Nonlinear Equations: Newton's Method

Math 131: Numerical Analysis

J.C. Meza

2/8/24

Newton's Method

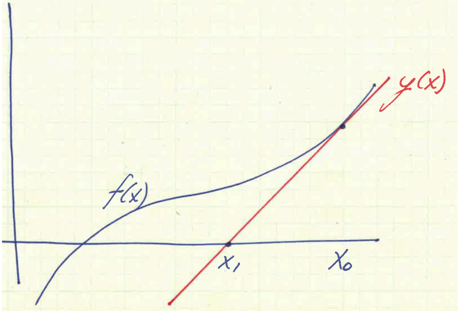
- In the last lecture we saw that the bisection method was robust and would always converge to a solution given the right set of initial values.
- However, it could exhibit slow convergence.
- As a result, most solvers use other types of root-finding algorithms.
- In this section we will study two such methods that can provide much faster convergence to a root.

Newton's Method (Quick Summary)

- Newton's method is likely the most popular and certainly the most powerful method for solving nonlinear equations [@meza2011newton].
- The idea behind Newton's method is to use the slope of the function at the current iterate to compute a new iterate.
- Naturally, this requires that we first assume that the given function $f(x)$ is differentiable.

Visually

- Note that if we take the derivative at the current iterate and use that to set up a linear equation, which we can solve for the new iterate.



.

Idea

Important

One approach for deriving Newton's method is to think about building a ***linear model of the function*** at the current iterate. Let's consider the linear model $m(x)$:

$$m(x) = f(x_0) + f'(x_0)(x - x_0). \tag{1}$$

Notice that at $x = x_0$ the model agrees with the function $f(x)$, in other words $m(x_0) = f(x_0)$. The idea is to then solve for the root of Equation 1 and use the root as the next guess of our iterative method:

Solving for new iterate

Using this idea let's solve for the root x^* of the linear model, $m(x)$, i.e.

$$
\begin{aligned}
m(x) = f'(x_0)(x - x_0) + f(x_0) = 0, \\
\implies x^\star = x_0 - \frac{f(x_0)}{f'(x_0)}.
\end{aligned}
$$

We can then set $x_1 = x^\star$ as the next iterate in our sequence and repeat the process. This gives us the general procedure for Newton's method:

Newton's Method

$$
x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}, \quad k = 0, 1, \ldots \tag{2}
$$

Another derivation

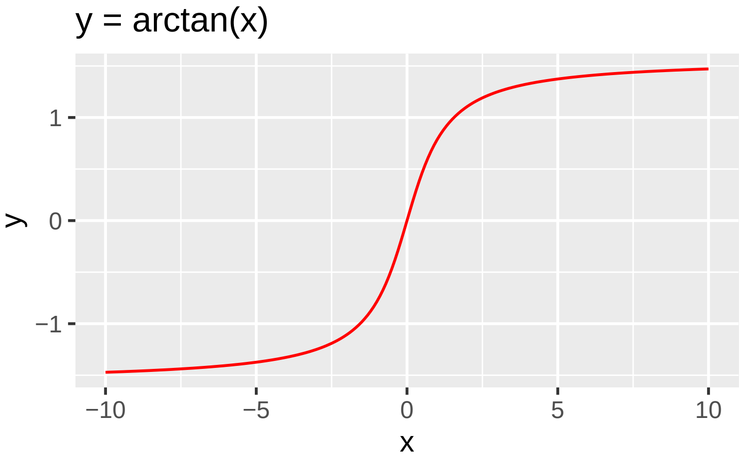
- We will note in passing that another derivation is to use Taylor's theorem to approximate our function $f(x)$ out to the first degree with a remainder term that includes the second derivative.
- We will ignore the second derivative term based on the argument that when we are near the solution the term would be small.
- Solving for our new iterate, we can derive the same equation as before.

Remark

A natural question to ask is under what conditions does Newton's method converge?

- In fact, it isn't hard to show that if the initial point x_0 is not chosen properly (i.e. close enough to a root), Newton's method will diverge.
- A typical example would be $y = \arctan(x)$, where if x_0 isn't close enough to the root the iterates quickly diverge to infinity.

Arctan(x)



Example

Let $f(x) = x^6 - x - 1 = 0$ and let $x_0 = 1.5$. It is easy to verify that one root is given by $x^* = 1.134724.$

To use Newton's method we first need to calculate the derivative - $f'(x) = 6x^5 - 1.$

Using Equation 2 allows us to compute the $k+1$ iteration:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)},$$
$$x_{k+1} = x_k - \frac{x_k^6 - x_k - 1}{6x_k^5 - 1}.$$

Example (cont.)

Proceeding in the natural way from x_0, we can generate the following sequence of iterates:

k		x_k
0	1.5	1.5
1	$x_1 = x_0 - \frac{x_0^6 - x_0 - 1}{6x_0^5 - 1} = 1.5 - \frac{8.8906}{44.5625}$	1.3005
2	$x_2 = x_1 - \frac{x_1^6 - x_1 - 1}{6x_1^5 - 1} = 1.3005 - \frac{2.5373}{21.3197}$	1.1815
3	$x_3 = x_2 - \frac{x_2^6 - x_2 - 1}{6x_2^5 - 1} = 1.1815 - \frac{0.5387}{12.8140}$	1.1395

Notice that after only 3 iterations, the iterates is already correct to 3 significant digits.

Analysis

Several questions one might consider at this point include:

- Under what conditions might we expect (local) convergence?
- Here by local we mean that the algorithm will converge if we start sufficiently close to a root. We will define this more carefully later.
- If Newton's method converges, how fast can we expect the convergence to be?

Error Analysis for Newton's Method

Let's consider the Taylor expansion about $x = x^*$.

$$0 = f(x_k) + (x^* - x_k)f'(x_k) + \frac{(x^* - x_k)^2}{2} f''(\xi).$$

Dividing by $f'(x_k)$ (we will assume for the time being that it's not equal to zero for any x_k) we get:

$$0 = \frac{f(x_k)}{f'(x_k)} + (x^* - x_k) + \frac{f''(\xi)}{f'(x_k)} \frac{(x^* - x_k)^2}{2}.$$

Using the equation for Newton's method we see that the first term is nothing but $x_k - x_{k+1}$ and substituting into the above equation we get:

$$0 = x_k - x_{k+1} + (x^* - x_k) + \frac{f''(\xi)}{f'(x_k)} \frac{(x^* - x_k)^2}{2}.$$

Erro Analysis (cont.)

- We see that the x_k terms cancel out. Rearranging to put the error on the left-hand side of the equation yields:

$$x^* - x_{k+1} = -\frac{f''(\xi)}{2f'(x_k)}\ (x^* - x_k)^2. \tag{3}$$

- The quantity on the left-hand side of the equation is just the error at the $k+1$ iteration, while the last term on the right-hand side is the error at the k iteration (squared).

Interpretation

$$
\begin{aligned}
|e_{k+1}| = |x^* - x_{k+1}| &= \left|\frac{f''(\xi)}{2f'(x_k)}\right| \cdot \ (x^* - x_k)^2, \\
&= \left|\frac{f''(\xi)}{2f'(x_k)}\right| \cdot |e_{k+1}|^2,
\end{aligned}
\tag{4}
$$

- We can interpret the equation to mean that the error at the $k+1$ iteration is proportional to the square of the error at the k iteration.

Important

This type of error bound is called ***quadratic convergence***

Remark

- If $f \in C^2[a, b]$ and $f'(x^*) = 0$, then Newton's method still converges but just not as rapidly.
- Consider for example $f(x) = x^4$, which has a root at $x = 0$, but where the first derivative is also equal to 0.

Summary for Newton's Method

Table 2: **Newton's Method Summary**

Advantages	Disadvantages
Doesn't require interval with function sign change	Need to have derivatives
Fast convergence rate – quadratic	May not converge from all starting points
Can generalize to higher dimension	Can be expensive (especially in higher dimensions)