve more easily.

In both cases, we have replaced the solution of

$$f(x) = 0$$

with the solution of a much simpler rootfinding problem for a linear equation.

GENERAL OBSERVATION

When dealing with problems involving differentiable functions $f(x)$, move to a nearby problem by approximating each such $f(x)$ with a linear problem.

The linearization of mathematical problems is common throughout applied mathematics and numerical analysis.

MATLAB contains the rootfinding routine $fzero$ that uses ideas involved in the bisection method and the secant method. As with many MATLAB programs, there are several possible calling sequences.

- The command
  $$\text{root} = \text{fzero}(\text{f\_name}, [a, b])$$
  produces a root within $[a, b]$, where it is assumed that $f(a)f(b) \leq 0$.
- The command
  $$\text{root} = \text{fzero}(\text{f\_name}, x0)$$
  tries to find a root of the function near $x0$.

The default error tolerance is the maximum precision of the machine, although this can be changed by the user.
This is an excellent rootfinding routine, combining guaranteed convergence with high efficiency.

There are three generalization of the Secant method that are also important. The **Method of False Position**, or **Regula Falsi**, is similar to the Bisection Method, but where the midpoint is replaced by a Secant Method-like approximation. Given an interval $[a, b]$ that brackets a root (assume that $f(a)f(b) < 0$), define the next point

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

as in the Secant Method, but unlike the Secant Method, the new point is guaranteed to lie in $[a, b]$, since the points $(a, f(a))$ and $(b, f(b))$ lie on separate sides of the x-axis. The new interval, either $[a, c]$ or $[c, b]$, is chosen according to whether $f(a)f(c) < 0$ or $f(c)f(b) < 0$, respectively, and still brackets a root.

Given interval $[a, b]$ such that $f(a)f(b) < 0$

for $i = 1, 2, 3, \ldots$

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

if $f(c) = 0,$ **stop, end**

if $f(a)f(c) < 0$

$b = c$

else

$a = c$

end

end

The Method of False Position at first appears to be an improvement on both the Bisection Method and the Secant Method, taking the best properties of each. However, while the Bisection method guarantees cutting the uncertainty by 1/2 on each step, False Position makes no such promise, and for some examples can converge very slowly.

Example

Apply the Method of False Position on initial interval [-1,1] to find the root $r = 1$ of $f(x) = x^3 - 2x^2 + \frac{3}{2}x$.

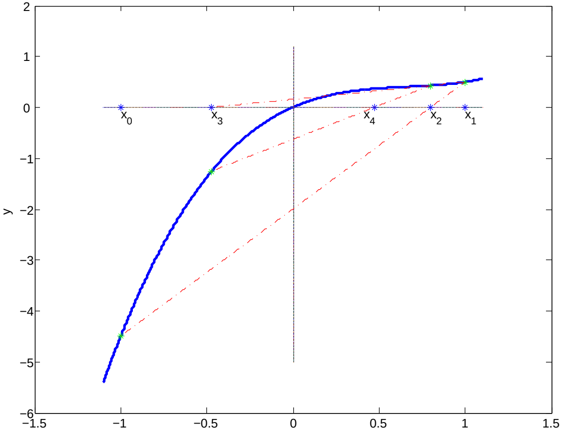
Given $x_0 = -1, x_1 = 1$ as the initial bracketing interval, we compute the new point

$$x_2 = \frac{x_1 f(x_0) - x_0 f(x_1)}{f(x_0) - f(x_1)} = \frac{1(-9/2) - (-1)1/2}{-9/2 - 1/2} = \frac{4}{5}.$$

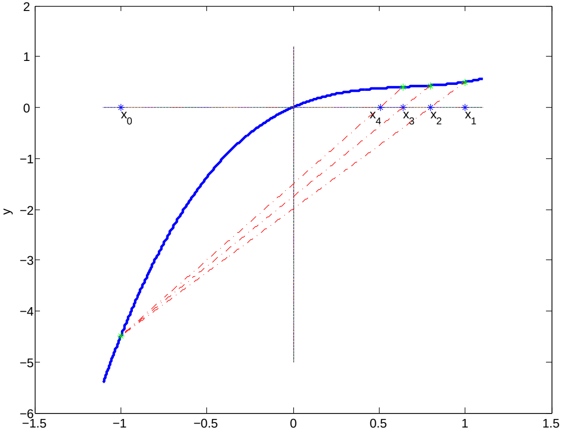
Since $f(-1)f(4/5) < 0$, the new bracketing interval is $[x_0, x_2] = [-1, 0.8]$. This completes the first step. Note that the uncertainty in the solution has decreased by far less than a factor of $1/2$. As seen in the Figure, further steps continue to make slow progress toward the root at $x = 0$.

Both the Secant Method and Method of False Position converge slowly to the root $r = 0$.

(a) The Secant Method converges slowly to the root $r = 0$.



(b) The Method of False Position converges slowly to the root $r = 0$.



- The *Newton method* (7.6)

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n = 0, 1, 2, \ldots$$

- and the *secant method* (7.20)

$$x_{n+1} = x_n - f(x_n) \cdot \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \qquad n \geq 1$$

are examples of one-point and two-point iteration methods, respectively.

In this section we give a more general introduction to iteration methods, presenting a general theory for one-point iteration formulae.

Solve the equation

$$x = g(x)$$

for a root $\alpha = g(\alpha)$ by the **iteration**

$$\begin{cases} \boldsymbol{x_0}, & \\ \boldsymbol{x_{n+1} = g(x_n)}, & n = 0, 1, 2, \ldots \end{cases}$$

Example: Newton's Method

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} := g(x_n)$$

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

Definition

The solution $\boldsymbol{\alpha}$ is called a **fixed point of $\boldsymbol{g}$**.

The solution of $f(x) = 0$ can always be rewritten as a fixed point of g, e.g.,

$$x + f(x) = x \quad \Longrightarrow \quad g(x) = x + f(x).$$

Example

As motivational example, consider solving the equation

$$x^2 - 5 = 0 \tag{7.24}$$

for the root $\alpha = \sqrt{5} \doteq 2.2361$.

We give four methods to solve this equation

I1. $x_{n+1} = 5 + x_n - x_n^2$ $\qquad x = x + c(x^2 - a), c \neq 0$

I2. $x_{n+1} = \frac{5}{x_n}$ $\qquad x = \frac{a}{x}$

I3. $x_{n+1} = 1 + x_n - \frac{1}{5}x_n^2$ $\qquad x = x + c(x^2 - a), c \neq 0$

I4. $x_{n+1} = \frac{1}{2}\left(x_n + \frac{5}{x_n}\right)$ $\qquad x = \frac{1}{2}(x + \frac{a}{x})$

All four iterations have the property that if the sequence $\{x_n : n \geq 0\}$ has a limit α, then α is a root of (7.24). For each equation, check this as follows: Replace x_n and x_{n+1} by α, and then show that this implies $\alpha = \pm\sqrt{5}$.

n	x_n:I1	x_n:I2	x_n :I3	x_n :I4
0	2.5	2.5	2.5	2.5
1	1.25	2.0	2.25	2.25
2	4.6875	2.5	2.2375	2.2361
3	-12.2852	2.0	2.2362	2.2361

Table: The iterations I1 to I4

To explain these numerical results, we present a general theory for one-point iteration formulae.

The iterations I1 to I4 all have the form

$$x_{n+1} = g(x_n)$$

for appropriate continuous functions $g(x)$. For example, with I1, $g(x) = 5 + x - x^2$. If the iterates x_n converge to a point α, then

$$\begin{aligned} \lim_{n\to\infty} x_{n+1} &= \lim_{n\to\infty} g(x_n) \\ \alpha &= g(\alpha) \end{aligned}$$

Thus α is a solution of the equation $x = g(x)$, and α is called a fixed point of the function g.

Existence of a fixed point

In this section, a general theory is given to explain when the iteration $x_{n+1} = f(x_n)$ will converge to a fixed point of g.
We begin with a lemma on existence of solutions of $x = g(x)$.

Lemma

Let $g \in C[a,b]$. Assume that $g([a,b]) \subset [a,b]$, i.e.,

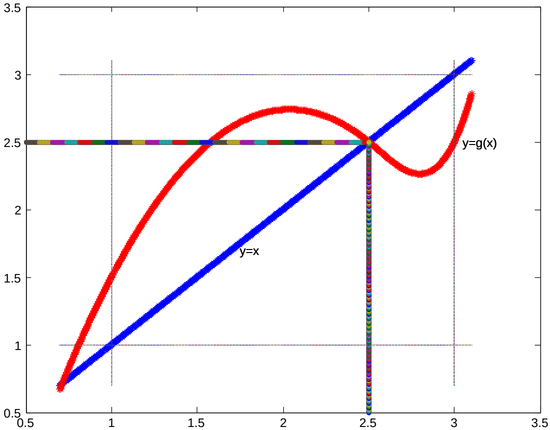
$$\forall x \in [a,b], \qquad g(x) \in [a,b].$$

Then $x = g(x)$ has at least one solution α in the interval $[a,b]$.

Proof. Define the function $f(x) = x - g(x)$. It is continuous for $a \leq x \leq b$. Moreover,

$$f(a) = a - g(a) \leq 0$$
$$f(b) = b - g(b) \geq 0$$

Intermediate value theorem $\Rightarrow \exists x \in [a,b]$ such that $f(x) = 0$, i.e. $x = g(x)$.
□



The solutions α are the x-coordinates of the intersection points of the graphs of $y = x$ and $y = g(x)$.

Lipschitz continuous

Definition

Given $g : [a, b] \to \mathbb{R}$, is called Lipschitz continuous with constant $\lambda > 0$ (denoted $g \in Lip_\lambda$[a,b]) if $\exists \lambda > 0$ such that

$$|g(x) - g(y)| \le \lambda |x - y| \qquad \forall x, y \in [a, b].$$

Definition

$g : [a, b] \to \mathbb{R}$ is called contraction map if $g \in Lip_\lambda[a, b]$ with $\lambda < 1$.

Existence and uniqueness of a fixed point

Lemma

Let $g \in Lip_\lambda[a,b]$ with $\lambda < 1$ and $g([a,b]) \subset [a,b]$.
Then $x = g(x)$ has exactly one solution α. Moreover, for $x_{n+1} = g(x_n)$, $x_n \to \alpha$ for any $x_0 \in [a,b]$ and

$$|\alpha - x_n| \leq \frac{\lambda^n}{1-\lambda}|x_1 - x_0|.$$

Proof.

Existence: follows from previous Lemma.

Uniqueness: assume $\exists \alpha, \beta$ solutions: $\alpha = g(\alpha)$, $\beta = g(\beta)$.

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| \leq \lambda|\alpha - \beta|$$
$$\underbrace{(1-\lambda)}_{>0}|\alpha - \beta| \leq 0.$$

Convergence of the iterates

If $x_n \in [a,b]$ then $g(x_n) = x_{n+1} \in [a,b]) \Rightarrow \{x_n\}_{n\geq 0} \subset [a,b]$.

Linear convergence with rate λ:

$$|\alpha - x_n| = |g(\alpha) - g(x_n)| \leq \lambda|\alpha - x_{n-1}| \leq \ldots \leq \lambda^n|\alpha - x_0| \quad (7.25)$$

$$\begin{aligned} |x_0 - \alpha| &= |x_0 - x_1 + x_1 - \alpha| \leq |x_0 - x_1| + |x_1 - \alpha| \\ &\leq |x_0 - x_1| + \lambda|x_0 - \alpha| \end{aligned}$$

$$\Longrightarrow |x_0 - \alpha| \leq \frac{x_0 - x_1}{1 - \lambda} \quad (7.26)$$

$$|x_n - \alpha| \leq \lambda^n|x_0 - \alpha| \leq \frac{\lambda^n}{1-\lambda}|x_1 - x_0|$$

□

Error estimate

$$|\alpha - x_n| \leq |\alpha - x_{n+1}| + |x_{n+1} - x_n| \leq \lambda|\alpha - x_n| + |x_{n+1} - x_n|$$

$$|\alpha - x_n| \leq \frac{1}{1-\lambda}|x_n - x_{n+1}|$$

$$|\alpha - x_{n+1}| \leq \lambda|\alpha - x_n|$$

$$\Longrightarrow |\alpha - x_{n+1}| \leq \frac{\lambda}{1-\lambda}|x_{n+1} - x_n|$$

Assume $g'(x)$ exists on $[a, b]$. By the mean value theorem:

$$g(x) - g(y) = g'(\xi)(x - y), \qquad \xi \in [a, b], \forall x, y \in [a, b]$$

Define

$$\lambda = \max_{x \in [a,b]} |g'(x)|.$$

Then $g \in Lip_\lambda[a, b]$:

$$|g(x) - g(y)| \leq |g'(\xi)||x - y| \leq \lambda|x - y|.$$

Theorem 2.6

Assume $g \in C^1[a,b], g([a,b]) \subset [a,b]$ and $\max_{x\in[a,]} |g'(x)| < 1$. Then

1. $x = g(x)$ has a unique solution α in $[a,b]$,
2. $x_n \to \alpha \qquad \forall x_0 \in [a,b]$,
3. $|\alpha - x_n| \leq \dfrac{\lambda^n}{1-\lambda}|x_1 - x_0|$,
4. $\displaystyle\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha - x_n} = g'(\alpha)$.

Proof.

$$\begin{aligned}\alpha - x_{n+1} &= g(\alpha) - g(x_n) \\ &= g'(\xi_n)(\alpha - x_n), \quad \xi_n \in [\alpha, x_n].\end{aligned}$$

□

Theorem 2.7

Assume α solves $x = g(x), g \in C^1[I_\alpha]$, for some $I_\alpha \ni \alpha$, $|g'(\alpha)| < 1$. Then Theorem 2.6 holds for x_0 close enough to α.

Proof.
Since $|g'(\alpha)| < 1$ by continuity

$$|g'(x)| < 1 \qquad \text{for } x \in I_\alpha = [\alpha - \varepsilon, \alpha + \varepsilon].$$

Take $x_0 \in I_\alpha$ close to $x_1 \in I_\alpha$

$$\begin{aligned} |x_1 - \alpha| &= |g(x_0) - g(\alpha)| = |g'(\xi)(x_0 - \alpha)| \\ &\leq |g'(\xi)||x_0 - \alpha| < |x_0 - \alpha| < \varepsilon \end{aligned}$$

$\Rightarrow x_1 \in I_\alpha$ and, by induction, $x_n \in I_\alpha$.
$\Rightarrow$ Theorem 2.6 holds with $[a, b] = I_\alpha$. □

Importance of $|g'(\alpha)| < 1$:

If $|g'(\alpha)| > 1$: and x_n is close to α then

$$\begin{aligned} |x_{n+1} - \alpha| &= |g'(\xi_n)||x_n - \alpha| \\ &\to |x_{n+1} - \alpha| > |x_n - \alpha| \\ &\implies \text{divergence} \end{aligned}$$

When $g'(\alpha) = 1$, no conclusion can be drawn; and even if convergence were to occur, the method would be far too slow for the iteration method to be practical.

Examples

Recall $\alpha = \sqrt{5}$.

1. $g(x) = 5 + x - x^2; , g'(x) = 1 - 2x, \qquad g'(\alpha) = 1 - 2\sqrt{5}$. Thus the iteration will not converge to $\sqrt{5}$.
2. $g(x) = 5/x, g'(x) = -\frac{5}{x^2}; \qquad g'(\alpha) = -\frac{5}{(\sqrt{5})^2} = -1.$

   We cannot conclude that the iteration converges or diverges. From the Table, it is clear that the iterates will not converge to α.
3. $g(x) = 1 + x - \frac{1}{5}x^2, g'(x) = 1 - \frac{2}{5}x, g'(\alpha) = 1 - \frac{2}{5}\sqrt{5} \doteq 0.106$, i.e., the iteration will converge. Also,

   $$|\alpha - x_{n+1}| \approx 0.106|\alpha - x_n|$$

   when x_n is close to α. The errors will decrease by approximately a factor of 0.1 with each iteration.
4. $g(x) = \frac{1}{2}\left(x + \frac{5}{x}\right); \qquad g'(\alpha) = 0 \qquad$ convergence

Note that this is Newton's method for computing $\sqrt{5}$.

Examples

Recall $\alpha = \sqrt{5}$.

1. $g(x) = 5 + x - x^2;, g'(x) = 1 - 2x, \qquad g'(\alpha) = 1 - 2\sqrt{5}$. Thus the iteration will not converge to $\sqrt{5}$.
2. $g(x) = 5/x, g'(x) = -\frac{5}{x^2}; \qquad g'(\alpha) = -\frac{5}{(\sqrt{5})^2} = -1.$
   We cannot conclude that the iteration converges or diverges.
   From the Table, it is clear that the iterates will not converge to α.
3. $g(x) = 1 + x - \frac{1}{5}x^2, g'(x) = 1 - \frac{2}{5}x, g'(\alpha) = 1 - \frac{2}{5}\sqrt{5} \doteq 0.106$, i.e., the iteration will converge. Also,
   $$|\alpha - x_{n+1}| \approx 0.106|\alpha - x_n|$$
   when x_n is close to α. The errors will decrease by approximately a factor of 0.1 with each iteration.
4. $g(x) = \frac{1}{2}\left(x + \frac{5}{x}\right); \qquad g'(\alpha) = 0 \qquad$ convergence

Note that this is Newton's method for computing $\sqrt{5}$.

Examples

Recall $\alpha = \sqrt{5}$.

1. $g(x) = 5 + x - x^2;, g'(x) = 1 - 2x, \qquad g'(\alpha) = 1 - 2\sqrt{5}$. Thus the iteration will not converge to $\sqrt{5}$.
2. $g(x) = 5/x, g'(x) = -\frac{5}{x^2}; \qquad g'(\alpha) = -\frac{5}{(\sqrt{5})^2} = -1.$
   We cannot conclude that the iteration converges or diverges.
   From the Table, it is clear that the iterates will not converge to α.
3. $g(x) = 1 + x - \frac{1}{5}x^2, g'(x) = 1 - \frac{2}{5}x, g'(\alpha) = 1 - \frac{2}{5}\sqrt{5} \doteq 0.106$, i.e., the iteration will converge. Also,
   $$|\alpha - x_{n+1}| \approx 0.106|\alpha - x_n|$$
   when x_n is close to α. The errors will decrease by approximately a factor of 0.1 with each iteration.
4. $g(x) = \frac{1}{2}\left(x + \frac{5}{x}\right); \qquad g'(\alpha) = 0 \qquad$ convergence

Note that this is Newton's method for computing $\sqrt{5}$.

Examples

Recall $\alpha = \sqrt{5}$.

1. $g(x) = 5 + x - x^2;, g'(x) = 1 - 2x, \qquad g'(\alpha) = 1 - 2\sqrt{5}$. Thus the iteration will not converge to $\sqrt{5}$.
2. $g(x) = 5/x, g'(x) = -\dfrac{5}{x^2}; \qquad g'(\alpha) = -\dfrac{5}{(\sqrt{5})^2} = -1.$
   We cannot conclude that the iteration converges or diverges. From the Table, it is clear that the iterates will not converge to α.
3. $g(x) = 1 + x - \frac{1}{5}x^2, g'(x) = 1 - \frac{2}{5}x, g'(\alpha) = 1 - \frac{2}{5}\sqrt{5} \doteq 0.106,$ i.e., the iteration will converge. Also,
   $$|\alpha - x_{n+1}| \approx 0.106|\alpha - x_n|$$
   when x_n is close to α. The errors will decrease by approximately a factor of 0.1 with each iteration.
4. $g(x) = \frac{1}{2}\left(x + \frac{5}{x}\right); \qquad g'(\alpha) = 0 \qquad$ convergence

Note that this is Newton's method for computing $\sqrt{5}$.

$x = g(x), \qquad g(x) = x + c(x^2 - 3)$

What value of c will give convergent iteration?

$$g'(x) = 1 + 2cx$$
$$\alpha = \sqrt{3}$$
$$\text{Need } |g'(\alpha)| < 1$$
$$-1 < 1 + 2c\sqrt{3} < 1$$

Optimal choice: $\quad 1 + 2c\sqrt{3} = 0 \quad \Longrightarrow c = -\frac{1}{2\sqrt{3}}$.

The possible behaviour of the fixed point iterates x_n for various sizes of $g'(\alpha)$.
To see the convergence, consider the case of $x_1 = g(x_0)$, the height of the graph of $y = g(x)$ at x_0.
We bring the number x_1 back to the x-axis by using the line $y = x$ and the height $y = x_1$.
We continue this with each iterate, obtaining a stairstep behaviour when $g'(\alpha) > 0$.

When $g'(\alpha) < 0$, the iterates oscillate around the fixed point α, as can be seen.

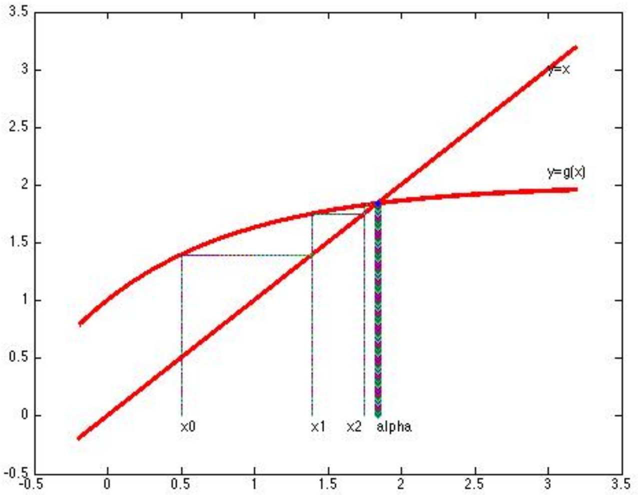


Figure: $0 < g'(\alpha) < 1$

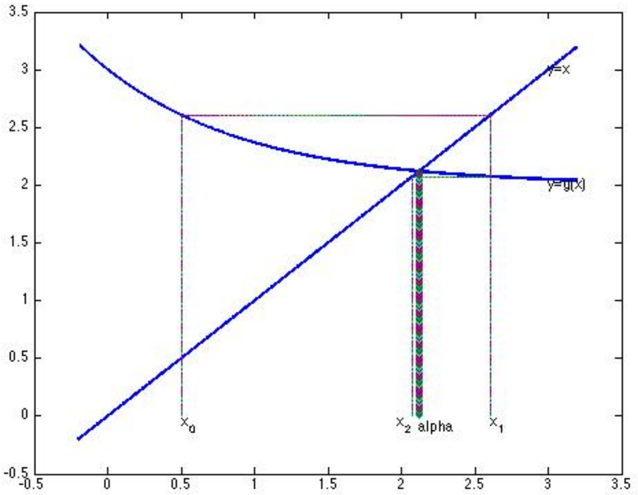


Figure: $-1 < g'(\alpha) < 0$

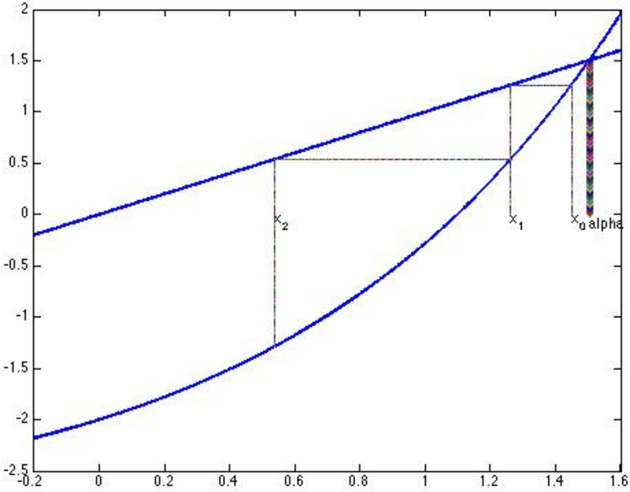


Figure: $1 < g'(\alpha)$

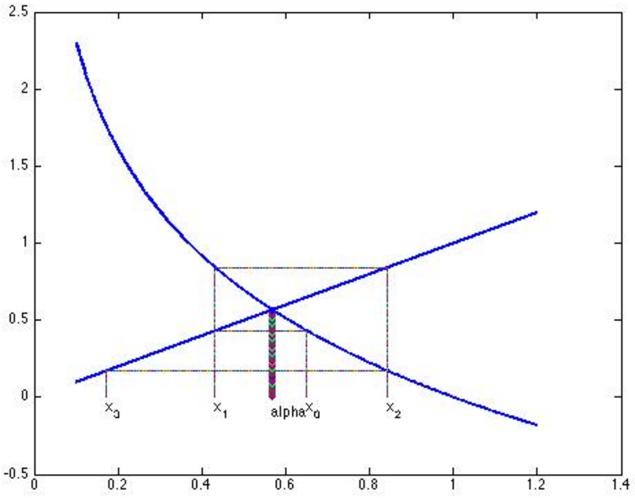


Figure: $g'(\alpha) < -1$

The results from the iteration for

$$g(x) = 1 + x - \frac{1}{5}x^2, \quad g'(\alpha) \doteq 0.106.$$

along with the ratios

$$r_n = \frac{\alpha - x_n}{\alpha - x_{n-1}}. \tag{7.27}$$

Empirically, the values of r_n converge to $g'(\alpha) \doteq 0.105573$, which agrees with

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{\alpha} - x_n = g'(\alpha).$$

n	x_n	$\alpha - x_n$	r_n
0	2.5	-2.64E-1	
1	2.25	-1.39E-2	0.0528
2	2.2375	-1.43E-3	0.1028
3	2.23621875	-1.51E-4	0.1053
4	2.23608389	-1.59E-5	0.1055
5	2.23606966	-1.68E-6	0.1056
6	2.23606815	-1.77E-7	0.1056
7	2.23606800	-1.87E-8	0.1056

Table: The iteration $x_{n+1} = 1 + x_n - \frac{1}{5}x_n^2$

We need a more precise way to deal with the concept of the speed of convergence of an iteration method.

Definition

We say that a sequence $\{x_n : n \geq 0\}$ converges to α with an **order of convergence** $p \geq 1$ if

$$|\alpha - x_{n+1}| \leq c|\alpha - x_n|^p, \qquad n \geq 0$$

for some constant $c \geq 0$.

The cases $p = 1, p = 2, p = 3$ are referred to as linear convergence, quadratic convergence and cubic convergence, respectively.

- Newton's method usually converges quadratically; and
- the secant method has a order of convergence $p = \frac{1+\sqrt{5}}{2}$.
- For linear convergence we make the additional requirement that $c < 1$; as otherwise, the error $\alpha - x_n$ need not converge to zero.

- If $g'(\alpha) < 1$, then formula

$$|\alpha - x_{n+1}| \leq |g'(\xi_n)||\alpha - x_n|$$

shows that the iterates x_n are linearly convergent.

- If in addition $g'(\alpha) \neq 0$, then formula

$$|\alpha - x_{n+1}| \approx |g'(\alpha)||\alpha - x_n|$$

proves that the convergence is exactly linear, with no higher order of convergence being possible. In this case, we call the value of $g'(\alpha)$ the linear rate of convergence.

High order one-point methods

Theorem 2.8

Assume $g \in C^p(I_\alpha)$ for some I_α containing α, and

$$g'(\alpha) = g''(\alpha) = \ldots = g^{(p-1)}(\alpha) = 0, \quad p \geq 2.$$

Then for x_0 close enough to α, $x_n \to \alpha$ and

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^p} = (-1)^{p-1} \frac{g^{(p)}(\alpha)}{p!}$$

i.e. convergence is of order p.

Proof:

$$\begin{aligned} x_{n+1} &= g(x_n) \\ &= \underbrace{g(\alpha)}_{=\alpha} + (x_n - \alpha)\underbrace{g'(\alpha)}_{=0} + \ldots + \frac{(x_n-\alpha)^{p-1}}{(p-1)!}\underbrace{g^{(p-1)}(\alpha)}_{=0} \\ &\quad + \frac{(x_n-\alpha)^p}{p!} g^{(p)}(\xi_n) \\ \alpha - x_{n+1} &= -\frac{(x_n-\alpha)^p}{p!} g^{(p)}(\xi_n). \\ \frac{\alpha - x_{n+1}}{(\alpha - x_n)^p} &= (-1)^{p-1}\frac{g^{(p)}(\xi_n)}{p!} \longrightarrow (-1)^{p-1}\frac{g^{(p)}(\alpha)}{p!} \qquad \blacksquare \end{aligned}$$

Example: Newton's method

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n > 0$$

$$= g(x_n), \qquad g(x_n) = x - \frac{f(x)}{f'(x)};$$

$$g'(x) = \frac{ff''}{(f')^2} \qquad\qquad g'(\alpha) = 0$$

$$g''(x) = \frac{f'f'' + ff'''}{(f')^2} - 2\frac{ff''}{(f')^3}, \qquad\qquad g''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$$

Theorem 2.8 with $p = 2$:

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^2} = -\frac{g''(\alpha)}{2} = -\frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)}.$$

Parallel Chords Method (two step fixed point method)

$$x_{n+1} = x_n - \frac{f(x_n)}{a}$$

Ex.: $a = f'(x_0)$.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)} = g(x_n).$$

Need $|g'(\alpha)| < 1$ for convergence: $\left|1 - \frac{f'(\alpha)}{a}\right|$

Linear convergence wit rate $1 - \frac{f'(\alpha)}{a}$. (Thm 2.6.)

If $a = f'(x_0)$ and x_0 is close enough to α, then $\left|1 - \frac{f'(\alpha)}{a}\right|$.

Aitken extrapolation for linearly convergent sequences

Recall

Theorem 2.6

$$x_{n+1} = g(x_n)$$
$$x_n \to \alpha$$
$$\frac{\alpha - x_{n+1}}{\alpha - x_n} \longrightarrow g'(\alpha)$$

Assuming linear convergence: $g'(\alpha) \neq 0$.

Derive an estimate for the error and use it to accelerate convergence.

$$\alpha - x_n = (\alpha - x_{n-1}) + (x_{n-1} - x_n) \tag{7.28}$$

$$\begin{aligned} \alpha - x_n &= g(\alpha) - g(x_{n-1}) \\ &= g'(\xi_{n-1})(\alpha - x_{n-1}) \end{aligned}$$

$$\alpha - x_{n-1} = \frac{1}{g'(\xi_{n-1})}(\alpha - x_n) \tag{7.29}$$

From (7.28)-(7.29)

$$\alpha - x_n = \frac{1}{g'(\xi_{n-1})}(\alpha - x_n) + (x_{n-1} - x_n)$$

$$\alpha - x_n = \frac{g'(\xi_{n-1})}{1 - g'(\xi_{n-1})}(x_{n-1} - x_n)$$

$$\alpha - x_n = \frac{g'(\xi_{n-1})}{1 - g'(\xi_{n-1})}(x_{n-1} - x_n)$$

$$\frac{g'(\xi_{n-1})}{1 - g'(\xi_{n-1})} \approx \frac{g'(\alpha)}{1 - g'(\alpha)}$$

Need an estimate for $g'(\alpha)$.

Define

$$\lambda_n = \frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}}$$

and

$$\alpha - x_{n+1} = g(\alpha) - g(x_n) = g'(\xi_n)(\alpha - x_n), \qquad \xi_n \in \overline{\alpha, x_n}, \quad n \geq 0.$$

$$
\begin{aligned}
\lambda_n =& \frac{(\alpha - x_{n-1}) - (\alpha - x_n)}{(\alpha - x_{n-2}) - (\alpha - x_{n-1})} \\
=& \frac{(\alpha - x_{n-1}) - g'(\xi_{n-1})(\alpha - x_{n-1})}{(\alpha - x_{n-1})/g'(\xi_{n-2}) - (\alpha - x_{n-1})} \\
=& \frac{1 - g'(\xi_{n-1})}{1 - g'(\xi_{n-2})} g'(\xi_{n-2})
\end{aligned}
$$

$$
\lambda_n \to g'(\alpha) \quad \text{as } \xi_n \to \alpha : \lambda_n \approx g'(\alpha)
$$

Aitken Error Formula

$$\alpha - x_n = \frac{\lambda_n}{1-\lambda_n}(x_n - x_{n-1}) \quad (7.30)$$

From (7.30)

$$\alpha \approx x_n + \frac{\lambda_n}{1-\lambda_n}(x_n - x_{n-1}) \quad (7.31)$$

Define

Aitken Extrapolation Formula

$$\hat{x}_n = x_n + \frac{\lambda_n}{1-\lambda_n}(x_n - x_{n-1}) \quad (7.32)$$

Example

Repeat the example for $I3$.

The Table contains the differences $x_n - x_{n-1}$, the ratios λ_n, and the estimated error from $\alpha - x_n \approx \frac{\lambda_n}{1-\lambda_n}(x_n - x_{n-1})$, given in the column Estimate. Compare the column Estimate with the error column in the previous Table.

n	x_n	$x_n - x_{n-1}$	λ_n	Estimate
0	2.5			
1	2.25	-2.50E-1		
2	2.2375	-1.25E-2	0.0500	-6.58E-4
3	2.23621875	-1.28E-3	0.1025	-1.46E-4
4	2.23608389	-1.35E-4	0.1053	-1.59E-5
5	2.23606966	-1.42E-5	0.1055	-1.68E-6
6	2.23606815	-1.50E-6	0.1056	-1.77E-7
7	2.23606800	-1.59E-7	0.1056	-1.87E-8

Table: The iteration $x_{n+1} = 1 + x_n - \frac{1}{5}x_n^2$ and Aitken Error Estimation

Algorithm (Aitken)

Given g, x_0, ε, root, assume $|g'(\alpha)| < 1$ and $x_n \to \alpha$ linearly.

1. $x_1 = g(x_0), x_2 = g(x_1)$
2. $\hat{x}_2 = x_2 + \frac{\lambda_2}{1-\lambda_2}(x_2 - x_1)$ where $\lambda_2 = \frac{x_2 - x_1}{x_1 - x_0}$
3. if $|\hat{x}_2 - x_2| \leq \varepsilon$ then root $= \hat{x}_2$; exit
4. set $x_0 = \hat{x}_2$, go to (1)

General remarks

There are a number of reasons to perform theoretical error analyses of numerical method. We want to better understand the method,

- when it will perform well,
- when it will perform poorly, and perhaps,
- when it may not work at all.

With a mathematical proof, we convinced ourselves of the correctness of a numerical method under precisely stated hypotheses on the problem being solved. Finally, we often can improve on the performance of a numerical method.

The use of the theorem to obtain the Aitken extrapolation formula is an illustration of the following:

> By understanding the behaviour of the error in a numerical method, it is often possible to improve on that method and to obtain another more rapidly convergent method.

Quasi-Newton Iterates

$$f(x) = 0 \qquad \begin{cases} x_0 \\ x_{k+1} = x_k - \frac{f(x_k)}{a_k}, \qquad k = 0, 1, \ldots \end{cases}$$

1. $a_k = f'(x_k) \Rightarrow$ *Newton's Method*
2. $a_k = \dfrac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \Rightarrow$ *Secant Method*
3. $a_k = a =$constant (e.g. $a_k = f'(x_0)$) $\Rightarrow$ *Parallel Chords Method*
4. $a_k = \dfrac{f(x_k + h_k) - f(x_k)}{h_k}$, $h_k > 0 \Rightarrow$ *Finite Diff. Newton Method*
   If $|h_k| < c|f(x_k)|$, then the convergence is quadratic. Need $h_k \geq h \approx \sqrt{\delta}$

Quasi-Newton Iterates

1. $a_k = \frac{f(x_k+f(x_k))-f(x_k)}{f(x_k)} \Rightarrow$ *Steffensen Method.* This is Finite Difference Method with $h_k = f(x_k) \Rightarrow$ quadratic convergence.
2. $a_k = \frac{f(x_k)-f(x_{k'})}{x_k-x_{k'}}$ where k' is the largest index $< k$ such that $f(x_k)f(x_{k'}) < 0 \Rightarrow$ *Regula False*
   Need $x_0, x_1 : f(x_0)f(x_1) < 0$

$$x_2 = x_1 - f(x_1)\frac{x_1 - x_0}{f(x_1) - f(x_0)}$$

$$x_3 = x_2 - f(x_2)\frac{x_2 - x_0}{f(x_2) - f(x_0)}$$

The convergence formula

$$\alpha - x_{n+1} \approx g'(\alpha)(\alpha - x_n)$$

gives less information in the case $g'(\alpha) = 0$, although the convergence is clearly quite good. To improve on the results in the Theorem, consider the Taylor expansion of $g(x_n)$ about α, assuming that $g(x)$ is twice continuously differentiable:

$$g(x_n) = g(\alpha) + (x_n - \alpha)g'(\alpha) + \frac{1}{2}(x_n - \alpha)^2 g''(c_n) \tag{7.33}$$

with c_n between x_n and α. Using $x_{n+1} = g(x_n)$, $\alpha = g(\alpha)$, and $g'(\alpha) = 0$, we have

$$x_{n+1} = \alpha + \frac{1}{2}(x_n - \alpha)^2 g''(c_n)$$

$$\alpha - x_{n+1} = -\frac{1}{2}(\alpha - x_n)^2 g''(c_n) \tag{7.34}$$

$$\lim_{n\to\infty} \frac{\alpha - x_{n+1}}{(\alpha - x_n)^2} = -\frac{1}{2} g''(\alpha) \tag{7.35}$$

If $g''(\alpha) \neq 0$, then this formula shows that the iteration $x_{n+1} = g(x_n)$ is of *order* 2 or is *quadratically convergent*.

If also $g''(\alpha) = 0$, and perhaps also some high-order derivatives are zero at α, then expand the Taylor series through higher-order terms in (7.33), until the final error term contains a derivative of g that is not nonzero at α. Thi leads to methods with an order of convergence greater than 2.

As an example, consider Newton's method as a fixed-point iteration:

$$x_{n+1} = g(x_n), \quad g(x) = x - \frac{f(x)}{f'(x)}. \tag{7.36}$$

Then,

$$g'(x) = \frac{f(x)f''(x)}{[f'(x)]^2}$$

and if $f'(\alpha) \neq 0$, then

$$g'(\alpha) = 0.$$

Similarly, it can be shown that $g''(\alpha) \neq 0$ if moreover, $f''(\alpha) \neq 0$. If we use (7.35), these results shows that Newton's method is of order 2, provided that $f'(\alpha) \neq 0$ and $f''(\alpha) \neq 0$.

We will examine two classes of problems for which the methods of Sections 3.1 to 3.4 do not perform well. Often there is little that a numerical analyst can do to improve these problems, but one should be aware of their existence and of the reason for their ill-behaviour.

We begin with functions that have a *multiple root*. The root α of $f(x)$ is said to be of multiplicity m if

$$f(x) = (x-\alpha)^m h(x), \qquad h(\alpha) \neq 0 \tag{7.37}$$

for some continuous function $h(x)$ with $h(\alpha) \neq 0$, m a positive integer. If we assume that $f(x)$ is sufficiently differentiable, an equivalent definition is that

$$f(\alpha) = f'(\alpha) = \cdots = f^{(m-1)}(\alpha) = 0, \qquad f^{(m)}(\alpha) \neq 0. \tag{7.38}$$

A root of multiplicity $m = 1$ is called a *simple root*.

Example.

(a) $f(x) = (x-1)^2(x+2)$ has two roots. The root $\alpha = 1$ has multiplicity 2, and $\alpha = -2$ is a simple root.

(b) $f(x) = x^3 - 3x^2 + 3x - 1$ has $\alpha = 1$ as a root of multiplicity 3. To see this, note that

$$f(1) = f'(1) = f''(1) = 0, \qquad f'''(1) = 6.$$

The result follows from (7.38).

(c) $f(x) = 1 - \cos(x)$ has $\alpha = 0$ as a root of multiplicity $m = 2$. To see this, write

$$f(x) = x^2 \left[\frac{2\sin^2(\frac{x}{2})}{x^2}\right] \equiv x^2 h(x)$$

with $h(0) = \frac{1}{2}$. The function $h(x)$ is continuous for all x.

When the Newton and secant methods are applied to the calculation of a multiple root α, the convergence of $\alpha - x_n$ to zero is much slower than it would be for simple root. In addition, there is a large *interval of uncertainty* as to where the root actually lies, because of the noise in evaluating $f(x)$.

The large interval of uncertainty for a multiple root is the most serious problem associated with numerically finding such a root.

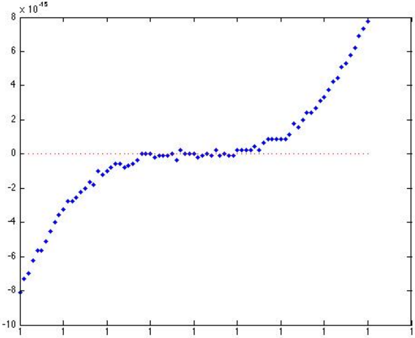


Figure: Detailed graph of $f(x) = x^3 - 3x^2 + 3x - 1$ near $x = 1$

The noise in evaluating $f(x) = (x-1)^3$, which has $\alpha = 1$ as a root of multiplicity 3. The graph also illustrates the large interval of uncertainty in finding α.

Example

To illustrate the effect of a multiple root on a rootfinding method, we use Newton's method to calculate the root $\alpha = 1.1$ of

$$\begin{aligned} f(x) = \quad & (x-1.1)^3(x-2.1) \\ & 2.7951 + x(-8.954 + x(10.56 + x(-5.4 + x))). \end{aligned} \tag{7.39}$$

The computer used is decimal with six digits in the significand, and it uses rounding. The function $f(x)$ is evaluated in the nested form of (7.39), and $f'(x)$ is evaluated similarly. The results are given in the Table.

The column "ratio" gives the values of

$$\frac{\alpha - x_n}{\alpha - x_{n-1}} \tag{7.40}$$

and we can see that these values equal about $\frac{2}{3}$.

n	x_n	$f(x_n)$	$\alpha - x_n$	Ratio
0	0.800000	0.03510	0.300000	
1	0.892857	0.01073	0.207143	0.690
2	0.958176	0.00325	0.141824	0.685
3	1.00344	0.00099	0.09656	0.681
4	1.03486	0.00029	0.06514	0.675
5	1.05581	0.00009	0.04419	0.678
6	1.07028	0.00003	0.02972	0.673
7	1.08092	0.0	0.01908	0.642

Table: Newton's Method for (7.39)

The iteration is linearly convergent with a rate of $\frac{2}{3}$.

It is possible to show that when we use Newton's method to calculate a root of multiplicity m, the ratios (7.40) will approach

$$\lambda = \frac{m-1}{m}, \qquad m \geq 1. \tag{7.41}$$

Thus, as x_n approaches α,

$$\alpha - x_n \approx \lambda(\alpha - x_{n-1}) \tag{7.42}$$

and the error decreases at about the constant rate. In our example, $\lambda = \frac{2}{3}$, since the root has multiplicity $m = 3$, which corresponds to the values in the last column of the table. The error formula (7.42) implies a much slower rate of convergence than is usual for Newton's method. With any root of multiplicity $m \geq 2$, the number $\lambda \geq \frac{1}{2}$; thus, the bisection method is always at least as fast as Newton's method for multiple roots. Of course, m must be an odd integer to have $f(x)$ change sign at $x = \alpha$, thus permitting the bisection method to be applied.

Newton' Method for Multiple Roots

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$
$$f(x) = (x-\alpha)^p h(x), \quad p \geq 0$$

Apply the fixed point iteration theorem

$$f'(x) = p(x-\alpha)^{p-1} h(x) + (x-\alpha)^p h'(x)$$
$$g(x) = x - \frac{(x-\alpha)^p h(x)}{p(x-\alpha)^{p-1} h(x) + (x-\alpha)^p h'(x)}$$

$$g(x) = x - \frac{(x-\alpha)h(x)}{ph(x) + (x-\alpha)h'(x)}$$

Differentiating

$$g'(x) = 1 - \frac{h(x)}{ph(x) + (x-\alpha)h'(x)} - (x-\alpha)\frac{d}{dx}\left[\frac{h(x)}{ph(x) + (x-\alpha)h'(x)}\right]$$

and

$$g'(\alpha) = 1 - \frac{1}{p} = \frac{p-1}{p}$$

Quasi-Newton Iterates

If $p = 1 \Rightarrow g'(\alpha) = 0$ Then by theorem 2.8 $\Rightarrow$ quadratic convergence

$$\frac{x_{k+1} - \alpha}{(x_k - \alpha)^2} \xrightarrow{k\to\infty} \frac{1}{2} g''(\alpha).$$

If $p > 1$ then by fixed point theory, theorem 2.6 $\Rightarrow$ linear convergence

$$|x_{k+1} - \alpha| \le \frac{p-1}{p} |x_k - \alpha|.$$

E.g. $p = 2, \frac{p-1}{p} = \frac{1}{2}$.

Acceleration of Newton's Method for Multiple Roots

$$f(x) = (x-\alpha)^p h(x), \qquad h(\alpha) \neq 0.$$

Assume p is known.

$$x_{k+1} = x_k - p\frac{f(x_k)}{f'(x_k)}$$

$$x_{k+1} = g(x_k)$$

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

$$g'(\alpha) = 1 - \frac{p}{p} = 0$$

$$\lim_{k\to\infty} \frac{\alpha - x_{k+1}}{(x-x_k)^2} = \frac{g''(\alpha)}{2}$$

Can run several Newton iterations to estimate p:

look at $\left|\frac{\alpha - x_{+1}}{\alpha - x_k}\right| \approx \frac{p-1}{p}$.

One way to deal with uncertainties in multiple roots:

$$\varphi(x) = f^{(p-1)}(x)$$
$$\varphi(x) = (x-\alpha)\psi(x), \qquad \psi(\alpha) \neq 0.$$

$\Rightarrow \alpha$ is a simple root for $\varphi(x)$.

Roots of polynomials

$$p(x) = 0$$

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n, \qquad a_n \neq 0.$$

Fundamental Theorem of Algebra:

$$p(x) = a_n(x - z_1)(x - z_2)\ldots(x - z_n), \qquad z_1, \ldots, z_n \in \mathbb{C}.$$

Location of real roots:

1. Descarte's rule of sign

Real coefficients

- $\nu = \#$ changes in sign of coefficients (ignore zero coefficients)
- $k = \#$ positive roots

$$k \leq \nu \qquad \text{and} \qquad k - \nu \quad \text{is even.}$$

Example: $p(x) = x^5 + 2x^4 - 3x^3 - 5x^2 - 1$.

$$\nu = 1 \Rightarrow k \leq 1 \Rightarrow k = 0 \text{ or } k = 1.$$

$$\nu - k = \begin{cases} 1, & k = 0 \qquad \text{not even} \\ 0, & k = 1. \end{cases}$$

For negative roots consider $q(x) = p(-x)$.
Apply rule to $q(x)$.

Ex.: $q(x) = -x^5 + 2x^4 + 3x^3 - 5x^2 - 1.$

$\nu = 2$

$k = 0$ or 2.

2. Cauchy

$$|\zeta_i| \leq 1 + \max_{0 \leq i \leq n-1} \left| \frac{a_i}{a_n} \right|$$

Book: Householder "The numerical treatment of single nonlinear equations", 1970.

Cauchy: given $p(x)$, consider

$$p_1(x) = |a_n|x^n + |a_{n-1}|x^{n-1} + \ldots + |a_1|x - |a_0| = 0$$
$$p_2(x) = |a_n|x^n - |a_{n-1}|x^{n-1} - \ldots - |a_1|x - |a_0| = 0$$

By Descarte's: p_i has a single positive root ρ_i

$$\rho_1 \leq |\zeta_j| \leq \rho_2.$$

Nested multiplication (Horner's method)

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n \tag{7.43}$$

$$p(x) = a_0 + x\left(a_1 + x\left(a_2 + \ldots + x\left(a_{n-1} + a_n x\right)\right)\ldots\right) \tag{7.44}$$

(7.44) requires n multiplications and n additions.

(7.43) to form $a_k x^k$:

$$x \cdot x^{k-1} : \quad 1*$$
$$a_k \cdot x^k : \quad 1*$$

$n+$ and $2n-1*$.

For any $\zeta \in \mathbb{R}$ define b_k, $k = 0, \ldots, n$.

$$\begin{aligned} b_n &= a_n \\ b_k &= a_k + \zeta b_{k+1}, \quad k = n-1, n-2, \ldots, 0 \\ p(\zeta) &= a_0 + \underbrace{\zeta \underbrace{\left(a_1 + \ldots + \zeta \underbrace{(a_{n-1} + a_n \zeta)}_{b_{n-1}} \ldots \right)}_{b_1}}_{b_0} \end{aligned}$$

Consider

$$q(x) = b_1 + b_2 x + \ldots + b_n x^{n-1}.$$

Claim:

$$p(x) = b_0 + (x - \zeta) q(x).$$

Proof.

$$\begin{aligned} & b_0 + (x - \zeta) q(x) \\ & = b_0 + (x - \zeta)(b_1 + b_2 x + \ldots + b_n x^{n-1}) \\ & = \underbrace{b_0 - \zeta b_1}_{a_0} + \underbrace{(b_1 - b_2 \zeta)}_{a_1} x + \ldots + \underbrace{(b_{n-1} - b_n \zeta)}_{a_{n-1}} x^{n-1} + \underbrace{b_n}_{a_n} x^n \\ & = a_0 + a_1 x + \ldots + a_n x^n = p(x). \qquad \square \end{aligned}$$

Note: if $p(\zeta) = 0$, then $b_0 = 0$: $\quad p(x) = (x - \zeta) q(x).$

Deflation

If ζ is found, continue with $q(x)$ to find the rest of the roots.

Newton's method for $p(x) = 0$.

$$x_{k+1} = x_k - \frac{p(x_k)}{p'(x_k)}, \qquad k = 0, 1, 2, \ldots$$

To evaluate p and p' at $x = \zeta$:

$$\begin{aligned} p(\zeta) &= b_0 \\ p'(x) &= q(x) + (x - \zeta)q'(x) \\ p'(\zeta) &= q(\zeta) \end{aligned}$$

Algorithm (Newton's method for $p(x) = 0$)

Given: $a = (a_0, a_1, \ldots, a_n)$
Output: $b = b(b_1, b_2, \ldots, b_n)$: coefficients of deflated polynomial $g(x)$; root.

Newton($a, n, x_0, \varepsilon, \texttt{itmax}, \text{root}, b, \texttt{ierr}$)

itnum = 1

1. $\zeta := x_0; b_n := a_n; c := a_n$

   for $k = n-1, \ldots, 1$ $\quad b_k := a_k + \zeta b_{k+1};$ $\quad c := b_k + \zeta c$ $\quad p'(\zeta)$

   $b_0 := a_0 + \zeta b_1$ $\quad p(\zeta)$

   - if $c = 0, \texttt{iter} = 2$, exit $\quad p'(\zeta) = 0$

     $x_1 = x_0 - b_0/c$ $\quad x_1 = x_0 - p(x_0)/p'(x_0)$
   - if $|x_0 - x_1| \leq \varepsilon$, then $\texttt{ierr} = 0$: root $= x$, exit
   - it `itnum = itmax`, then `ierr = 1`, exit

     `itnum = itnum + 1`, $x_0 := x_1$,

     quad go to 1.

Conditioning

$p(x) = a_0 + a_1 x + \ldots + a_n x^n$

roots: $\zeta_1, \ldots, \zeta_n$

Perturbation polynomial $\quad q(x) = b_0 + b_1 x + \ldots + b_n x^n$

Perturbed polynomial

$$\begin{aligned} p(x,\varepsilon) &= p(x) + \varepsilon q(x) \\ &= (a_0 + \varepsilon b_0) + (a_1 + \varepsilon b_1)x + \ldots + (a_n + \varepsilon b_n)x^n \end{aligned}$$

roots: $\zeta_j(\varepsilon)$ - continuous functions of ε, $\zeta_i(0) = \zeta_i$.

(Absolute) Conditioning number

$$k_{\zeta_j} = \lim_{\varepsilon \to 0} \frac{|\zeta_j(\varepsilon) - \zeta_j|}{|\varepsilon|}$$

Example

$$(x-1)^3 = 0 \qquad \zeta_1 = \zeta_2 = \zeta_3 = 1$$

$$(x-1)^3 - \varepsilon = 0 \qquad (q(x) = -1)$$

Set $y = x - 1$ and $a = \varepsilon^{1/3}$

$$p(x, \varepsilon) = y^3 - \varepsilon = y^3 - a^3 = (y-a)(y^2 + ya + a^2)$$

$$y_1 = a$$

$$y_{2,3} = \frac{-a \pm \sqrt{-3a^2}}{2} = \frac{-a(1 \pm i\sqrt{3})}{2}$$

$$y_2 = -a\omega, y_3 = -a\omega^2, \qquad \omega = \frac{1 - i\sqrt{3}}{2}, |\omega| = 1$$

$$\zeta_1(\varepsilon) = 1 + \varepsilon^{1/3}$$
$$\zeta_2(\varepsilon) = 1 - \omega\varepsilon^{1/3}$$
$$\zeta_3(\varepsilon) = 1 - \omega^2\varepsilon^{1/3}$$
$$|\zeta_j(\varepsilon) - 1| = \varepsilon^{1/3}$$

Conditioning number $\left|\frac{\zeta_j(\varepsilon) - 1}{\varepsilon}\right| = \frac{\varepsilon^{1/3}}{\varepsilon} \xrightarrow{\varepsilon\to 0} \infty$

If $\varepsilon = 0.001, \varepsilon^{1/3} = 0.1$, $|\zeta_j(\varepsilon) - 1| = 0.1$.

General argument

$p(x)$, Simple root ζ

$$p(\zeta) = 0$$
$$p'(\zeta) \neq 0$$
$$p(x,\varepsilon) : \text{root } \zeta(\varepsilon)$$
$$\zeta(\varepsilon) = \zeta + \sum_{\ell=1}^{\infty} \gamma_\ell \varepsilon^\ell$$
$$= \underbrace{\zeta}_{\text{this is what matters}} + \gamma_1 \varepsilon + \underbrace{\gamma_2 \varepsilon^2 + \dots}_{\text{negligeable if } \varepsilon \text{is small}}$$

$$\frac{\zeta(\varepsilon) - \zeta}{\varepsilon} = \gamma_1 + \gamma_2 \varepsilon + \dots \underset{\varepsilon \to \infty}{\longrightarrow} \gamma_1$$

To find γ_1:

$$\zeta'(0) = \gamma_1$$
$$p(\zeta(\varepsilon), \varepsilon) = 0$$
$$p(\zeta(\varepsilon)) + \varepsilon q(\zeta(\varepsilon)) = 0$$
$$p'(\zeta(\varepsilon))\zeta'(\varepsilon) + q(\zeta(\varepsilon)) + \varepsilon q'(\zeta(\varepsilon))\zeta'(\varepsilon) =$$
$$\varepsilon = 0$$
$$p'(\zeta)\underbrace{\zeta'(0)}_{\gamma_1} + q(\zeta) = 0 \Longrightarrow \gamma_1 = -\frac{q(\zeta)}{p'(\zeta)}$$
$$k_\zeta = |\gamma_1| = \left|\frac{q(\zeta)}{p'(\zeta)}\right|$$

k is large if $p'(\zeta)$ is close to zero.

Example

$$p(x) = W_7 = \prod_{i=1}^{7}(x-i)$$
$$q(x) = x^6, \varepsilon = -0.002$$
$$p'(\zeta_j) = \prod_{\ell=1,\ell\neq j}^{7}(j-\ell) \qquad \zeta_j = j$$
$$k_{\zeta j} = \left|\frac{q(\zeta_j)}{p'(\zeta_j)}\right| = \frac{j^6}{\prod_{\ell=1}^{7}(j-\ell)}$$

In particular,

$$\zeta_j(\varepsilon) \approx j + \varepsilon\frac{q(\zeta_j)}{p'(\zeta_j)} = j + \delta(j).$$

Systems of Nonlinear Equations

$$\begin{cases} f_1(x_1, \ldots, x_m) = 0, \\ f_2(x_1, \ldots, x_m) = 0, \\ \vdots \\ f_m(x_1, \ldots, x_m) = 0. \end{cases} \tag{7.45}$$

If we denote

$$\mathbf{F} = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} : \mathbb{R}^m \to \mathbb{R}^m,$$

then (7.45) is equivalent to writing

$$\mathbf{F}(\mathbf{x}) = 0. \tag{7.46}$$

Fixed Point Iteration

$$\mathbf{x} = \mathbf{G}(\mathbf{x}), \qquad \mathbf{G} : \mathbb{R}^m \to \mathbb{R}^m$$

Solution $\boldsymbol{\alpha}$: $\boldsymbol{\alpha} = \mathbf{G}(\boldsymbol{\alpha})$ is called a fixed point of $\mathbf{G}$.

Example: $\mathbf{F}(\mathbf{x}) = 0$

$$\begin{aligned} \mathbf{x} &= \mathbf{x} - A\mathbf{F}(\mathbf{x}) \quad \text{for some } A \in \mathbb{R}^{m\times m}\text{, nonsingular matrix.} \\ &= \mathbf{G}(\mathbf{x}) \end{aligned}$$

Iteration:

initial guess x_0

$$\mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n), \quad n = 0, 1, 2, \ldots$$

Recall $\mathbf{x} \in \mathbb{R}^m$

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}, \qquad 1 \le p < \infty,$$

$$\|\mathbf{x}\|_\infty = \max_{1 \le i \le n} |x_i|.$$

Matrix norms: operator induced

$A \in \mathbb{R}^{m \times m}$

$$\|A\|_p = \sup_{x \in \mathbb{R}^m, \mathbf{x} \neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}, \quad 1 \le p < \infty$$

$$\|A\|_\infty = \max_{1 \le i \le m} \|\text{Row}_i(A)\|_1$$

$$= \max_{1 \le i \le m} \sum_{j=1}^{m} |a_{ij}|$$

Recall $\mathbf{x} \in \mathbb{R}^m$

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{m} |x_i|^p \right)^{1/p}, \qquad 1 \le p < \infty,$$

$$\|\mathbf{x}\|_\infty = \max_{1\le i\le n} |x_i|.$$

Matrix norms: operator induced

$A \in \mathbb{R}^{m\times m}$

$$\|A\|_p = \sup_{x\in\mathbb{R}^m, \mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p}, \quad 1 \le p < \infty$$

$$\|A\|_\infty = \max_{1\le i\le m} \|\mathrm{Row}_i(A)\|_1$$

$$= \max_{1\le i\le m} \sum_{j=1}^{m} |a_{ij}|$$

Let $\|\cdot\|$ be any norm in $\mathbb{R}^m$.

Definition

$\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^m$ is called a contractive mapping if

$$\|\mathbf{G}(\mathbf{x}) - \mathbf{G}(\mathbf{y})\| \leq \lambda \|\mathbf{x} - \mathbf{y}\|, \qquad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^m,$$

for some $\lambda < 1$.

Contractive mapping theorem

Theorem (Contractive mapping theorem)

Assume

1. D *is a closed, bounded subset of* $\mathbb{R}^m$.
2. $\mathbf{G} : D \to D$ *is a contractive mapping.*

Then

- $\exists$ *unique* $\boldsymbol{\alpha} \in D$ *such that* $\boldsymbol{\alpha} = \mathbf{G}(\boldsymbol{\alpha})$ *(unique fixed point).*
- *For any* $\mathbf{x}_0 \in D, \mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n)$ *converges linearly to* $\boldsymbol{\alpha}$ *with rate* λ.

Proof

We will show that $\|\mathbf{x}_n\| \to \boldsymbol{\alpha}$.

$$\|\mathbf{x}_{i+1} - \mathbf{x}_i\| = \|\mathbf{G}(\mathbf{x}_i) - \mathbf{G}(\mathbf{x}_{i-1})\| \leq \lambda \|\mathbf{x}_i - \mathbf{x}_{i-1}\|$$
$$\leq \ldots \leq \lambda^i \|\mathbf{x}_1 - \mathbf{x}_0\| \qquad \text{(by induction)}$$

$$\|\mathbf{x}_k - \mathbf{x}_0\| = \|\sum_{i=0}^{k-1}(\mathbf{x}_{i+1} - \mathbf{x}_i)\| \leq \sum_{i=0}^{k-1} \|\mathbf{x}_{i+1} - \mathbf{x}_i\|$$
$$\leq \sum_{i=0}^{k-1} \lambda^i \|\mathbf{x}_1 - \mathbf{x}_0\| = \frac{1-\lambda^k}{1-\lambda}\|\mathbf{x}_1 - \mathbf{x}_0\|$$
$$< \frac{1}{1-\lambda}\|\mathbf{x}_1 - \mathbf{x}_0\|.$$

$\forall k, \ell$:

$$\begin{aligned}\|\mathbf{x}_{k+\ell} - \mathbf{x}_k\| &= \|\mathbf{G}(\mathbf{x}_{k+\ell-1}) - \mathbf{G}(\mathbf{x}_{k-1})\| \\ &\leq \lambda \|\mathbf{x}_{k+\ell-1} - \mathbf{x}_{k-1}\| \\ &\leq \ldots \leq \lambda^k \|\mathbf{x}_\ell - \mathbf{x}_0\| \\ &< \frac{\lambda^k}{1-\lambda} \|\mathbf{x}_1 - \mathbf{x}_0\| \underset{k\to\infty}{\longrightarrow} 0\end{aligned}$$

$\Rightarrow \{\mathbf{x}_n\}$ is a Cauchy sequence $\Rightarrow \{\mathbf{x}_n\} \to \boldsymbol{\alpha}$.

$$\begin{array}{c}\mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n) \\ \downarrow n\to\infty \\ \boldsymbol{\alpha} = \mathbf{G}(\boldsymbol{\alpha})\end{array}$$

$\Rightarrow$ $\boldsymbol{\alpha}$ is a fixed point.

Uniqueness: Assume $\boldsymbol{\beta} = \mathbf{G}(\boldsymbol{\beta})$

$$\|\boldsymbol{\alpha}-\boldsymbol{\beta}\| = \|\mathbf{G}(\boldsymbol{\alpha})-\mathbf{G}(\boldsymbol{\beta})\| \le \lambda\|\boldsymbol{\alpha}-\boldsymbol{\beta}\|,$$

$$\underbrace{(1-\lambda)}_{>0}\|\boldsymbol{\alpha}-\boldsymbol{\beta}\| \le 0 \Rightarrow \|\boldsymbol{\alpha}-\boldsymbol{\beta}\| = 0 \Rightarrow \boldsymbol{\alpha} = \boldsymbol{\beta}.$$

Linear convergence with rate λ:

$$\|\mathbf{x}_{n+1}-\boldsymbol{\alpha}\| = \|\mathbf{G}(\mathbf{x}_n)-\mathbf{G}(\boldsymbol{\alpha})\| \le \lambda\|\mathbf{x}_n-\boldsymbol{\alpha}\|.$$

■

Jacobian matrix

Definition

$\mathbf{F} : \mathbb{R}^m \to \mathbb{R}^m$ is continuously differentiable ($\mathbf{F} \in C^1(\mathbb{R}^m)$) if, for every $\mathbf{x} \in \mathbb{R}^m$,

$$\frac{\partial f_i(\mathbf{x})}{\partial x_j}, \qquad i, j = 1, \ldots, m$$

exist.

$$\mathbf{F}'(\mathbf{x}) \stackrel{\text{def}}{:=} \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_m} \end{pmatrix}_{m \times m}$$

$$(\mathbf{F}'(\mathbf{x}))_{ij} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}, \qquad i, j = 1, \ldots, m.$$

Mean Value Theorem

Theorem (Mean Value Theorem)

$f : \mathbb{R}^m \to \mathbb{R}$,

$$f(\mathbf{x}) - f(\mathbf{y}) = \nabla f(\mathbf{z})^T(\mathbf{x} - \mathbf{y})$$

for some $\mathbf{z} \in \overline{\mathbf{x}, \mathbf{y}}$, *where* $\nabla f(\mathbf{z}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_1} \end{pmatrix}$.

Proof: Follows immediately from Taylor's theorem (linear Taylor expansion). Since

$$\nabla f(\mathbf{z})^T(\mathbf{x} - \mathbf{y}) = \frac{\partial f(\mathbf{z})}{\partial x_1}(x_1 - y_1) + \ldots + \frac{\partial f(\mathbf{z})}{\partial x_m}(x_m - y_m).$$

No Mean Value Theorem for vector value functions

Note:

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix},$$

$$f_i(\mathbf{x}) - f_i(\mathbf{y}) = \nabla f_i(\mathbf{z}_i)^T(\mathbf{x}-\mathbf{y}), \qquad i = 1, \ldots, n.$$

It is not true that

$$\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y}) = \mathbf{F}'(\mathbf{z})(\mathbf{x}-\mathbf{y})$$

Consider $\mathbf{x} = \mathbf{G}(\mathbf{x})$, $(\mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n))$ with solution $\boldsymbol{\alpha} = \mathbf{G}(\boldsymbol{\alpha})$.

$$\begin{aligned}\alpha_i - (\mathbf{x}_{n+1})_i &= g_i(\boldsymbol{\alpha}) - g_i(\mathbf{x}_n) \\ &\overset{MVT}{=} \nabla g_i(z_n^i)^T(\boldsymbol{\alpha} - \mathbf{x}_n), \qquad i = 1, \ldots, m\end{aligned}$$

$$\boldsymbol{\alpha} - \mathbf{x}_{n+1} = \underbrace{\begin{pmatrix} \nabla g_1(\mathbf{z}_1)^T \\ \vdots \\ \nabla g_m(\mathbf{z}_m)^T \end{pmatrix}}_{J_n} (\boldsymbol{\alpha} - \mathbf{x}_n) \qquad \mathbf{z}_j \in \overline{\boldsymbol{\alpha}, \mathbf{x}_n}$$

$$\boldsymbol{\alpha} - \mathbf{x}_{n+1} = J_n(\boldsymbol{\alpha} - \mathbf{x}_n) \tag{7.47}$$

If $\mathbf{x}_n \to \boldsymbol{\alpha}$, $J_n \to \begin{pmatrix} \nabla g_1(\boldsymbol{\alpha})^T \\ \vdots \\ \nabla g_m(\boldsymbol{\alpha})^T \end{pmatrix} = \mathbf{G}'(\boldsymbol{\alpha})$.

The size of $\mathbf{G}'(\boldsymbol{\alpha})$ will affect convergence.

Theorem 2.9

Assume

- D is closed, bounded, convex subset of $\mathbb{R}^m$.
- $\mathbf{G} \in C^1(D)$
- $\mathbf{G}(D) \subset D$
- $\lambda = \max_{\mathbf{x}\in D} \|\mathbf{G}'(\mathbf{x})\|_\infty < 1$.

Then

(i) $\mathbf{x} = \mathbf{G}(\mathbf{x})$ has a unique solution $\alpha \in D$

(ii) $\forall \mathbf{x}_0 \in D, \mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n)$ converges to $\boldsymbol{\alpha}$.

(iii) $\|\boldsymbol{\alpha} - \mathbf{x}_{n+1}\|_\infty \le (\|\mathbf{G}'(\boldsymbol{\alpha})\|_\infty + \varepsilon_n)\, \|\boldsymbol{\alpha} - \mathbf{x}_n\|_\infty$,

whenever $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$.

Proof: $\forall \mathbf{x}, \mathbf{y} \in D$

$$
\begin{aligned}
|g_i(\mathbf{x}) - g_i(\mathbf{y})| &\le \left|\nabla g_i(\mathbf{z}_i)^T(\mathbf{x}-\mathbf{y})\right|, \quad \mathbf{z}_i \in \overline{\mathbf{x},\mathbf{y}} \\
&= \left|\sum_{j=1}^{m} \frac{\partial g_i(\mathbf{z}_i)}{\partial x_j}(x_j - y_j)\right| \le \sum_{j=1}^{m}\left|\frac{\partial g_i(\mathbf{z}_i)}{\partial x_j}\right| |x_j - y_j| \\
&\le \sum_{j=1}^{m}\left|\frac{\partial g_i(\mathbf{z}_i)}{\partial x_j}\right| \|\mathbf{x}-\mathbf{y}\|_\infty \le \|\mathbf{G}'(\mathbf{z}_i)\|_\infty \|\mathbf{x}-\mathbf{y}\|_\infty
\end{aligned}
$$

$$\Rightarrow \|\mathbf{G}(\mathbf{x}) - \mathbf{G}(\mathbf{y})\|_\infty \le \|\mathbf{G}'(\mathbf{z}_i)\|_\infty \|\mathbf{x}-\mathbf{y}\|_\infty \le \lambda \|\mathbf{x}-\mathbf{y}\|_\infty$$

$\Rightarrow$ $\mathbf{G}$ is a contractive mapping. $\Rightarrow$ (i) and (ii).

To show (iii), from (7.47):

$$
\begin{aligned}
\|\boldsymbol{\alpha} - \mathbf{x}_{n+1}\|_\infty &\le \|J_n\|_\infty \|\boldsymbol{\alpha} - \mathbf{x}_n\|_\infty \\
&\le \Big(\underbrace{\|J_n - \mathbf{G}'(\boldsymbol{\alpha})\|_\infty}_{\varepsilon_n \xrightarrow[n\to\infty]{} 0} + \|\mathbf{G}'(\boldsymbol{\alpha})\|_\infty\Big)\|\boldsymbol{\alpha} - \mathbf{x}_n\|_\infty. \qquad \blacksquare
\end{aligned}
$$

Example (p.104)

Solve

$$\begin{cases} f_1 \equiv 3x_1^2 + 4x_2^2 - 1 = 0 \\ f_2 \equiv x_2^3 - 8x_1^3 - 1 = 0 \end{cases}, \text{ for } \boldsymbol{\alpha} \text{ near } (x_1, x_2) = (-.5, .25).$$

Iteratively

$$\begin{bmatrix} x_{1,n+1} \\ x_{2,n+1} \end{bmatrix} = \begin{bmatrix} x_{1,n} \\ x_{2,n} \end{bmatrix} - \begin{bmatrix} .016 & -.17 \\ .52 & -.26 \end{bmatrix} \begin{bmatrix} 3x_{1,n}^2 + 4x_{2,n}^2 - 1 \\ x_{2,n}^3 - 8x_{1,n}^3 - 1 \end{bmatrix}$$

Example (p.104)

$$\mathbf{x}_{n+1} = \underbrace{\mathbf{x}_n - A\mathbf{F}(\mathbf{x}_n)}_{\mathbf{G}(\mathbf{x})}$$

$$\|\mathbf{G}'(\boldsymbol{\alpha})\|_\infty \approx 0.04, \qquad \frac{\|\boldsymbol{\alpha} - \mathbf{x}_{n+1}\|_\infty}{\|\boldsymbol{\alpha} - \mathbf{x}_n\|_\infty} \longrightarrow 0.04, \qquad A = \left(\mathbf{F}'(\mathbf{x}_0)\right)^{-1}$$

Why?

$$\mathbf{G}'(\mathbf{x}) = \mathbf{I} - A\mathbf{F}'(\mathbf{x}), \qquad \mathbf{G}'(\boldsymbol{\alpha}) = \mathbf{I} - A\mathbf{F}'(\boldsymbol{\alpha})$$

Need

$$\|\mathbf{G}'(\boldsymbol{\alpha})\|_\infty \approx 0$$

$$\mathbf{A} \approx \left(\mathbf{F}'(\boldsymbol{\alpha})\right)^{-1}, \qquad A = \left(\mathbf{F}'(\mathbf{x}_0)\right)^{-1}$$

m dimensional Parallel Chords Method

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \left(\mathbf{F}'(\mathbf{x}_0)\right)^{-1}\mathbf{F}(\mathbf{x}_n)$$

Newton's Method for $\mathbf{F}(\mathbf{x}) = 0$

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \left(\mathbf{F}'(\mathbf{x}_n)\right)^{-1} \mathbf{F}(\mathbf{x}_n), \qquad n = 0, 1, 2, \ldots$$

Given initial guess:

$$f_i(\mathbf{x}) = f_i(\mathbf{x}_0) + \nabla f_i(\mathbf{x}_0)^T (\mathbf{x} - \mathbf{x}_0) + \underbrace{O\left(\|\mathbf{x} - \mathbf{x}_0)\|^2\right)}_{neglect}$$

$$\mathbf{F}(\mathbf{x}) \approx \mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \equiv M_0(\mathbf{x})$$

$M_0(\mathbf{x})$: linear model of $\mathbf{F}(\mathbf{x})$ around $\mathbf{x}_0$.
Set $\mathbf{x}_1$: $\quad M_0(\mathbf{x}_1) = 0$

$$\mathbf{F}(\mathbf{x}_0) + \mathbf{F}'(\mathbf{x}_0)(\mathbf{x}_1 - \mathbf{x}_0) = 0$$

$$\mathbf{x}_1 = \mathbf{x}_0 - \left(\mathbf{F}'(\mathbf{x}_0)\right)^{-1} \mathbf{F}(\mathbf{x}_0)$$

In general, Newton's method:

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \left(\mathbf{F}'(\mathbf{x}_n)\right)^{-1}\mathbf{F}(\mathbf{x}_n)$$

Geometric interpretation:

$$m_i(\mathbf{x}) = f_i(\mathbf{x}_0) + \nabla f_i(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0), \qquad i = 1, \ldots, m$$

$f_i(\mathbf{x})$: surface

$m_i(\mathbf{x})$: tangent at $\mathbf{x}_0$.

In practice:

1. Solve a linear system $\mathbf{F}'(\mathbf{x}_n)\boldsymbol{\delta}_n = -\mathbf{F}(\mathbf{x}_n)$
2. Set $\mathbf{x}_{n+1} = \mathbf{x}_n + \boldsymbol{\delta}_n$

Convergence Analysis: 1. Use the fixed point iteration theorem

$$\mathbf{F}(\mathbf{x}) = 0, \qquad \mathbf{x} = \mathbf{x} - (\mathbf{F}'(\mathbf{x}))^{-1}\mathbf{F}(\mathbf{x}) = \mathbf{G}(\mathbf{x}), \qquad \mathbf{x}_{n+1} = \mathbf{G}(\mathbf{x}_n)$$

Assume $\mathbf{F}(\boldsymbol{\alpha}) = 0$, $\mathbf{F}'(\boldsymbol{\alpha})$ is nonsingular.
Then $\mathbf{G}'(\boldsymbol{\alpha}) = 0$ (exercise !)

$$\|\mathbf{G}'(\boldsymbol{\alpha})\|_\infty = 0$$

If $\mathbf{G}' \in C^1(B_r(\boldsymbol{\alpha}))$ where $B_r(\boldsymbol{\alpha}) = \{\mathbf{y} : \|\mathbf{y} - \boldsymbol{\alpha}\| \leq r\}$, by continuity:

$$\|\mathbf{G}'(\boldsymbol{\alpha})\|_\infty < 1 \qquad \text{for } \mathbf{x} \in B_{\overline{r}}(x)$$

for some $\overline{r}$.
By Theorem 2.9 $D = B_{\overline{r}} \Rightarrow$ linear convergence.

Convergence Analysis: 2. Assume: $\mathbf{F}'(\mathbf{x}) \in Lip_\gamma(D)$
$(\|\mathbf{F}'(\mathbf{x}) - \mathbf{F}'(\mathbf{y})\| \leq \gamma\|\mathbf{x} - \mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in D)$

Theorem

Assume

- $\mathbf{F}' \in C^1(D)$
- $\exists \boldsymbol{\alpha} \in D$ such that $\mathbf{F}(\boldsymbol{\alpha}) = 0$
- $\mathbf{F}'(\boldsymbol{\alpha}) \in Lip_\gamma(D)$
- $\exists \quad (\mathbf{F}'(\boldsymbol{\alpha}))^{-1}$ and $\|\mathbf{F}'(\boldsymbol{\alpha})\|^{-1} \leq \beta$

Then $\exists \varepsilon > 0$ such that if $\|\mathbf{x}_0 - \boldsymbol{\alpha}\| < \varepsilon$,
$\Longrightarrow \mathbf{x}_{n+1} = \mathbf{x}_n - (\mathbf{F}'(\mathbf{x}_n))^{-1}\mathbf{F}(\mathbf{x}_n) \to \alpha$ and $\|\mathbf{x}_{n+1} - \boldsymbol{\alpha}\| \leq \beta\gamma\|\mathbf{x}_n - \boldsymbol{\alpha}\|^2$.

($\beta\gamma$: measure of nonlinearity)
So, need $\varepsilon < \frac{1}{\beta\gamma}$.

Reference: Dennis & Schnabel, SIAM.

Quasi - Newton Methods

$$\mathbf{x}_{n+1} = \mathbf{x}_n - A_n^{-1}\mathbf{F}(\mathbf{x}_n), \qquad A_n \approx \mathbf{F}'(\mathbf{x}_n)$$

Ex.: Finite Difference Newton

$$A_n = a_{ij} = \frac{f_i(\mathbf{x}_n + h_n\mathbf{e}_j) - f_i(\mathbf{x}_n)}{h_n} \approx \frac{\partial f_i(\mathbf{x}_n)}{\partial x_j},$$

$$h_n \approx \sqrt{\delta} \text{ and } \mathbf{e}_j = (0\dots0 \quad \underset{j^{th} position}{\underset{\uparrow}{1}} \quad 0\dots0)^T.$$

Global Convergence

Newton's method:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + s_n \mathbf{d}_n$$

when

$$d_n = -(\mathbf{F}(\mathbf{x}_n)')^{-1}\mathbf{F}(\mathbf{x}_n)$$
$$s_n = 1$$

If Newton step s_n not satisfactory, e.g. $\|\mathbf{F}(\mathbf{x}_{n+1})\|_2 > \|\mathbf{F}(\mathbf{x}_n)\|_2$

$$s_n \longleftarrow g s_n \qquad \text{for some } g < 1 \qquad \text{(backtracking)}$$

We can choose s_n such that

$$\varphi(s) = \|\mathbf{F}(\mathbf{x}_n + s\mathbf{d}_n)\|_2$$

is minimized. Line Search

In practice: minimize a quadratic model of $\varphi(s)$.

Trust region

Set a region in which the model of the function is reliable. If Newton step takes us outside this region, cut it to be inside the region

(See Optimization Toolbox of Matlab)

The MATLAB instruction

```
zero = fsolve('fun', x0)
```

allows the computation of one zero of a nonlinear system

$$\begin{cases} f_1(x_1, x_2, \ldots, x_n) = 0, \\ f_2(x_1, x_2, \ldots, x_n) = 0, \\ \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0, \end{cases}$$

defined through the user function `fun` starting from the vector x0 as initial guess.

The function `fun` returns the n values $f_1(x), \ldots, f_n(x)$ for any value of the input vector x.

For instance, let us consider the following system:

$$\begin{cases} x^2 + y^2 = 1, \\ \sin(\pi x/2) + y^3 = 0, \end{cases}$$

whose solutions are (0.4761, -0.8794) and (-0.4761,0.8794).
The corresponding Matlab user function, called systemnl, is defined as:

```
function fx=systemnl(x)
fx = [x(1)^2+x(2)^2-1;
      sin(pi*0.5*x(1) )+x(2)^3;]
```

The Matlab instructions to solve this system are therefore:

```
>> x0 = [1 1];
>> options=optimset('Display','iter');
>> [alpha,fval] = fsolve('systemnl',x0,options)
alpha =
    0.4761   -0.8794
```

Using this procedure we have found only one of the two roots. The other can be computed starting from the initial datum -x0.

$$f(x_1, x_2, \ldots, x_n) : \mathbb{R}^n \to \mathbb{R};$$
$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$$

Theorem (first order necessary condition for a minimizer)

If $f \in C^1(D), D \subset \mathbb{R}^n$, and $\mathbf{x} \in D$ is a local minimizer then $\nabla f(\mathbf{x}) = 0$.

Solve:

$$\nabla f(\mathbf{x}) = 0 \text{ with } \nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}. \tag{7.48}$$

Hamiltonian

Apply Newton's method for (7.48) with $\mathbf{F}(\mathbf{x}) = \nabla f(\mathbf{x})$.
Need $\mathbf{F}'(\mathbf{x}) = \nabla^2 f(\mathbf{x}) = \mathbf{H}(\mathbf{x})$.

$$\mathbf{H}_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{H}(\mathbf{x}_n)^{-1} \nabla f(\mathbf{x}_n)$$

If $\mathbf{H}(\boldsymbol{\alpha})$ is nonsingular and is Lip_γ, then $\mathbf{x}_n \to \boldsymbol{\alpha}$ quadratically.
Problems:

1. Not globally convergent.
2. Requires solving a linear system each iteration.
3. Requires ∇f and $\mathbf{H}$.
4. May not converge to a minimum.

Could converge to a maximum or saddle point.

1. Globalization strategy (Line Search, Trust Region).
2. Secant Approximation to $\mathbf{H}$.
3. Finite Difference derivatives for ∇f not for $\mathbf{H}$.

Theorem (necessary and sufficient conditions for a minimizer):

4. Assume $f \in C^2(D), D \subset \mathbb{R}^2$, $\exists \mathbf{x} \in D$ such that $\nabla f(\mathbf{x}) = 0$. Then $\mathbf{x}$ is a local minimum if and only if $\mathbf{H}(\mathbf{x})$ is symmetric positive semidefinite $(\mathbf{v}^T\mathbf{H}\mathbf{v} \geq 0 \quad \forall \mathbf{v} \in \mathbb{R}^n)$

$$\mathbf{x} \text{ is a local minimum} \qquad f(\mathbf{x}) \leq f(\mathbf{y}) \qquad \text{for } \forall \mathbf{y} \in B_r(\mathbf{x}).$$

Quadratic model for $f(\mathbf{x})$

Taylor:

$$m_n(\mathbf{x}) = f(\mathbf{x}_n) + \nabla f(\mathbf{x}_n)^T(\mathbf{x} - \mathbf{x}_n) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_n)^T \mathbf{H}(\mathbf{x}_n)(\mathbf{x} - \mathbf{x}_n)$$

$$m_n(\mathbf{x}) \approx f(\mathbf{x}) \qquad \text{for } \mathbf{x} \text{ near } \mathbf{x}_n.$$

Newton's method:

$$\mathbf{x}_{n+1} \text{ such that } \nabla m_n(\mathbf{x}_{n+1}) = 0$$

We need to guarantee that Hessian of the quadratic model is symmetric positive definite

$$\nabla^2 m_n(\mathbf{x}_n) = \mathbf{H}(\mathbf{x}_n).$$

Modify

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \widetilde{\mathbf{H}}^{-1}(\mathbf{x}_n)\nabla f(\mathbf{x}_n)$$

where

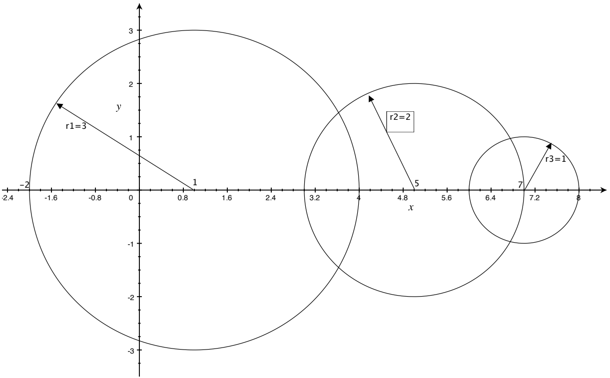
$$\widetilde{\mathbf{H}}(\mathbf{x}_n) = \mathbf{H}(\mathbf{x}_n) + \mu_n I,$$

for some $\mu_n \geq 0$.
If $\lambda_1, \ldots, \lambda_n$ eigenvalues of $\mathbf{H}(\mathbf{x}_n)$ and $\widetilde{\lambda}_1, \ldots, \widetilde{\lambda}_n$ eigenvalues of $\widetilde{\mathbf{H}}$.

$$\widetilde{\lambda}_i = \lambda_i + \mu_n$$

Need: $\mu_n : \lambda_{mm} + \mu_n > 0$.
Ghershgorin Theorem: $A = (a_{ij})$, eigenvalues lie in circles with centers a_{ii} and radius $r = \sum_{j=1, j\neq i}^{n} |a_{ij}|$.

Gershgorin Circles



Descent Methods

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{H}(\mathbf{x}_n)^{-1} \nabla f(\mathbf{x}_n)$$

Definition

$\mathbf{d}$ is a descent direction for $f(\mathbf{x})$ at point $\mathbf{x}_0$ if

$f(\mathbf{x}_0) > f(\mathbf{x}_0 + \boldsymbol{\alpha}\mathbf{d})$ for $0 \leq \boldsymbol{\alpha} < \boldsymbol{\alpha}_0$

Lemma

$\mathbf{d}$ *is descent direction if and only if* $\nabla f(\mathbf{x}_0)^T \mathbf{d} < 0$.

Newton: $\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{d}_n$, $\mathbf{d}_n = -\mathbf{H}(\mathbf{x}_n)^{-1} \nabla f(\mathbf{x}_n)$.
$\mathbf{d}_n$ is a descent direction if $\mathbf{H}(\mathbf{x}_n)$ is symmetric positive definite.

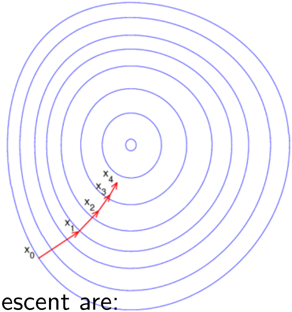
$$\nabla f(\mathbf{x}_n)^T \mathbf{d}_n = -\nabla f(\mathbf{x}_n)^T \mathbf{H}(\mathbf{x}_n)^{-1} \nabla f(\mathbf{x}_n) < 0$$

since $\mathbf{H}(\mathbf{x}_n)$ is symmetric positive definite.

Method of Steepest Descent

$$\mathbf{x}_{n+1} = \mathbf{x}_n + s_n \mathbf{d}_n, \qquad \mathbf{d}_n = -\nabla f(\mathbf{x}_n), \qquad s_n = \min_{s>0} f(\mathbf{x}_n + s\mathbf{d}_n)$$

Level curve: $C = \{\mathbf{x} | f(\mathbf{x}) = f(\mathbf{x}_0)\}$.
If C is closed and contains $\boldsymbol{\alpha}$ in the interior, then the method of steepest descent converges to $\boldsymbol{\alpha}$. Convergence is linear.



Weaknesses of gradient descent are:

1. The algorithm can take many iterations to converge towards a local minimum, if the curvature in different directions is very different.
2. Finding the optimal s_n per step can be time-consuming. Conversely, using a fixed s_n can yield poor results. Methods based on Newton's method and inversion of the Hessian using conjugate gradient techniques are often a better alternative.