2 Methods for Solving Nonlinear Problems

2.1 Preliminary Discussion

In this chapter we will learn methods for approximating solutions of nonlinear algebraic equations. We will limit our attention to the case of finding roots of a single equation of one variable. Thus, given a function, $f(x)$, we will be be interested in finding points x^*, for which $f(x^*) = 0$. A classical example that we are all familiar with is the case in which $f(x)$ is a quadratic equation. If, $f(x) = ax^2 + bx + c$, it is well known that the roots of $f(x)$ are given by

$$x^*_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

These roots may be complex or repeat (if the discriminant vanishes). This is a simple case in which the can be computed using a closed analytic formula. There exist formulas for finding roots of polynomials of degree 3 and 4, but these are rather complex. In more general cases, when f(x) is a polynomial of degree that is $\geqslant 5$, formulas for the roots no longer exist. Of course, there is no reason to limit ourselves to study polynomials, and in most cases, when $f(x)$ is an arbitrary function, there are no analytic tools for calculating the desired roots. Instead, we must use approximation methods. In fact, even in cases in which exact formulas are available (such as with polynomials of degree 3 or 4) an exact formula might be too complex to be used in practice, and approximation methods may quickly provide an accurate solution.

An equation $f(x) = 0$ may or may not have solutions. We are not going to focus on finding methods to decide whether an equation has a solutions or not, but we will look for approximation methods assuming that solutions actually exist. We will also assume that we are looking only for real roots. There are extensions of some of the methods that we will describe to the case of complex roots but we will not deal with this case. Even with the simple example of the quadratic equation, it is clear that a nonlinear equation $f(x) = 0$ may have more than one root. We will not develop any general methods for calculating the number of the roots. This issue will have to be dealt with on a case by case basis. We will also not deal with general methods for finding *a*ll the solutions of a given equation. Rather, we will focus on approximating *o*ne of the solutions.

The methods that we will describe, all belong to the category of *i*terative methods. Such methods will typically start with an initial guess of the root (or of the neighborhood of the root) and will gradually attempt to approach the root. In some cases, the sequence of iterations will converge to a limit, in which case we will then ask if the limit point is actually a solution of the equation. If this is indeed the case, another question of interest is how fast does the method converge to the solution? To be more precise, this question can be formulated in the following way: how many iterations of the method are required to guarantee a certain accuracy in the approximation of the solution of the equation.

2.1.1 Are there any roots anywhere?

There really are not that many general tools to knowing up front whether the root-finding problem can be solved. For our purposes, there most important issue will be to obtain some information about whether a root exists or not, and if a root does exist, then it will be important to make an attempt to estimate an interval to which such a solution belongs. One of our first attempts in solving such a problem may be to try to plot the function. After all, if the goal is to solve $f(x) = 0$, and the function $f(x)$ can be plotted in a way that the intersection of $f(x)$ with the x-axis is visible, then we should have a rather good idea as of where to look for for the root. There is absolutely nothing wrong with such a method, but it is not always easy to plot the function. There are many cases, in which it is rather easy to miss the root, and the situation always gets worse when moving to higher dimensions (i.e., more equations that should simultaneously be solved). Instead, something that is sometimes easier, is to verify that the function $f(x)$ is continuous (which hopefully it is) in which case all that we need is to find a point a in which $f(a) > 0$, and a point b, in which $f(b) < 0$. The continuity will then guarantee (due to the intermediate value theorem) that there exists a point c between a and b for which $f(c) = 0$, and the hunt for that point can then begin. How to find such points a and b? Again, there really is no general recipe. A combination of intuition, common sense, graphics, thinking, and trial-and-error is typically helpful. We would now like to consider several examples:

Example 2.1
A standard way of attempting to determine if a continuous function has a root in an interval is to try to find a point in which it is positive, and a second point in which it is negative. The intermediate value theorem for continuous functions then guarantees the existence of at least one point for which the function vanishes. To demonstrate this method, consider $f(x) = \sin(x) - x + 0.5$. At $x = 0$, $f(0) = 0.5 > 0$, while at $x = 5$, clearly $f(x)$ must be negative. Hence the intermediate value theorem guarantees the existence of at least one point $x^* \in (0, 5)$ for which $f(x^*) = 0$.

Example 2.2
Consider the problem $e^{-x} = x$, for which we are being asked to determine if a solution exists. One possible way to approach this problem is to define a function $f(x) = e^{-x} - x$, rewrite the problem as $f(x) = 0$, and plot $f(x)$. This is not so bad, but already requires a graphic calculator or a calculus-like analysis of the function $f(x)$ in order to plot it. Instead, it is a reasonable idea to start with the original problem, and plot both functions e^{-x} and x. Clearly, these functions intersect each other, and the intersection is the desirable root. Now, we can return to $f(x)$ and use its continuity (as a difference between continuous functions) to check its sign at a couple of points. For example, at $x = 0$, we have that $f(0) = 1 > 0$, while at $x = 1$, $f(1) = 1/e - 1 < 0$. Hence, due to the intermediate value theorem, there must exist a point x^* in the interval $(0, 1)$ for which $f(x^*) = 0$. At that point x^* we have $e^{-x^*} = x^*$. Note that while the graphical

argument clearly indicates that there exists one and only one solution for the equation, the argument that is based on the intermediate value theorem provides the existence of *at* least one solution.

A tool that is related to the intermediate value theorem is Brouwer's fixed point theorem:

Theorem 2.3 (Brouwer's Fixed Point Theorem) *Assume that $g(x)$ is continuous on the closed interval $[a, b]$. Assume that the interval $[a, b]$ is mapped to itself by $g(x)$, i.e., for any $x \in [a, b]$, $g(x) \in [a, b]$. Then there exists a point $c \in [a, b]$ such that $g(c) = c$. The point c is a fixed point of $g(x)$.*

The theorem is demonstrated in Figure 2.1. Since the interval $[a, b]$ is mapped to itself, the continuity of $g(x)$ implies that it must intersect the line x in the interval $[a, b]$ at least once. Such intersection points are the desirable fixed points of the function $g(x)$, as guaranteed by Theorem 2.3.

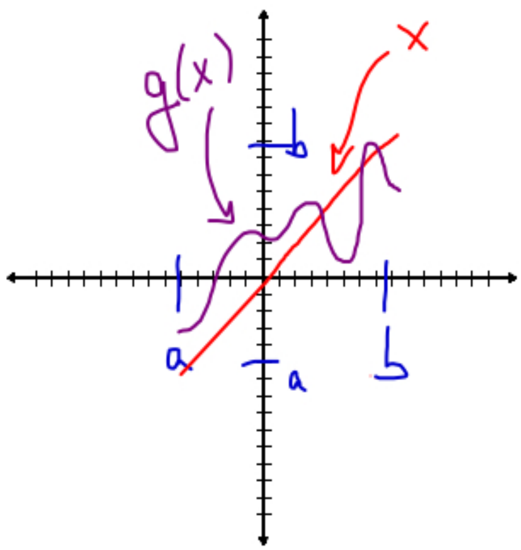

Figure 2.1: An illustration of the Brouwer fixed point theorem

Proof. Let $f(x) = x - g(x)$. Since $g(a) \in [a, b]$ and also $g(b) \in [a, b]$, we know that $f(a) = a - g(a) \leqslant 0$ while $f(b) = b - g(b) \geqslant 0$. Since $g(x)$ is continuous in $[a, b]$, so is $f(x)$, and hence according to the intermediate value theorem, there must exist a point $c \in [a, b]$ at which $f(c) = 0$. At this point $g(c) = c$. ■

How much does Theorem 2.3 add in terms of tools for proving that a root exists in a certain interval? In practice, the actual contribution is rather marginal, but there are cases where it adds something. Clearly if we are looking for roots of a function $f(x)$, we can always reformulate the problem as a fixed point problem for a function

$g(x)$ by defining $g(x) = f(x) + x$. Usually this is not the only way in which a root finding problem can be converted into a fixed point problem. In order to be able to use Theorem 2.3, the key point is always to look for a fixed point problem in which the interval of interest is mapped to itself.

Example 2.4
To demonstrate how the fixed point theorem can be used, consider the function $f(x) = e^x - x^2 - 3$ for $x \in [1, 2]$. Define $g(x) = \ln(x^2 + 3)$. Fixed points of $g(x)$ is a root of $f(x)$. Clearly, $g(1) = \ln 4 > \ln e = 1$ and $g(2) = \ln(7) < \ln(e^2) = 2$, and since $g(x)$ is continuous and monotone in $[1, 2]$, we have that $g([1, 2]) \subset [1, 2]$. Hence the conditions of Theorem 2.3 are satisfied and $f(x)$ must have a root in the interval $[1, 2]$.

2.1.2 Examples of root-finding methods

So far our focus has been on attempting to figure out if a given function has any roots, and if it does have roots, approximately where can they be. However, we have not went into any details in developing methods for approximating the values of such roots. Before we start with a detailed study of such methods, we would like to go over a couple of the methods that will be studied later on, emphasizing that they are all iterative methods. The methods that we will briefly describe are Newton's method and the secant method. A more detailed study of these methods will be conducted in the following sections.

1. **Newton's method.** Newton's method for finding a root of a differentiable function $f(x)$ is given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2.1}$$

   We note that for the formula (2.1) to be well-defined, we must require that $f'(x_n) \neq 0$ for any x_n. To provide us with a list of successive approximation, Newton's method (2.1) should be supplemented with one initial guess, say x_0. The equation (2.1) will then provide the values of $x_1, x_2, \ldots$

   One way of obtaining Newton's method is the following: Given a point x_n we are looking for the next point x_{n+1}. A linear approximation of $f(x)$ at x_{n+1} is

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

   Since x_{n+1} should be an approximation to the root of $f(x)$, we set $f(x_{n+1}) = 0$, rearrange the terms and get (2.1).

2. **The secant method.** The secant method is obtained by replacing the derivative in Newton's method, $f'(x_n)$, by the following finite difference approximation:

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}. \tag{2.2}$$

The secant method is thus:

$$x_{n+1} - x_n - f(x_n)\left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right]. \tag{2.3}$$

The secant method (2.3) should be supplemented by two initial values, say x_0, and x_1. Using these two values, (2.3) will provide the values of $x_2, x_3, \ldots$.

2.2 Iterative Methods

At this point we would like to explore more tools for studying iterative methods. We start by considering simple iterates, in which given an initial value x_0, the iterates are given by the following recursion:

$$x_{n+1} = g(x_n), \qquad n = 0, 1, \ldots \tag{2.4}$$

If the sequence $\{x_n\}$ in (2.4) converges, and if the function $g(x)$ is continuous, the limit must be a fixed point of the function $g(x)$. This is obvious, since if $x_n \to x^*$ as $n \to \infty$, then the continuity of $g(x)$ implies that in the limit we have

$$x^* = g(x^*).$$

Since things seem to work well when the sequence $\{x_n\}$ converges, we are now interested in studying exactly how can the convergence of this sequence be guaranteed? Intuitively, we expect that a convergence of the sequence will occur if the function $g(x)$ is "shrinking" the distance between any two points in a given interval. Formally, such a concept is known as "contraction" and is given by the following definition:

Definition 2.5 Assume that $g(x)$ is a continuous function in $[a, b]$. Then $g(x)$ is a contraction on $[a, b]$ if there exists a constant L such that $0 < L < 1$ for which for any x and y in $[a, b]$:

$$|g(x) - g(y)| \leqslant L|x - y|. \tag{2.5}$$

The equation (2.5) is referred to as a *Lipschitz condition* and the constant L is the *Lipschitz constant.*

Indeed, if the function $g(x)$ is a contraction, i.e., if it satisfies the Lipschitz condition (2.5), we can expect the iterates (2.4) to converge as given by the Contraction Mapping Theorem.

Theorem 2.6 (Contraction Mapping) *Assume that $g(x)$ is a continuous function on $[a, b]$. Assume that $g(x)$ satisfies the Lipschitz condition (2.5), and that $g([a, b]) \subset [a, b]$. Then $g(x)$ has a unique fixed point $c \in [a, b]$. Also, the sequence $\{x_n\}$ defined in (2.4) converges to c as $n \to \infty$ for any $x_0 \in [a, b]$.*

Proof. We know that the function $g(x)$ must have at least one fixed point due to Theorem 2.3. To prove the uniqueness of the fixed point, we assume that there are two fixed points c_1 and c_2. We will prove that these two points must be identical.

$$|c_1 - c_2| = |g(c_1) - g(c_2)| \leqslant L|c_1 - c_2|,$$

and since $0 < L < 1$, c_1 must be equal to c_2.
Finally, we prove that the iterates in (2.4) converge to c for any $x_0 \in [a, b]$.

$$|x_{n+1} - c| = |g(x_n) - g(c)| \leqslant L|x_n - c| \leqslant \ldots \leq L^{n+1}|x_0 - c|. \tag{2.6}$$

Since $0 < L < 1$, we have that as $n \to \infty$, $|x_{n+1} - c| \to 0$, and we have convergence of the iterates to the fixed point of $g(x)$ independently of the starting point x_0. ■

Remarks.

1. In order to use the Contraction Mapping Theorem, we must verify that the function $g(x)$ satisfies the Lipschitz condition, but what does it mean? The Lipschitz condition provides information about the "slope" of the function. The quotation marks are being used here, because we never required that the function $g(x)$ is differentiable. Our only requirement had to do with the continuity of $g(x)$. The Lipschitz condition can be rewritten as:

$$\frac{|g(x) - g(y)|}{|x - y|} \leqslant L, \qquad \forall x, y \in [a, b], \qquad x \neq y,$$

with $0 < L < 1$. The term on the LHS is a discrete approximation to the slope of $g(x)$. In fact, if the function $g(x)$ is differentiable, according to the Mean Value Theorem, there exists a point ξ between x and y such that

$$g'(\xi) = \frac{g(x) - g(y)}{x - y}.$$

Hence, in practice, if the function $g(x)$ is differentiable in the interval (a, b), and if there exists $L \in (0, 1)$, such that $|g'(x)| < L$ for any $x \in (a, b)$, then the assumptions on $g(x)$ satisfying the Lipshitz condition in Theorem 2.6 hold. Having $g(x)$ differentiable is more than the theorem requires but in many practical cases, we anyhow deal with differentiable g's so it is straightforward to use the condition that involves the derivative.

2. Another typical thing that can happen is that the function $g(x)$ will be differentiable, and $|g'(x)|$ will be less than 1, but only in a neighborhood of the fixed point. In this case, we can still formulate a "local" version of the contraction mapping theorem. This theorem will guarantee convergence to a fixed point, c, of $g(x)$ if we start the iterations sufficiently close to that point c.

Starting "far" from c may or may not lead to a convergence to c. Also, since we consider only a neighborhood of the fixed point c, we can no longer guarantee the uniqueness of the fixed point, as away from there, we do not post any restriction on the slope of $g(x)$ and therefore anything can happen.

3. When the contraction mapping theorem holds, and convergence of the iterates to the unique fixed point follows, it is of interest to know how many iterations are required in order to approximate the fixed point with a given accuracy. If our goal is to approximate c within a distance ε, then this means that we are looking for n such that

$$|x_n - c| \leqslant \varepsilon.$$

We know from (2.6) that

$$|x_n - c| \leqslant L^n |x_0 - c|, \qquad n \geqslant 1. \tag{2.7}$$

In order to get rid of c from the RHS of (2.7), we compute

$$|x_0 - c| = |x_c - x_1 + x_1 - c| \leqslant |x_0 - x_1| + |x_1 - c| \leqslant L|x_0 - c| + |x_1 - x_0|.$$

Hence

$$|x_0 - c| \leqslant \frac{|x_1 - x_0|}{1 - L}.$$

We thus have

$$|x_n - c| \leqslant \frac{L^n}{1 - L}|x_1 - x_0|,$$

and for $|x_n - c| < \varepsilon$ we require that

$$L^n \leqslant \frac{\varepsilon(1 - L)}{|x_1 - x_0|},$$

which implies that the number of iterations that will guarantee that the approximation error will be under ε must exceed

$$n \geqslant \frac{1}{\ln(L)} \cdot \ln\left[\frac{(1 - L)\varepsilon}{|x_1 - x_0|}\right]. \tag{2.8}$$

2.3 The Bisection Method

Before returning to Newton's method, we would like to present and study a method for finding roots which is one of the most intuitive methods one can easily come up with. The method we will consider is known as the "bisection method" .

We are looking for a root of a function $f(x)$ which we assume is continuous on the interval $[a, b]$. We also assume that it has opposite signs at both edges of the interval, i.e., $f(a)f(b) < 0$. We then know that $f(x)$ has at least one zero in $[a, b]$. Of course $f(x)$ may have more than one zero in the interval. The bisection method is only going to converge to one of the zeros of $f(x)$. There will also be no indication as of how many zeros $f(x)$ has in the interval, and no hints regarding where can we actually hope to find more roots, if indeed there are additional roots.

The first step is to divide the interval into two equal subintervals,

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

This generates two subintervals, $[a, c]$ and $[c, b]$, of equal lengths. We want to keep the subinterval that is guaranteed to contain a root. Of course, in the rare event where $f(c) = 0$ we are done. Otherwise, we check if $f(a)f(c) < 0$. If yes, we keep the left subinterval $[a, c]$. If $f(a)f(c) > 0$, we keep the right subinterval $[c, b]$. This procedure repeats until the stopping criterion is satisfied: we fix a small parameter $\varepsilon > 0$ and stop when $|f(c)| < \varepsilon$. To simplify the notation, we denote the successive intervals by $[a_0, b_0]$, $[a_1, b_1]$,... The first two iterations in the bisection method are shown in Figure 2.2. Note that in the case that is shown in the figure, the function $f(x)$ has multiple roots but the method converges to only one of them.

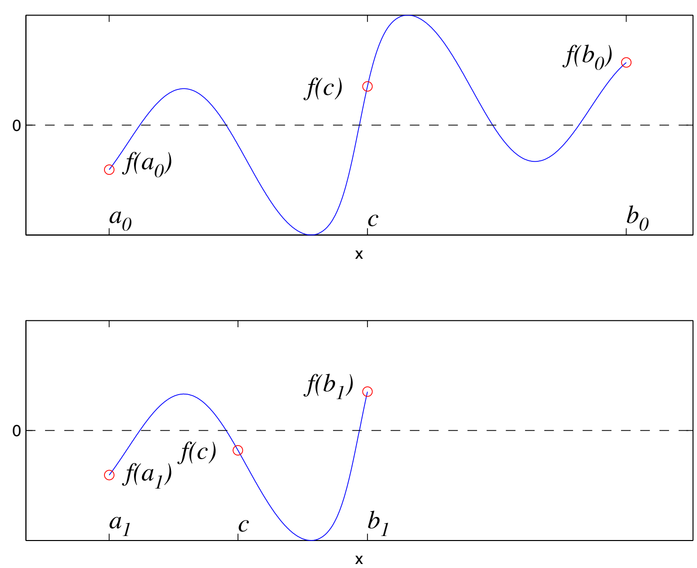

Figure 2.2: The first two iterations in a bisection root-finding method

We would now like to understand if the bisection method always converges to a root. We would also like to figure out how close we are to a root after iterating the algorithm several times. We first note that

$$a_0 \leqslant a_1 \leqslant a_2 \leqslant \ldots \leqslant b_0,$$

and

$$b_0 \geqslant b_1 \geqslant b_2 \geqslant \ldots \geqslant a_0.$$

We also know that every iteration shrinks the length of the interval by a half, i.e.,

$$b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n), \qquad n \geqslant 0,$$

which means that

$$b_n - a_n = 2^{-n}(b_0 - a_0).$$

The sequences $\{a_n\}_{n\geqslant 0}$ and $\{b_n\}_{n\geqslant 0}$ are monotone and bounded, and hence converge. Also

$$\lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n = \lim_{n\to\infty} 2^{-n}(b_0 - a_0) = 0,$$

so that both sequences converge to the same value. We denote that value by r, i.e.,

$$r = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n.$$

Since $f(a_n)f(b_n) \leqslant 0$, we know that $(f(r))^2 \leqslant 0$, which means that $f(r) = 0$, i.e., r is a root of $f(x)$.

We now assume that we stop in the interval $[a_n, b_n]$. This means that $r \in [a_n, b_n]$. Given such an interval, if we have to guess where is the root (which we know is in the interval), it is easy to see that the best estimate for the location of the root is the center of the interval, i.e.,

$$c_n = \frac{a_n + b_n}{2}.$$

In this case, we have

$$|r - c_n| \leqslant \frac{1}{2}(b_n - a_n) = 2^{-(n+1)}(b_0 - a_0).$$

We summarize this result with the following theorem.

Theorem 2.7 *If $[a_n, b_n]$ is the interval that is obtained in the n^{th} iteration of the bisection method, then the limits $\lim_{n\to\infty} a_n$ and $\lim_{n\to\infty} b_n$ exist, and*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = r,$$

where $f(r) = 0$. In addition, if

$$c_n = \frac{a_n + b_n}{2},$$

then

$$|r - c_n| \leqslant 2^{-(n+1)}(b_0 - a_0).$$

2.4 Newton's Method

Newton's method is a relatively simple, practical, and widely-used root finding method. It is easy to see that while in some cases the method rapidly converges to a root of the function, in some other cases it may fail to converge at all. This is one reason as of why it is so important not only to understand the construction of the method, but also to understand its limitations.

As always, we assume that $f(x)$ has at least one (real) root, and denote it by r. We start with an initial guess for the location of the root, say x_0. We then let $l(x)$ be the tangent line to $f(x)$ at x_0, i.e.,

$$l(x) - f(x_0) = f'(x_0)(x - x_0).$$

The intersection of $l(x)$ with the x-axis serves as the next estimate of the root. We denote this point by x_1 and write

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

which means that

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \tag{2.9}$$

In general, **the Newton method** (also known as the Newton-Raphson method) for finding a root is given by iterating (2.9) repeatedly, i.e.,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}. \tag{2.10}$$

Two sample iterations of the method are shown in Figure 2.3. Starting from a point x_n, we find the next approximation of the root x_{n+1}, from which we find x_{n+2} and so on. In this case, we do converge to the root of $f(x)$.

It is easy to see that Newton's method does not always converge. We demonstrate such a case in Figure 2.4. Here we consider the function $f(x) = \tan^{-1}(x)$ and show what happens if we start with a point which is a fixed point of Newton's method, iterated twice. In this case, $x_0 \approx 1.3917$ is such a point.

In order to analyze the error in Newton's method we let the error in the n^{th} iteration be

$$e_n = x_n - r.$$

We assume that $f''(x)$ is continuous and that $f'(r) \neq 0$, i.e., that r is a simple root of $f(x)$. We will show that the method has a quadratic convergence rate, i.e.,

$$e_{n+1} \approx c e_n^2. \tag{2.11}$$

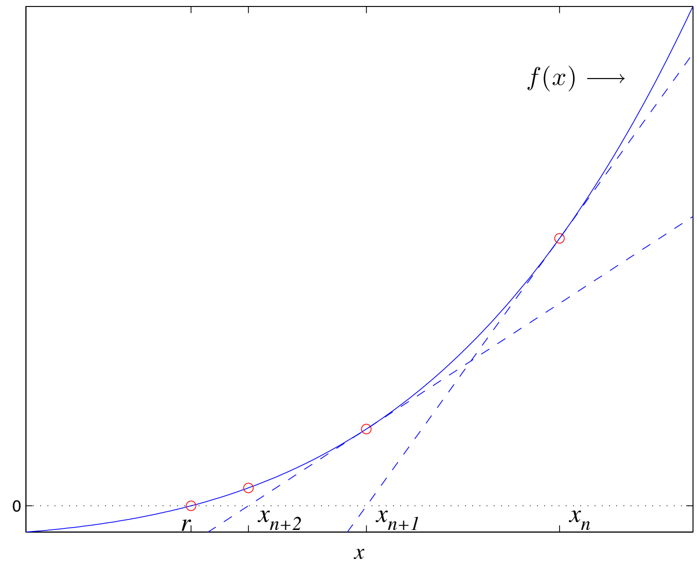

Figure 2.3: Two iterations in Newton's root-finding method. r is the root of $f(x)$ we approach by starting from x_n, computing x_{n+1}, then x_{n+2}, etc.

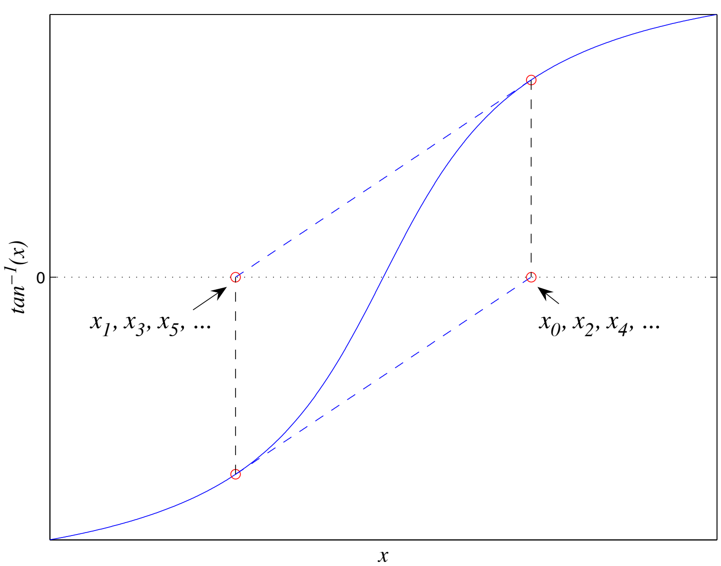

Figure 2.4: Newton's method does not always converge. In this case, the starting point is a fixed point of Newton's method iterated twice

A convergence rate estimate of the type (2.11) makes sense, of course, only if the method converges. Indeed, we will prove the convergence of the method for certain functions $f(x)$, but before we get to the convergence issue, let's derive the estimate (2.11). We rewrite e_{n+1} as

$$e_{n+1} = x_{n+1} - r = x_n - \frac{f(x_n)}{f'(x_n)} - r = e_n - \frac{f(x_n)}{f'(x_n)} = \frac{e_n f'(x_n) - f(x_n)}{f'(x_n)}.$$

Writing a Taylor expansion of $f(r)$ about $x = x_n$ we have

$$0 = f(r) = f(x_n - e_n) = f(x_n) - e_n f'(x_n) + \frac{1}{2}e_n^2 f''(\xi_n),$$

which means that

$$e_n f'(x_n) - f(x_n) = \frac{1}{2} f''(\xi_n) e_n^2.$$

Hence, the relation (2.11), $e_{n+1} \approx c e_n^2$, holds with

$$c = \frac{1}{2}\frac{f''(\xi_n)}{f'(x_n)} \tag{2.12}$$

Since we assume that the method converges, in the limit as $n \to \infty$ we can replace (2.12) by

$$c = \frac{1}{2}\frac{f''(r)}{f'(r)}. \tag{2.13}$$

We now return to the issue of convergence and prove that for certain functions Newton's method converges regardless of the starting point.

Theorem 2.8 *Assume that $f(x)$ has two continuous derivatives, is monotonically increasing, convex, and has a zero. Then the zero is unique and Newton's method will converge to it from every starting point.*

Proof. The assumptions on the function $f(x)$ imply that $\forall x$, $f''(x) > 0$ and $f'(x) > 0$. By (2.12), the error at the $(n+1)^{\text{th}}$ iteration, e_{n+1}, is given by

$$e_{n+1} = \frac{1}{2}\frac{f''(\xi_n)}{f'(x_n)} e_n^2,$$

and hence it is positive, i.e., $e_{n+1} > 0$. This implies that $\forall n \geqslant 1$, $x_n > r$, Since $f'(x) > 0$, we have

$$f(x_n) > f(r) = 0.$$

Now, subtracting r from both sides of (2.10) we may write

$$e_{n+1} = e_n - \frac{f(x_n)}{f'(x_n)}, \tag{2.14}$$

which means that $e_{n+1} < e_n$ (and hence $x_{n+1} < x_n$). Hence, both $\{e_n\}_{n\geqslant 0}$ and $\{x_n\}_{n\geqslant 0}$ are decreasing and bounded from below. This means that both series converge, i.e., there exists e^* such that,

$$e^* = \lim_{n\to\infty} e_n,$$

and there exists x^* such that

$$x^* = \lim_{n\to\infty} x_n.$$

By (2.14) we have

$$e^* = e^* - \frac{f(x^*)}{f'(x^*)},$$

so that $f(x^*) = 0$, and hence $x^* = r$. ■

Theorem 2.8 guarantees global convergence to the unique root of a monotonically increasing, convex smooth function. If we relax some of the requirements on the function, Newton's method may still converge. The price that we will have to pay is that the convergence theorem will no longer be global. Convergence to a root will happen only if we start sufficiently close to it. Such a result is formulated in the following theorem.

Theorem 2.9 *Assume $f(x)$ is a continuous function with a continuous second derivative, that is defined on an interval $I = [r-\delta, r+\delta]$, with $\delta > 0$. Assume that $f(r) = 0$, and that $f''(r) \neq 0$. Assume that there exists a constant A such that*

$$\frac{|f''(x)|}{|f'(y)|} \leqslant A, \qquad \forall x, y \in I.$$

If the initial guess x_0 is sufficiently close to the root r, i.e., if $|r - x_0| \leq \min\{\delta, 1/A\}$, then the sequence $\{x_n\}$ defined in (2.10) converges quadratically to the root r.

Proof. We assume that $x_n \in I$. Since $f(r) = 0$, a Taylor expansion of $f(x)$ at $x = x_n$, evaluated at $x = r$ is:

$$0 = f(r) = f(x_n) + (r - x_n)f'(x_n) + \frac{(r-x_n)^2}{2} f''(\xi_n), \tag{2.15}$$

where ξ_n is between r and x_n, and hence $\xi \in I$. Equation (2.15) implies that

$$r - x_n = \frac{-2f(x_n) - (r-x_n)^2 f''(\xi_n)}{2f'(x_n)}.$$

Since x_{n+1} are the Newton iterates and hence satisfy (2.10), we have

$$r - x_{n+1} = r - x_n + \frac{f(x_n)}{f'(x_n)} = -\frac{(r - x_n^2)f''(\xi_n)}{2f'(x_n)}. \tag{2.16}$$

Hence

$$|r - x_{n+1}| \leqslant \frac{(r - x_n)^2}{2} A \leqslant \frac{|r - x_n|}{2} \leqslant \ldots \leqslant 2^{-(n-1)}|r - x_0,|$$

which implies that $x_n \to r$ as $n \to \infty$.
It remains to show that the convergence rate of $\{x_n\}$ to r is quadratic. Since ξ_n is between the root r and x_n, it also converges to r as $n \to \infty$. The derivatives f' and f'' are continuous and therefore we can take the limit of (2.16) as $n \to \infty$ and write

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

which implies the quadratic convergence of $\{x_n\}$ to r. ■

2.5 The Secant Method

We recall that Newton's root finding method is given by equation (2.10), i.e.,

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

We now assume that we do not know that the function $f(x)$ is differentiable at x_n, and thus can not use Newton's method as is. Instead, we can replace the derivative $f'(x_n)$ that appears in Newton's method by a difference approximation. A particular choice of such an approximation,

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

leads to **the secant method** which is given by

$$x_{n+1} = x_n - f(x_n) \left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \right], \qquad n \geqslant 1. \tag{2.17}$$

A geometric interpretation of the secant method is shown in Figure 2.5. Given two points, $(x_{n-1}, f(x_{n-1}))$ and $(x_n, f(x_n))$, the line $l(x)$ that connects them satisfies

$$l(x) - f(x_n) = \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}(x - x_n).$$

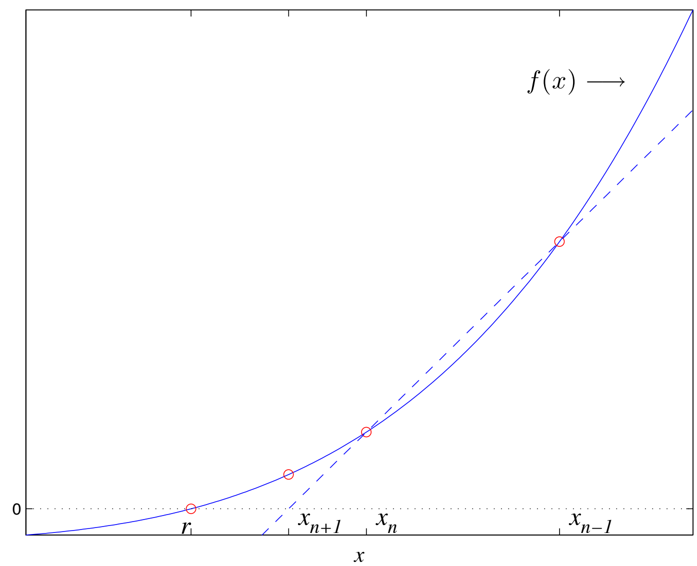

Figure 2.5: The Secant root-finding method. The points x_{n-1} and x_n are used to obtain x_{n+1}, which is the next approximation of the root r

The next approximation of the root, x_{n+1}, is defined as the intersection of $l(x)$ and the x-axis, i.e.,

$$0 - f(x_n) = \frac{f(x_{n-1}) - f(x_n)}{x_{n-1} - x_n}(x_{n+1} - x_n). \tag{2.18}$$

Rearranging the terms in (2.18) we end up with the secant method (2.17).

We note that the secant method (2.17) requires two initial points. While this is an extra requirement compared with, e.g., Newton's method, we note that in the secant method there is no need to evaluate any derivatives. In addition, if implemented properly, every stage requires only one new function evaluation.

We now proceed with an error analysis for the secant method. As usual, we denote the error at the n^{th} iteration by $e_n = x_n - r$. We claim that the rate of convergence of the secant method is **superlinear** (meaning, better than linear but less than quadratic). More precisely, we will show that it is given by

$$|e_{n+1}| \approx |e_n|^{\alpha}, \tag{2.19}$$

with

$$\alpha = \frac{1+\sqrt{5}}{2}. \tag{2.20}$$

We start by rewriting e_{n+1} as

$$e_{n+1} = x_{n+1} - r = \frac{f(x_n)x_{n-1} - f(x_{n-1})x_n}{f(x_n) - f(x_{n-1})} - r = \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})}.$$

Hence

$$e_{n+1} = e_n e_{n-1} \left[\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \right] \left[\frac{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}}{x_n - x_{n-1}} \right]. \tag{2.21}$$

A Taylor expansion of $f(x_n)$ about $x = r$ reads

$$f(x_n) = f(r + e_n) = f(r) + e_n f'(r) + \frac{1}{2} e_n^2 f''(r) + O(e_n^3),$$

and hence

$$\frac{f(x_n)}{e_n} = f'(r) + \frac{1}{2} e_n f''(r) + O(e_n^2).$$

We thus have

$$\begin{aligned} \frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}} &= \frac{1}{2}(e_n - e_{n-1}) f''(r) + O(e_{n-1}^2) + O(e_n^2) \\ &= \frac{1}{2}(x_n - x_{n-1}) f''(r) + O(e_{n-1}^2) + O(e_n^2). \end{aligned}$$

Therefore,

$$\frac{\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}}{x_n - x_{n-1}} \approx \frac{1}{2} f''(r),$$

and

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

The error expression (2.21) can be now simplified to

$$e_{n+1} \approx \frac{1}{2} \frac{f''(r)}{f'(r)} e_n e_{n-1} = c e_n e_{n-1}. \tag{2.22}$$

Equation (2.22) expresses the error at iteration $n+1$ in terms of the errors at iterations n and $n-1$. In order to turn this into a relation between the error at the $(n+1)^{\text{th}}$ iteration and the error at the n^{th} iteration, we now assume that the order of convergence is α, i.e.,

$$|e_{n+1}| \sim A|e_n|^{\alpha}. \tag{2.23}$$

Since (2.23) also means that $|e_n| \sim A|e_{n-1}|^\alpha$, we have

$$A|e_n|^\alpha \sim C|e_n|A^{-\frac{1}{\alpha}}|e_n|^{\frac{1}{\alpha}}.$$

This implies that

$$A^{1+\frac{1}{\alpha}}C^{-1} \sim |e_n|^{1-\alpha+\frac{1}{\alpha}}. \tag{2.24}$$

The left-hand-side of (2.24) is non-zero while the right-hand-side of (2.24) tends to zero as $n \to \infty$ (assuming, of course, that the method converges). This is possible only if

$$1 - \alpha + \frac{1}{\alpha} = 0,$$

which, in turn, means that

$$\alpha = \frac{1+\sqrt{5}}{2}.$$

The constant A in (2.23) is thus given by

$$A = C^{\frac{1}{1+\frac{1}{\alpha}}} = C^{\frac{1}{\alpha}} = C^{\alpha-1} = \left[\frac{f''(r)}{2f'(r)}\right]^{\alpha-1}.$$

We summarize this result with the theorem:

Theorem 2.10 *Assume that $f''(x)$ is continuous $\forall x$ in an interval I. Assume that $f(r) = 0$ and that $f'(r) \neq 0$. If x_0, x_1 are sufficiently close to the root r, then $x_n \to r$. In this case, the convergence is of order $\frac{1+\sqrt{5}}{2}$.*