Newton's Method

Suppose $f : \mathbb{R} \to \mathbb{R}$ is differentiable and we want to find a number x^* such that

$$f(x^*) = 0.$$

One of the most recurring themes in applied mathematics is to replace a problem we don't know how to solve with one that we can solve. This often yields some useful information, and possibly a good approximate solution. Let's do that here by replacing $f(x) = 0$ with $p(x) = 0$, where p is the line tangent to f at the point $(x_0, f(x_0))$. This type of simplification is called *linearizing* f at $(x_0, f(x_0))$. That's it! Well, that's the key, at least. Let's do that that much: Recall that if $y = p(x)$ is the line with slope m passing through (x_0, y_0), then $y - y_0 = m(x - x_0)$. Substituting $m = f'(x_0)$ and $y_0 = f(x_0)$ gives the tangent line as

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

Let x_1 be the solution to $p(x) = 0$ (x_1 is the x-intercept of our tangent line)

$$x_1 = x_0 - f(x_0)/f'(x_0).$$

So the solution to the linearized problem is x_1. So what? Draw a picture. Literally (I mean actually get out pen and paper and) sketch a smooth function crossing the x-axis. Pick a point x_0 not too far from the true root and linearize. Your x_1 is closer to the solution than was x_0, right? Now linearize about x_1 to get an x_2, and you will see why this is an important method.

Newton's method (also commonly called the Newton-Raphson method) is simply this iteration:

$$x_{k+1} = x_k - f(x_k)/f'(x_k).$$

It's not really a method, yet, but the kernel of a method. To have a useful method we will need to (i) have x_0, (ii) be able to evaluate f at x_k, (iii) be able to evaluate f' at x_k, and (iv) know when to stop. Even given these things, we may divide by zero along the way, or have our iterates bounce around or cycle, never converging.

The good news is that we have a theorem that guarantees that if x_0 is close to x^*, then $x_k \to x^*$ as $k \to \infty$. The bad news is that since we don't know x^* it is hard to say if x_0 is close enough (the theorem also requires some information about f that we probably don't have). We can say that when it converges, it nearly always converges very fast (more on that later...).

Because of item (iii) above, we do *not* consider Newton's method a *general purpose* method. A general purpose root finder would typically require that the user provide a subroutine that would evaluate f at x_k, and optionally a starting value x_0 or possibly an interval to search. An iteration of Newton's method requires two function evaluations $f(x_k)$ and $f'(x_k)$; this requires more of the user than a typical general purpose method.