Numerical Analysis (10th ed)

R.L. Burden, J. D. Faires, A. M. Burden

Chapter 2

Solutions of Equations in One Variable

Chapter 2.1: The Bisection Method*

One of the most basic root-finding methods is the Bisection method. This method is based on the Intermediate Value Theorem and generates a sequence of approximate solutions to $f(x) = 0$ that converge to a root of *f,* provided *f* is continuous on the interval where we believe a root exists. The main "take-aways" in this section are:

• How do we use the Bisection method to generate the sequence?
• How many steps do we need to insure that the root is within a certain specified tolerance?

MOTIVATION: Given a function f that is continuous on [a,b] where a solution to f(x) = 0 is expected to exist, the Bisection method begins with finding the functional values at the endpoints, namely, f(a) and f(b) . If these functional values change sign, then we know that the graph will cross the x-axis at least once on the interval [a,b]. That is, f(x) = 0 will have at least one solution on that interval.

We use this knowledge to approximate the solution by the Bisection method

1. Find the midpoint, c, of the interval [a,b] and check for sign changes between f(a), f(b) , and f(c) .
2. Let's assume, for illustration, that the functional values change sign at x = a and x = c . We then modify the interval from [a,b] to [a,c] and begin the process again on the modified interval.
3. We continue this process until we reach an approximation to the solution of f(x) = 0 that is accurate to within a specified tolerance.

The absolute error for the Bisection method is $|p_n - p| \le \frac{b-a}{2^n}$, where p is a zero of f and $n \ge 1$. If we want to determine how many steps are needed to insure that $|p_n - p| \le \frac{b-a}{2^n} < Tol$ then we use logarithms to find the value of n that will give us the tolerance we are looking for. In other words, $2^n > (b-a) \cdot Tol$ or $n > \frac{\log_{10}((b-a) \cdot Tol)}{\log_{10}(2)}$. A really nice animation of the Bisection algorithm can be seen by clicking on the following Bisection example found at:
http://mathfaculty.fullerton.edu/mathews//a2001/animations/rootfinding/BisectionMethod/BisectionMethod.html

Chapter 2.2: Fixed-point Iteration*

The questions we would like to answer in this section are:
• What is a fixed point of a function?
• How do we find a fixed point of a function?
• What are fixed points used for?

A fixed point is a an invariant value, p, for which the functional value at p equals p. That is $f(p) = p$. Not all functions have fixed points. Suppose $f(x) = x^2 - 2x$. We can easily see that $x = 0$ and $x = 3$ are fixed points of f since $f(0) = 0$ and $f(3) = 3$. However, $f(x) = x^2 - 2x + 3$ has no fixed points since there are no values, p, for which $f(p) = p$.

Example 1

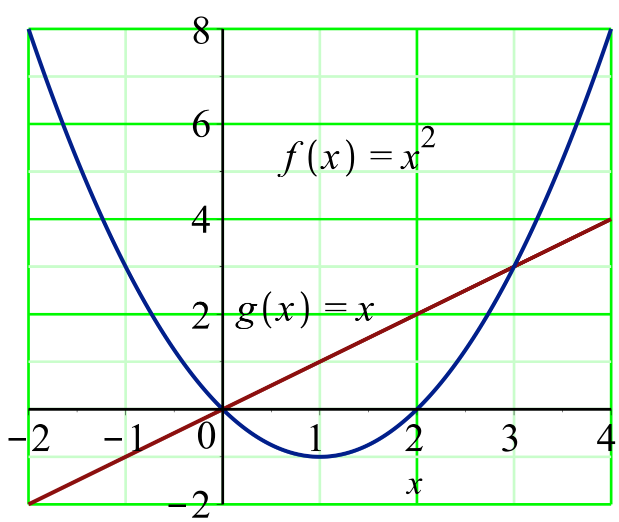
If we solve $x^2 - 2x = x$ for x, we see that $x = 0$ and $x = 3$. Plugging these values back into f will show that $f(0) = 0$ and $f(3) = 3$.

$solve(x^2 - 2x = x)$

0, 3 **(2.1.1)**

Graphically, we see that the curve $f(x) = x^2 - 2x$ and the straight line $g(x) = x$ intersect at $x = 0$ and $x = 3$ indicating that these values are fixed points of f.

$plot(\{x^2 - 2x, x\}, x = -2..4, axis = [gridlines = [colour = green, majorlines = 2]])$



Example 2

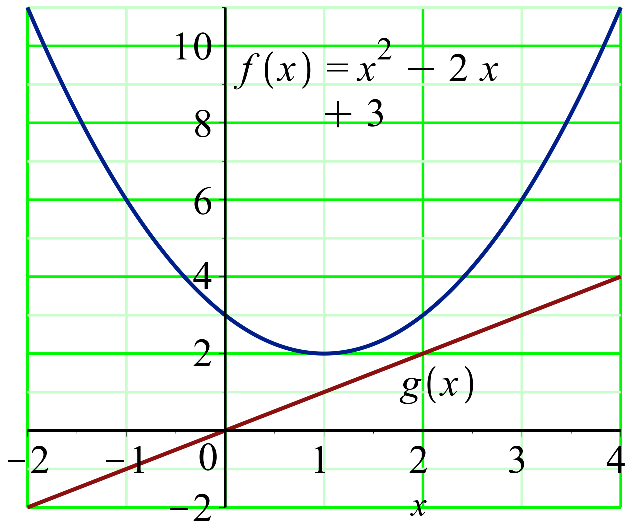
If we solve $x^2 - 2x + 3 = x$ for x, we see that there are no real solutions, and hence, no fixed points.

$solve(x^2 - 2x + 3 = x)$

$$\frac{3}{2}+\frac{1}{2}\,\mathrm{I}\sqrt{3},\ \frac{3}{2}-\frac{1}{2}\,\mathrm{I}\sqrt{3} \tag{2.2.1}$$

Graphically, we see that the curve $f(x) = x^2 - 2x + 3$ and the straight line $g(x) = x$ indicating that there are no fixed points of *f*.

$plot\left(\{x^2 - 2x + 3, x\}, x = -2..4, axis = [gridlines = [colour = green, majorlines = 2]]\right)$



Theorem 2.3 in this section gave sufficient conditions for the existence and uniqueness of a fixed point. In Example 1, because $f(x)$ is continuous on the interval $\left[-\frac{1}{4}, 1\right]$ and *f* maps every *x* in that interval back into the interval, theorem 2.3 guarantees at least one fixed point on the interval. We can also find a similar interval about $x = 3$ for which this theorem applies. However, this clearly is not the case for the second example.

Fixed points are important in many disciplines. In general, they describe points of equilibrium of a process (function), if they exist, which can then be analyzed. The fixed point iteration generates a sequence of values which, depending upon the starting value used, will converge to a fixed point if it exists.

The following steps describe the fixed point iteration:

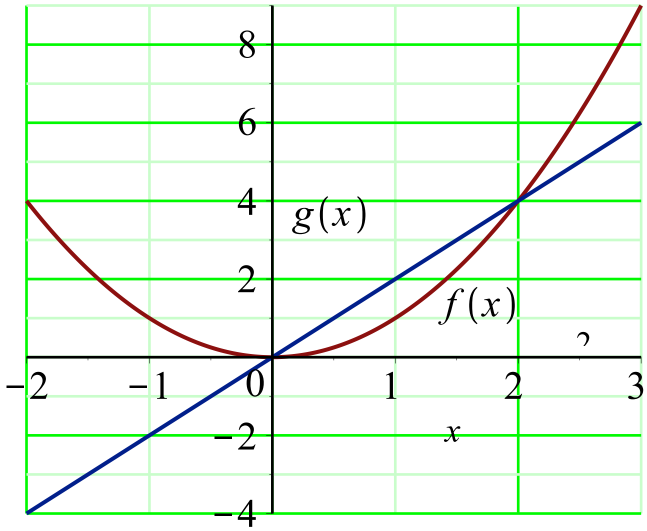
1. Identify a starting value for *x,* call it p_0.
2. Compute $f(p_0) = p_1$
3. Compute $f(p_1) = p_2$
4. We continue this process until we reach an approximation to the solution of f(x) = 0 that is accurate to within a specified tolerance.

Example 3

How can we use the fixed point iteration to find the roots of the $x^2 - 2x = 0$? To do this, we solve the equation for x^2 and proceed as before. Of course, we know that the roots are $x = 0$ and $x = 2$.

Graphically, we see that the curve $f(x) = x^2$ and the straight line $g(x) = 2x$ intersect at the two fixed points of f, namely, $x = 0$ and $x = 2$.

$plot(\{x^2, 2x\}, x = -2..3, axis = [gridlines = [colour = green, majorlines = 2]])$



Several really nice animations of the fixed point method can be seen by clicking on the following URL:

http://mathfaculty.fullerton.edu/mathews//a2001/animations/rootfinding/FixedPoint/FixedPoint.html

Chapter 2.3: Newton's Method and Its Extensions*

The key questions we want to address in this section are:
• How do we use Newton's method to find the root of an equation?
• How do we use the Secant method to find the root of an equation?
• How do we use the method of False Position used to find the root of an equation?

Newton's Method:
Newton's method is derived from the first Taylor polynomial. As long as the function f is twice continuously differentiable on $[a, b]$
and an approximation $p_0 \in [a, b]$ to the solution p is such that $f'(p_0) \neq 0$, Taylor's polynomial of degree one expanded about p_0 and evaluated at $x = p$
is given by

$$f(p) = f(p_0) + (p - p_0) f'(p_0) + \frac{(p - p_0)}{2} f''(\zeta(p))$$

where $\zeta(p)$ lies between p and p_0. Recall that the last term is called the remainder and provides a formula for the bound on the error for this method. Since p is a solution to the equation, $f(p) = 0$ and we have

$$0 = f(p_0) + (p - p_0) f'(p_0) + \frac{(p - p_0)}{2} f''(\zeta(p))$$

If we ignore the error term, solving for p, we then have that

$$p \approx p_0 - \frac{f(p_0)}{f'(p_0)}.$$

Thus, the initial approximation to the solution p is found by computing $p_1 = p_0 - \frac{f(p_0)}{f'(p_0)}$ and this represents the point at which the tangent line at the point $(p_0, f(p_0))$ crosses the x-axis . We can compute a second approximation p_2 to p if we replace p_0 with p_1 in our computation. We can continue this process until we achieve our desired accuracy.

As is stated in the text, the major drawback for this method is the need to know the value of the derivative at each approximation.

Several nice animations of Newton's method can be found at the following URL:
http://mathfaculty.fullerton.edu/mathews//a2001/animations/rootfinding/NewtonMethod/Newtonee.html

Secant Method:

The secant is a slight modification of Newton's method. The definition of the first derivative is used to replace the derivative in Newton's method.
This leads to the formula

$$p_n = p_{n-1} - \frac{f(p_n)(p_{n-1} - p_{n-2})}{f(p_{n-1}) - f(p_{n-2})}$$

which is, in fact, the equation of the secant line through $(p_{n-1}, f(p_{n-1}))$ and $(p_{n-2}, f(p_{n-2}))$ with x-intercept $(p_n, 0)$. Sequential application of the formula to successive approximations to p leads to a sequence that hopefully converges to the root of f.

Several nice animations of the Secant method can be found at the following URL:
http://mathfaculty.fullerton.edu/mathews//a2001/animations/rootfinding/SecantMethod/SecantMethod.html

Method of False Position (Regula Falsi):

This method differs from the Secant method only in the way the results are handled once we have them. In the Secant method we simply compute successive approximations to the root using the formula until we have reach the desired level of accuracy. In the method of False Position, we refine interval $[p_{n-2}, p_{n-1}]$ so that it always brackets the root, as we do with the with the Bisection method.

A nice animation of the method of False Position can be found at the following URL:
http://mathfaculty.fullerton.edu/mathews//a2001/animations/rootfinding/RegulaFalsi/RegulaFalsi.html

Chapter 2.4: Error Analysis for Iterative Methods*

The key questions we want to address in this section are:
• How do we determine the rate of convergence of a method?
• How do we determine the number of iterations necessary to achieve the desired level of convergence?

To determine the convergence rate of a sequence of approximations $\{p_n\}_{n=0}^{\infty}$ to p, we look at the following equation and two cases:

$$\lambda = \lim_{n \to \infty} \frac{|p_{n+1} - p|}{|p_n - p|^{\alpha}}$$

• if $\alpha = 1$(and $\lambda < 1$), the sequence is linearly convergent
• if $\alpha = 2$, the sequence is quadratically convergent.

The Illustration in the text shows a good example as to how this computation is done.

To determine the number of iterations necessary to achieve the desired level of convergence, given that we know the sequence formula in terms of N, we can solve the following inequality for N.
$|p_{n+1} - p| < TOL$

Chapter 2.5: Accelerating Convergence*

The key question in this section is:
• How do we accelerate convergence?

Theorem 2.8 from the last section implies that higher-order convergence for fixed-point methods of the form $g(p) = p$ can occur only when $g'(p) = 0$.

The method this discussed in this section is Aitken's Δ2 method. We begin by making an important assumption, namely that the finite sequence $\{p_n\}_{n=0}^{N}$ or the infinite sequence $\{p_n\}_{n=0}^{\infty}$ linearly converges to some limit p. We define a new sequence using the following formula:

$$\hat{p}_n = p_n - \frac{(p_{n+1} - p_n)^2}{p_{n+2} - 2p_{n+1} + p_n} \text{ for } n = 0, 1, ..., N-2$$

or

$$\hat{p}_n = p_n - \frac{(p_{n+1} - p_n)^2}{p_{n+2} - 2p_{n+1} + p_n} \text{ for } n = 0, 1, ..., \infty$$

For instance, suppose we want to generate the first three terms of a sequence {pn} generated by using Aitken's Δ2 method where $p_n = 3^{-p_{n-1}}$ and $p_0 = 0.5$.

n	p_n	$\hat{p}_n$
0	0.5	$0.5 - \dfrac{(0.5773502692 - 0.5)^2}{0.5303150046 - 2\cdot 0.5773502692 + 0.5} =$ 0.5481009646
1	$3^{-0.5} =$ 0.5773502692	$0.5773502692 - \dfrac{(0.5303150046 - 0.5773502692)^2}{0.5584386128 - 2\cdot 0.5303150046 + 0.5773502692}$ = 0.5479150738
2	$3^{-0.5773502692} =$ 0.5303150046	$0.5303150046 - \dfrac{(0.5584386128 - 0.5303150046)^2}{0.5414483921 - 2\cdot 0.5584386128 + 0.5303150046}$ = 0.5478470419
3	$3^{-0.5303150046} =$ 0.5584386128	$0.5584386128 - \dfrac{(0.5414483921 - 0.5584386128)^2}{0.5516497984 - 2\cdot 0.5414483921 + 0.5584386128}$ = 0.5478225655
4	$3^{-0.5584386128} =$ 0.5414483921	
5	$3^{-0.5414483921} =$ 0.5516497984	

Applying Aitken's Δ^2 method to a linearly convergent fixed point iteration yields Steffensen's method. This method upgrades the convergence to quadratic. Recall that a fixed point iteration is one in which $p = g(p)$. Steffensen's method proceeds as follows:

- begin with an initial approximation p_0^0 to p
- set $p_1^0 = g\left(p_0^0\right)$ and $p_2^0 = g\left(p_1^0\right)$
- determine the first Aitken's Δ^2 approximation p_0^1
- set $p_1^1 = g\left(p_0^1\right)$ and $p_2^0 = g\left(p_1^0\right)$
- continue the process until the desired TOL is achieved

The text gives a nice example of this process.

Chapter 2.6: Zeros of Polynomials and Müller's Method

The main concepts in this section are:
• $P(x)$?
• How do we deal with a polynomial $P(x)$ that has complex zeros?

Horner's method:

Horner's method uses synthetic division to re-write the polynomial
$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots a_1 x + a_0$
as
$P(x) = (x - x_0) Q(x) + b_0$ where $Q(x) = b_n x^{n-1} + b_{n-1} x^n + \ldots b_2 x + b_1$.

With the assumptions that $b_n = a_n$ and $b_k = a_k + b_{k+1} x_0$, for $k = n-1, n-2, \ldots, 1, 0$, we see that for an initial
approximation, x_0, of the zero of our polynomial as our divisor, we have that
$P(x_0) = (x_0 - x_0) Q(x_0) + b_0 = b_0$.

Using the product rule, we see that the derivative with respect to x of $P(x) = (x - x_0) Q(x) + b_0$ is
$P'(x) = Q(x)$.

Therefore, applying synthetic division to $Q(x)$ will give us $P'(x_0) = Q(x_0)$.
With that information, we are able to apply Newton's method to find the first approximation to a zero of $P(x)$:

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

If we begin the entire process again, applying synthetic division on $P(x)$ and $Q(x)$ using the approximation x_1 as our divisor,
we will find our next approximation, x_2, to the zero of $P(x)$.

The procedure described above is called ***deflation*** and unfortunately, inaccuracy results from the computation of the zeros of
successive polynomials.

Müllers method:

What do we do if the polynomial has complex zeros? Müller's, which is an extension of the Secant method, is useful in this case.
We want to consider the zeros of a quadratic polynomial

$$P(x) = a(x - p_2)^2 + b(x - p_2) + \mathrm{c}$$

that passes through three points

$$\left(p_0, f\left(x_0\right)\right), \left(p_1, f\left(x_1\right)\right) \text{ and } \left(p_3, f\left(x_3\right)\right).$$

That is, we have the following system of equations that can be solved for a, b, and c:

$f\left(p_0\right) = \mathrm{a}\left(p_0 - p_2\right)^2 + b\left(p_0 - p_2\right) + \mathrm{c}$:

$f\left(p_1\right) = \mathrm{a}\left(p_1 - p_2\right)^2 + b\left(p_1 - p_2\right) + \mathrm{c}$:

$f\left(p_2\right) = \mathrm{a}\left(p_2 - p_2\right)^2 + b\left(p_2 - p_2\right) + \mathrm{c} = \mathrm{c}$:

to obtain:

$$a = \frac{\left(p_0 - p_2\right)^2\left[f\left(p_1\right) - f\left(p_2\right)\right] - \left(p_1 - p_2\right)^2\left[f\left(p_0\right) - f\left(p_2\right)\right]}{\left(p_0 - p_2\right)\left(p_1 - p_2\right)\left(p_0 - p_1\right)} :$$

$$b = \frac{\left(p_1 - p_2\right)\left[f\left(p_0\right) - f\left(p_2\right)\right] - \left(p_0 - p_2\right)^2\left[f\left(p_1\right) - f\left(p_2\right)\right]}{\left(p_0 - p_2\right)\left(p_1 - p_2\right)\left(p_0 - p_1\right)} :$$

$$c = f\left(p_2\right) :$$

Substituting the formulas for a, b, and c into $P(x)$ and using the quadratic formula to solve $P(x) = 0$ for p_3 yields:

$$p_3 = p_2 - \frac{2\,c}{b + sgn(b)\sqrt{b^2 - 4\,ac}} :$$

where the denominator is chosen as the largest in magnitude that will result in p_3 being selected as the closest zero of P to p_2. Table 2.12 of the text shows the values computed by the method for the function below:

$f := x \rightarrow x^4 - 3\,x^3 + x^2 + x + 1$

$$x \rightarrow x^4 - 3\,x^3 + x^2 + x + 1 \qquad \textbf{(6.1)}$$

where p_0, p_1, p_2 are given and $f\left(p_0\right), f\left(p_1\right), f\left(p_2\right)$ are computed as follows:

given values of p and computations of f(p) for i = 0,1,2

$p0 := 0.5; f(p0)$

$$0.5$$

$$1.4375 \qquad \textbf{(6.1.1)}$$

$p1 := -0.5; f(p1)$

$$-0.5$$

$$1.1875 \qquad \textbf{(6.1.2)}$$

$p2 := 0; f(p2)$

$$0$$

$$1 \tag{6.1.3}$$

So that the student can better appreciate the steps within the algorithm, the computations for the first row of the table as defined in the algorithm are shown below:

computations for first pass of algorithm

$h1 := p1 - p0$

$$-1.0 \tag{6.2.1}$$

$h2 := p2 - p1$

$$0.5 \tag{6.2.2}$$

$del1 := \dfrac{f(p1) - f(p0)}{h1}$

$$0.2500000000 \tag{6.2.3}$$

$del2 := \dfrac{f(p2) - f(p1)}{h2}$

$$-0.3750000000 \tag{6.2.4}$$

$dd := \dfrac{del2 - del1}{h2 + h1}$

$$1.250000000 \tag{6.2.5}$$

$b := del2 + h2 \cdot dd$

$$0.2500000000 \tag{6.2.6}$$

$CapD := \left(b^2 - 4 \cdot f(p2) \cdot dd\right)^{\frac{1}{2}}$

$$2.222048604\,\mathrm{I} \tag{6.2.7}$$

if $|b - CapD| < |b + CapD|$ **then** $E := b + CapD$ **else** $E := b - CapD$ **end if**

$$0.2500000000 - 2.222048604\,\mathrm{I} \tag{6.2.8}$$

$h := -\dfrac{2 \cdot f(p2)}{E}$

$$-0.1000000000 - 0.8888194418\,\mathrm{I} \tag{6.2.9}$$

The compution below, then, is the first approximation, p_3, to the zero that appears in the first row of the table.

$p = p2 + h$

$$p = -0.1000000000 - 0.8888194418\,\mathrm{I} \tag{6.2.10}$$

The values are then reset and the process repeats to obtain the next row of the table.