Newton's Method

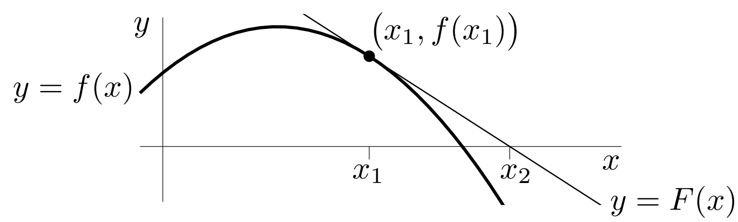
Newton's method is a technique for generating numerical approximate solutions to equations of the form $f(x) = 0$. For example, one can easily get a good approximation to $\sqrt{2}$ by applying Newton's method to the equation $x^2 - 2 = 0$. This will be done in Example 1, below.

Here is the derivation of Newton's method. We start by simply making a guess for the solution. For example we could base the guess on a sketch of the graph of $f(x)$. Call the initial guess x_1. Next find the linear (tangent line) approximation to $f(x)$ near x_1. Let's call the linear approximation $F(x)$. It is

$$F(x) = f(x_1) + f'(x_1)\,(x - x_1)$$

Now, instead of trying to solve $f(x) = 0$, we solve the linear equation $F(x) = 0$ and call the solution x_2.

$$0 = F(x) = f(x_1) + f'(x_1)\,(x - x_1) \iff x - x_1 = -\frac{f(x_1)}{f'(x_1)} \iff x = x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$



Now we repeat, but starting with the (second) guess x_2 rather than x_1. This gives the (third) guess $x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$. And so on. By way of summary, Newton's method is

1) Make a preliminary guess x_1.
2) Define $x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$.
3) Iterate. That is, for each natural number n, once you have computed x_n, define $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$.

Example 1 In this example we compute, approximately, the square root of two by applying Newton's method to the equation

$$f(x) = x^2 - 2 = 0$$

Since $f'(x) = 2x$, Newton's method says that we should generate approximate solutions by iteratively applying

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

Since $1^2 = 1 < 2$ and $2^2 = 4 > 2$, the square root of two must be between 1 and 2, so let's start Newton's method with the initial guess $x_1 = 1.5$ Here goes:

$$\begin{aligned}
x_1 &= 1.5 \\
x_2 &= \tfrac{1}{2}x_1 + \tfrac{1}{x_1} = \tfrac{1}{2}(1.5) + \tfrac{1}{1.5} \\
&= 1.416666667 \\
x_3 &= \tfrac{1}{2}x_2 + \tfrac{1}{x_2} = \tfrac{1}{2}(1.416666667) + \tfrac{1}{1.416666667} \\
&= 1.414215686 \\
x_4 &= \tfrac{1}{2}x_3 + \tfrac{1}{x_3} = \tfrac{1}{2}(1.414215686) + \tfrac{1}{1.414215686} \\
&= 1.414213562 \\
x_5 &= \tfrac{1}{2}x_4 + \tfrac{1}{x_4} = \tfrac{1}{2}(1.414213562) + \tfrac{1}{1.414213562} \\
&= 1.414213562
\end{aligned}$$

Since $f(1.4142135615) = -2.5 \times 10^{-9} < 0$ and $f(1.4142135625) = 3.6 \times 10^{-10} > 0$ the square root of two must be between 1.4142135615 and 1.4142135625.

Example 2 In this example we compute, approximately, π by applying Newton's method to the equation

$$f(x) = \sin x = 0$$

starting with $x_1 = 3$. Since $f'(x) = \cos x$, Newton's method says that we should generate approximate solutions by iteratively applying

$$x_{n+1} = x_n - \tfrac{f(x_n)}{f'(x_n)} = x_n - \tfrac{\sin x_n}{\cos x_n} = x_n - \tan x_n$$

Here goes

$$\begin{aligned}
x_1 &= 3 \\
x_2 &= x_1 - \tan x_1 = 3 - \tan 3 \\
&= 3.142546543 \\
x_3 &= 3.142546543 - \tan 3.142546543 \\
&= 3.141592653 \\
x_4 &= 3.141592653 - \tan 3.141592653 \\
&= 3.141592654 \\
x_5 &= 3.141592654 - \tan 3.141592654 \\
&= 3.141592654
\end{aligned}$$

Since $f(3.1415926535) = 9.0 \times 10^{-11} > 0$ and $f(3.1415926545) = -9.1 \times 10^{-11} < 0$ the square root of two must be between 3.1415926535 and 3.1415926545.

Example 3 This example illustrates how Newton's method can go badly wrong if your initial guess is not good enough. We'll try to solve the equation

$$f(x) = \tan^{-1} x = 0$$

starting with $x_1 = 1.5$. Of course the solution to this equation is just $x = 0$. Since $f'(x) = \frac{1}{1+x^2}$ Newton's method gives

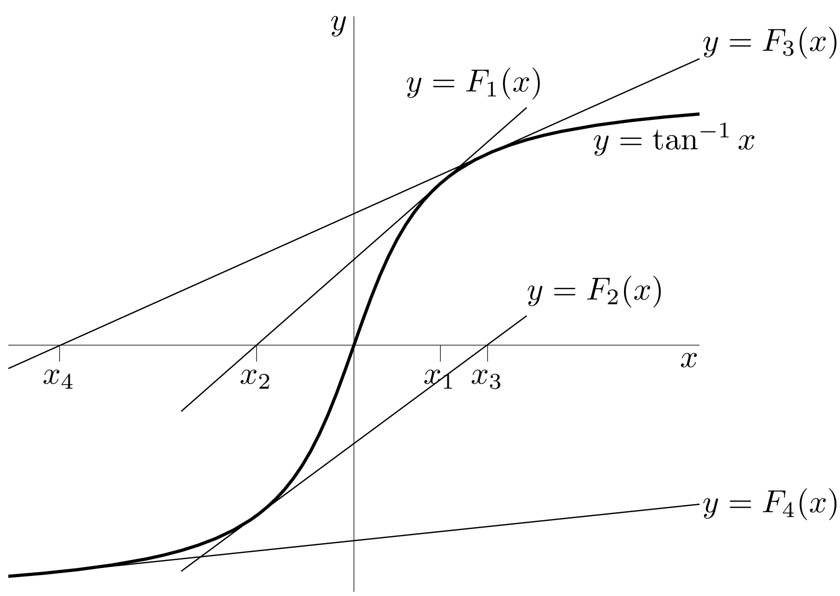
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - (1 + x_n^2)\tan^{-1} x_n$$

So

$$\begin{aligned}
x_1 &= 1.5\\
x_2 &= 1.5 - (1 + 1.5^2)\tan^{-1} 1.5 = -1.69\\
x_3 &= -1.69 - (1 + 1.69^2)\tan^{-1}(-1.69) = 2.32\\
x_4 &= 2.32 - (1 + 2.32^2)\tan^{-1}(2.32) = -5.11\\
x_5 &= -5.11 - (1 + 5.11^2)\tan^{-1}(-5.11) = 32.3\\
x_6 &= 32.3 - (1 + 32.3^2)\tan^{-1}(32.3) = -1575\\
x_7 &= 3,894,976
\end{aligned}$$

Here is a figure which shows what went wrong. In this figure, $y = F_1(x)$ is the tangent line to $y = \tan^{-1} x$ at $x = x_1$. Under Newton's method, this tangent line crosses the x–axis at $x = x_2$. Then $y = F_2(x)$ is the tangent to $y = \tan^{-1} x$ at $x = x_2$. Under Newton's method, this tangent line crosses the x–axis at $x = x_3$. And so on.

The problem arose because the x_n's, and especially x_1, were far enough from the solution $x = 0$, that the tangent line approximations, while good approximations to $f(x)$



for $x \approx x_n$, were very poor approximations to $f(x)$ for $x \approx 0$. If we had started with $x_1 = 0.5$ instead of $x_1 = 1.5$, Newton's method would not have failed:

$$x_1 = 0.5 \qquad x_2 = -0.0796 \qquad x_3 = 0.000335 \qquad x_4 = -2.51 \times 10^{-11}$$

Error Behaviour of Newton's Method

Newton's method usually works spectacularly well, provided your initial guess is reasonably close to a solution of $f(x) = 0$. A good way to select this initial guess is to sketch the graph of $y = f(x)$. We now see why "Newton's method usually works spectacularly well, provided your initial guess is reasonably close to a solution of $f(x) = 0$".

Let r be any solution of $f(x) = 0$. Then $f(r) = 0$. Suppose that we have already computed x_n. The error in x_n is $\big|x_n - r\big|$. We now derive a formula that relates the error after the next step, $\big|x_{n+1} - r\big|$, to $\big|x_n - r\big|$. We have seen in class that

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \tfrac{1}{2} f'(c)(x - x_n)^2$$

for some c between x_n and x. In particular, choosing $x = r$,

$$0 = f(r) = f(x_n) + f'(x_n)(r - x_n) + \tfrac{1}{2} f'(c)(r - x_n)^2 \tag{1}$$

By the definition of x_{n+1},

$$0 = f(x_n) + f'(x_n)(x_{n+1} - x_n) \tag{2}$$

(In fact, we defined x_{n+1} as the solution of $0 = f(x_n) + f'(x_n)(x - x_n)$.) Subtracting (2) from (1).

$$\begin{aligned} 0 = f'(x_n)(r - x_{n+1}) + \tfrac{1}{2} f''(c)(r - x_n)^2 \ &\Rightarrow\ x_{n+1} - r = \tfrac{f''(c)}{2f'(x_n)}(x_n - r)^2 \\ &\Rightarrow\ \big|x_{n+1} - r\big| = \tfrac{|f''(c)|}{2|f'(x_n)|}|x_n - r|^2 \end{aligned}$$

If the guess x_n is close to r, then c, which must be between x_n and r, is also close to r and $\big|x_{n+1} - r\big| \approx \frac{|f''(r)|}{2|f'(r)|}|x_n - r|^2$. Even when x_n is not close to r, if we know that there are two numbers $L, M > 0$ such that f obeys:

H1) $\big|f'(x_n)\big| \ge L$

H2) $\big|f''(c)\big| \le M$

(we'll see examples of this below) then we will have

$$\big|x_{n+1} - r\big| \le \tfrac{M}{2L}|x_n - r|^2 \tag{3}$$

Let's denote by ε_1 the error $|x_1 - r|$ of our initial guess. In fact, let's denote by ε_n the error $|x_n - r|$ in x_n. Then (3) says

$$\varepsilon_{n+1} \le \tfrac{M}{2L}\varepsilon_n^2$$

In particular

$$\begin{aligned}
\varepsilon_2 &\le \tfrac{M}{2L}\varepsilon_1^2 \\
\varepsilon_3 &\le \tfrac{M}{2L}\varepsilon_2^2 \le \tfrac{M}{2L}\big(\tfrac{M}{2L}\varepsilon_1^2\big)^2 &&= \big(\tfrac{M}{2L}\big)^3\varepsilon_1^4 \\
\varepsilon_4 &\le \tfrac{M}{2L}\varepsilon_3^2 \le \tfrac{M}{2L}\Big(\big(\tfrac{M}{2L}\big)^3\varepsilon_1^4\Big)^2 &&= \big(\tfrac{M}{2L}\big)^7\varepsilon_1^8 \\
\varepsilon_5 &\le \tfrac{M}{2L}\varepsilon_4^2 \le \tfrac{M}{2L}\Big(\big(\tfrac{M}{2L}\big)^7\varepsilon_1^8\Big)^2 &&= \big(\tfrac{M}{2L}\big)^{15}\varepsilon_1^{16}
\end{aligned}$$

By now we can see a pattern forming, that is easily verified by induction

$$\varepsilon_n \le \big(\tfrac{M}{2L}\big)^{2^{n-1}-1}\varepsilon_1^{2^{n-1}} = \tfrac{2L}{M}\Big(\tfrac{M}{2L}\varepsilon_1\Big)^{2^{n-1}} \tag{4}$$

As long as $\frac{M}{2L}\varepsilon_1 < 1$ (which tells us quantitatively how good our first guess has to be in order for Newton's method to converge), this goes to zero extremely quickly as n increases. For example, suppose that $\frac{M}{2L}\varepsilon_1 \le \frac{1}{2}$. Then

$$\varepsilon_n \le \tfrac{2L}{M}\big(\tfrac{1}{2}\big)^{2^{n-1}} \le \tfrac{2L}{M}\begin{cases} 0.25 & \text{if } n=2 \\ 0.0625 & \text{if } n=3 \\ 0.0039 = 3.9\times 10^{-3} & \text{if } n=4 \\ 0.000015 = 1.5\times 10^{-5} & \text{if } n=5 \\ 0.00000000023 = 2.3\times 10^{-10} & \text{if } n=6 \\ 0.000000000000000000054 = 5.4\times 10^{-20} & \text{if } n=7 \end{cases}$$

Each time you increase n by one, the number of zeroes after the decimal place roughly doubles.

Example 1 (continued) Let's consider, as we did in Example 1, $f(x) = x^2 - 2$, starting with $x_1 = \frac{3}{2}$. Then

$$f'(x) = 2x \qquad f''(x) = 2$$

So we may certainly take $M = 2$ and if, for example, $x_n \ge 1$ for all n (as happened in Example 1), we may take $L = 2$ too. While we do not know what r is, we do know that $1 \le r \le 2$ (since $f(1) = 1^1 - 2 < 0$ and $f(2) = 2^2 - 2 > 0$). As we took $x_1 = \frac{3}{2}$, we have $\varepsilon_1 = |x_1 - r| \le \frac{1}{2}$, so that $\frac{M}{2L}\varepsilon_1 \le \frac{1}{4}$ and

$$\varepsilon_{n+1} \le \tfrac{2L}{M}\Big(\tfrac{M}{2L}\varepsilon_1\Big)^{2^{n-1}} \le 2\big(\tfrac{1}{4}\big)^{2^{n-1}}$$

Example 2 (continued) Let's consider, as we did in Example 2, $f(x) = \sin x$, starting with $x_1 = 3$. Then

$$f'(x) = \cos x \qquad f''(x) = -\sin x$$

As $|-\sin x| \le 1$, we may certainly take $M = 1$.

In Example 2, all x_n's were between 3 and 3.2. Since (to three decimal places)

$$\sin(3) = 0.141 > 0 \qquad \sin(3.2) = -0.058 < 0$$

we necessarily have $3 < r < 3.2$ and $\varepsilon_1 = |x_1 - r| < 0.2$.

So r and all x_n's and hence all c's lie in the interval $(3, 3.2)$. Since

$$\cos(3) < -0.9 \qquad \cos(3.2) < -0.9$$

we necessarily have $\big|f''(c)\big| = \big|-\cos c\big| \ge 0.9$ and we may take $L = 0.9$. So

$$\varepsilon_{n+1} \le \tfrac{2L}{M}\Big(\tfrac{M}{2L}\varepsilon_1\Big)^{2^{n-1}} \le \tfrac{2\times 0.9}{1}\Big(\tfrac{1}{2\times 0.9}0.2\Big)^{2^{n-1}} \le 2\big(\tfrac{1}{9}\big)^{2^{n-1}}$$