0.1 Newton’s Method

An unfortunate truth of mathematics is that the majority of problems cannot be solved exactly - instead we need to use numerical approximations. Let’s consider the problem of solving an algebraic equation. For a second-order polynomial we can use the quadratic formula to find its roots, and for third and fourth order polynomials there are even more complicated methods to find the roots. However, for a fifth-order or higher polynomial there is no general method for finding roots. Right now we’re only considering polynomials. What about equations involving trigonometric functions? We previously considered the bisection method, which is one numerical method we can use to solve such equations. Newton’s method is another method for solving such equations, which can be very useful, because it can converge to a result much more quickly, or provide us with a better estimate after fewer iterations.

The basic idea of Newton’s method is of first-order approximation. As stated previously, the line tangent to a curve at a given point is the best first-order approximation of the curve at that point. This means that very close to the point of tangency the tangent line is an excellent approximation to a function, but as we move further away from the point of tangency the accuracy of the approximation decreases. The key here is that the function behaves similar to the tangent line. For this reason, it is reasonable to believe that a tangent line will cross the $x$-axis at point somewhere similar to where the function itself crosses the $x$-axis (this is useful because we can reforge any equation into a question of where a specific function crosses the $x$-axis, or is 0).

If this were as far as we could go Newton’s method would not be that useful. The trick comes in as follows. By finding the intercept of the tangent line we get a new $x$-coordinate, which is presumably closer to the 0 of our function than the original point we chose. Now we construct a tangent line at the new point and see where it crosses the $x$-axis. This is presumably once again closer to our real solution. By continuing this process we get a sequence of approximations to our solution, which hopefully converge to, or become increasingly closer to, the real solution. If we look in the limit as the number of iterations of this process tends to $\infty$, we should hopefully find the actual solution to the problem at hand. See figure 1 for an illustration of Newton’s method.

![Figure 1](image1.png)

Figure 1: The initial guess generally provides a poor estimate of the actual root, but by iterating the process we hope to move closer and closer to the actual root.

In practice, we apply Newton’s method in the following way. First, move all terms to one side of the equation in order to have an equation of the form $f(x) = 0$ (this transforms the problem of finding where two sides of an equation are equal to finding where a new function crosses the $x$-axis). Then, make an initial guess $x_0$ at a possible solution. Next, find the equation of the line
tangent to the function $f(x)$ at the point $x_0$. Solve for the zero of the tangent line (which is a simple first-order equation), and use the result as a second guess for the solution to $f(x) = 0$. Continue iterating this process until solutions either converge or diverge.

**Example 1.** Solve the equation $x^2 = 3$ using Newton’s method.

**Solution** We already know that $\pm \sqrt{3}$ are solutions to this equation, but let’s try and find them using Newton’s method. The first step is to formulate the equation as $f(x) = x^2 - 3$, and set it equal to 0. Let us make an initial guess that $x = 2$ is a solution to the equation (even though we know it is a little high). We find $f'(x) = 2x$, so the equation of the tangent line is given by

$$
\hat{f}_2(x) = f(2) + f'(2)(x - 2) = (4 - 3) + 4(x - 2) = 4x - 7.
$$

If we set $\hat{f}_2(x) = 4x - 7 = 0$ we can see by inspection that the solution is $x = 1.75$, which is closer than 2 to $\sqrt{3} \approx 1.73205$. If we continue this process we can find

$$
\hat{f}_{1.75}(x) = f(1.75) + f'(1.75)(x - 1.75) = (1.75^2 - 3) + 3.5(x - 1.75) = 3.5x - 6.0625.
$$

Now when we set $\hat{f}_{1.75}(x) = 3.5x - 6.0625 = 0$ we find $x = 6.0625/3.5 \approx 1.73214$. After only two iterations we have found a solution that is relatively close to $\sqrt{3}$.

Since Newton’s is an iterative process it is very useful to recast the process in a different form that simplifies calculations, and makes it much easier to implement Newton’s method using a computer. Let us suppose we are on the $n^{th}$ iteration of Newton’s method, and we have found an $x$ value of $x_n$. Setting the line tangent to $f(x)$ at $x_n$ to 0 we get

$$
0 = \hat{f}_{x_n} = f(x_n) + f'(x_n)(x - x_n),
$$

where $x$ is the solution to this equation that gives us our next value $x$. Solving this equation for $x$ we find

$$
\begin{align*}
  f(x_n) + f'(x_n)(x - x_n) &= 0 \\
  f'(x_n)(x - x_n) &= -f(x_n) \\
  x - x_n &= -\frac{f(x_n)}{f'(x_n)} \\
  x &= x_n - \frac{f(x_n)}{f'(x_n)}
\end{align*}
$$

This will be $x_{n+1}$, the value we use in the next iteration. In conclusion, we have that

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

The above equation holds as long as $f'(x_n) \neq 0$ (if $f'(x_n) = 0$ then we have a horizontal tangent line, which does not have any zeros, unless it is zero everywhere). This simplifies the process of finding a numerical solution as follows. First, we set the equation $f(x) = 0$, and make an initial guess at a solution. Next, use the above formula to find a second guess, and continue iterating until the answers converge to a value up to the desired accuracy. In this way we construct a sequence $\{x_n\}_n$ that hopefully converges to a solution.

**Example 2.** Use Newton’s method to find a nonzero solution to $x = \frac{\pi}{2} \sin(x)$.
**Solution** We can see by inspection that there are no possible solutions for $|x| > \pi/2$, because the sine function has an amplitude of $\pi/2$. Let us reframe the question to solving

$$f(x) = x - \frac{\pi}{2} \sin(x) = 0.$$ 

We will choose $x_0 = 1$ for an initial guess. Since

$$f'(x) = 1 - \frac{\pi}{2} \cos(x),$$

we find that

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{1 - (\pi/2) \sin(1)}{1 - (\pi/2) \cos(1)} \approx 3.12683$$

Iterating again we find

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 3.12683 - \frac{3.12683 - \frac{\pi}{2} \sin(3.12683)}{1 - \frac{\pi}{2} \cos(3.12683)} \approx 1.91948$$

Continuing these iterations, we find $x_3 = 1.63106$, $x_4 = 1.5734$, $x_5 = 1.5708$, and $x_6 = 1.570796$. At this point we can feel confident we’ve converged to a solution, which happens to be $x = \pi/2 \approx 1.570796$. We have found a solution that is accurate up to at least the first 6 decimal places (actually more) of the solution $\pi/2$.

For this particular equation, there are actually three solutions. Depending on where the initial point is chosen, Newton’s method may find a different root ($x_0 = -1$ yields $x = -\pi/2$), or may not converge at all. Despite being very effective, Newton’s method can be a bit of trial and error. Because of this, in practice one would use some method of determining a reasonable first guess, and then use Newton’s method for increased accuracy.

**Example 3.** Use Newton’s method to solve $e^x = x + 2$ with at least 3 decimal places of accuracy, beginning with $x_0 = 1$.

**Solution** First we set $f(x) = e^x - x - 2$ and differentiate to find $f'(x) = e^x - 1$. Now we iterate to find

$$x_1 = 1 - \frac{e^1 - 1 - 2}{e^1 - 1} \approx 1.16395$$

$$x_2 = 1.16395 - \frac{e^{1.16395} - 1.16395 - 2}{e^{1.16395} - 1} \approx 1.14642$$

$$x_3 = 1.14642 - \frac{e^{1.14642} - 1.14642 - 2}{e^{1.14642} - 1} \approx 1.14619.$$ 

In order to be confident we have achieved 3 decimal points of accuracy, we continue iterating until the first 3 decimal places do not change. To really be confident we might consider iterating one step further.