Newton's Method

Suppose we seek a critical point for some univariate function $f(x)$. We know this happens where $f'(x) = 0$. In other words, we are looking for r such that $f'(r) = 0$. Unfortunately solving the equation $f'(x) = 0$ is not always easy or even possible analytically. But we still might need an answer, so we could initially guess at a solution x_0. See Figure 1. Note this is a picture of $f'(x)$, not $f(x)$.

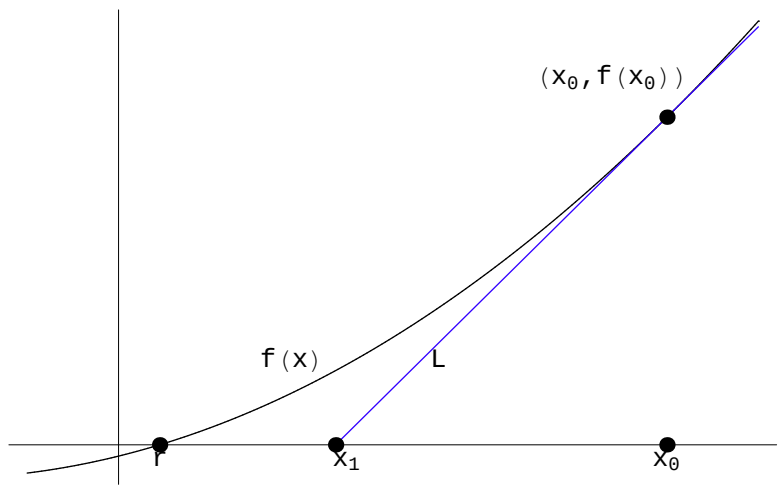

Figure 1: $f'(x)$

Now consider a tangent line L to $f'(x)$ at the point $(x_0, f'(x_0))$. The slope of this line is $f''(x_0)$, so using point-slope form we see that the equation of the line L is

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

Suppose that L intersects the x-axis at x_1 as shown in the figure. If we consider (1) at this point of intersection $(x_1, 0)$, we obtain

$$0 - f'(x_0) = f''(x_0)(x_1 - x_0). \tag{2}$$

Solving (2) for x_1, we obtain

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

This gives us a way to calculate x_1 as long as we have x_0 and can compute $f'(x_0)$ and $f''(x_0)$. Of course if we repeat the exact same line of reasoning we obtain

$$x_2 = x_1 - \frac{f'(x_1)}{f''(x_1)}$$

which gives rise to the idea that we could continue to compute $x_1, x_2, x_3, ...$ in hopes of getting closer to the true value r that we were looking for in the first place. The general form of this equation is

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)} \tag{3}$$

In higher dimension, (2) becomes

$$-\nabla f(x_0) = H(x_0)(x_1 - x_0).$$

We can solve this for x_1, we obtain

$$x_1 = x_0 - H^{-1}(x_0)\nabla f(x_0) \tag{4}$$

which generalizes to

$$x_{i+1} = x_i - H^{-1}(x_i)\nabla f(x_i). \tag{5}$$

Note that x_i and x_{i+1} are column *vectors* that each describe a point in the domain. $H^{-1}(x_i)$ is the inverse of the Hessian matrix evaluated at the point x_i, and $\nabla f(x_i)$ is the column vector representation of the gradient of f evaluated at x_i.

In (5), we can think of x_{i+1} as being equal to the previous point x_i plus an update $-H^{-1}(x_i)\nabla f(x_i)$. This allows us to compute $x_1, x_2, x_3, ...$ until we decide that we are "close enough". Normally, we stop when $\| \nabla f(x_i) \| < \epsilon$ where ϵ is set by that analyst (that's you!!).

An Example

Let's do a simple two variable example to make this clear. First recall we are using the notation x_i to describe the ith point in the domain that we find. You are probably used to using the notation $f(x, y)$ to describe a two variable

function. In this case, that will be confusing since we are already using x to describe the point; not just one coordinate. Accordingly, we will denote our two variable function as $f(a, b)$. Each point x_i has two coordinates a_i and b_i. In other words, $x_i = \begin{bmatrix} a_i \\ b_i \end{bmatrix}$.

Now we can move forward with our example. Suppose $f(a, b) = (a - 2)^4 + (a - 2b)^2$, with the initial point $x_0 = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$. To determine x_1, we follow (4). First, we compute

$$H(x_0) = \begin{bmatrix} 50 & -4 \\ -4 & 8 \end{bmatrix}.$$

Now it is easy to compute

$$H^{-1}(x_0) = \frac{1}{192} \begin{bmatrix} 4 & 2 \\ 2 & 25 \end{bmatrix}. \tag{6}$$

Finally we can evaluate the gradient at x_0, obtaining

$$\nabla f(x_0) = \begin{bmatrix} -44 \\ 24 \end{bmatrix}. \tag{7}$$

Next, we use our initial point x_0 along with (6) and (7) to compute x_1 as follows:

$$x_1 = x_0 - H^{-1}(x_0)\nabla f(x_0) = \begin{bmatrix} 0 \\ 3 \end{bmatrix} - \frac{1}{192} \begin{bmatrix} 4 & 2 \\ 2 & 25 \end{bmatrix} \begin{bmatrix} -44 \\ 24 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Thus $x_1 = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$ and we can continue to find new points from here using (5).