Lecture Notes: Root Finding

In order to use the normal equations to fit a model to a set of data, the model must be a linear combination of basis functions. To solve for nonlinear terms a different technique must be taken. There are several alternatives, including the Levenberg-Marquardt algorithm, the Gauss-Newton algorithm, and other variations of gradient descent. All these methods are iterative. They start with an initial guess (also called approximation) and then repeatedly calculate a next guess based upon the previous guess. Each of these successive guesses/approximations gets closer to the true solution. Typically the iterations are halted when the difference between successive approximations drops below a threshold. These techniques are known as **nonlinear regression**.

To demonstrate the technique we will look at the Newton-Raphson method (circa 1700) for **root finding**. Consider the following problem: how do we calculate the square root of 44? We know that

$$x^2 = 44 \tag{1}$$

and can rewrite this as

$$x^2 - 44 = 0 \tag{2}$$

Now consider a function $f(x)$ such that

$$f(x) = x^2 - 44 \tag{3}$$

We can plot this function using a curve over the $x, f(x)$ space. Using only integers we can calculate some points by hand that we know are near the solution. At $x = 6$, $f(x) = -8$, and at $x = 7$, $f(x) = 5$. Interpolating between these two points we can imagine a curve as shown in figure 1.

We desire to find where this curve crosses zero, or more precisely where $f(x) = 0$. As shown in the figure the true solution is at $x = 6.63324958....$

The Newton-Raphson method is an iterative approach that uses the derivative of the function to successively draw near the true solution. First we make a guess, say, 7:

$$x_0 = 7 \tag{4}$$

Using the derivative of the function, we iteratively refine the guess. The derivative of our example is

$$f'(x) = 2x \tag{5}$$

We can calculate where the derivative crosses zero because the derivative is linear. Figure 2 is useful to visualize the technique. The derivative can be defined as the slope of a line,

f(x)

5

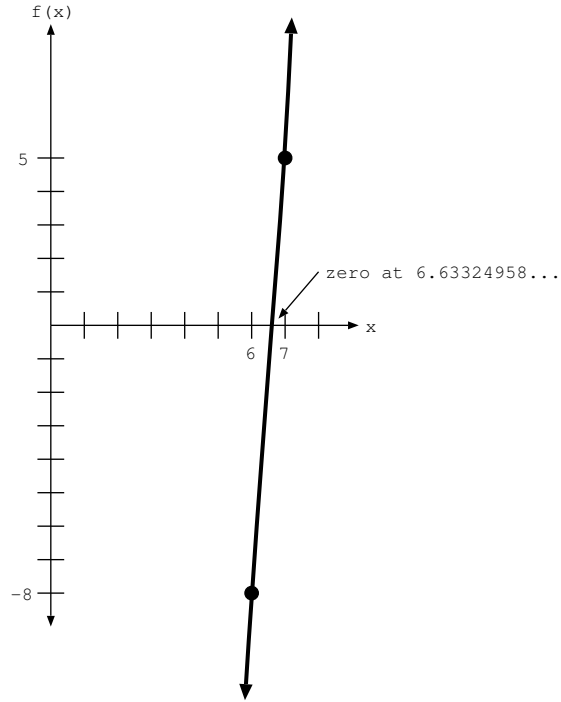zero at 6.63324958...

x

6 7

−8

Figure 1: Plotting a function to see where it crosses zero.

which is the rise (change in y) over the run (change in x):

$$f'(x_0) = \frac{f(x_0) - 0}{x_0 - x_1} \tag{6}$$

Rearranging terms yields

$$x_0 - x_1 = \frac{f(x_0)}{f'(x_0)} \tag{7}$$

and finally

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \tag{8}$$

In general, for each iteration we can write

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{9}$$

Continuing our example, we can use equation 9 iteratively until our approximation is
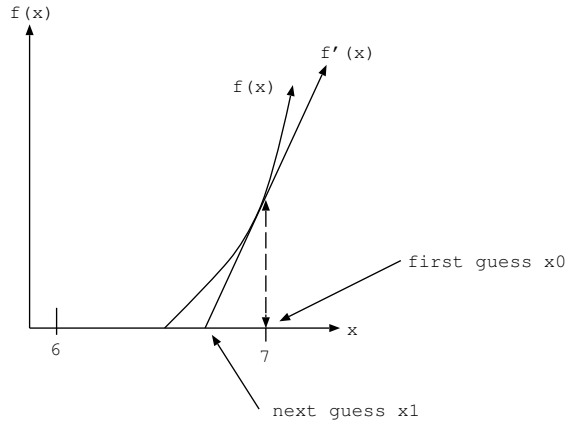
2

Figure 2: Plot of the derivative crossing zero.

sufficiently close:

$$x_0 = 7 \tag{10}$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 7 - \frac{7^2 - 44}{2(7)} = 6.6429 \tag{11}$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 6.6429 - \frac{6.6429^2 - 44}{2(6.6429)} = 6.6333 \tag{12}$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 6.6333 - \frac{6.6333^2 - 44}{2(6.6333)} = 6.6332 \tag{13}$$

$$\tag{14}$$

For this example the iterations converge quickly towards the true solution.

It is important to note that the calculations are assumed to take place with a fixed level of precision. (In the case of Netwon circa 1700, this would be equivalent to halting the by-hand calculations and rounding the final calculated digit; in modern times, this is equivalent to the level of precision provided by a computer program.) The level of precision is related to the desired accuracy in the solution. In the above example, I used 4 digits after the decimal point, but this is an arbitrary decision that can be changed as needed.

The following is a second example of the same technique. Suppose we wish to find x for the equation $sin(x) = x^3$. Let

$$f(x) = sin(x) - x^3 \tag{15}$$

and we seek to find the true solution where $f(x) = 0$. The derivative of $f(x)$ is

$$f'(x) = cos(x) - 3x^2 \tag{16}$$

Suppose an initial guess of $x_0 = 0.3$. Using equation 9 the successive approximations (all

3

calculations assume radians) are as follows:

$$x_0 = 0.3 \tag{17}$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0.3 - \frac{sin(0.3) - (0.3)^3}{cos(0.3) - 3(0.3)^2} = -0.0918 \tag{18}$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = -0.0918 - \frac{sin(-0.0918) - (-0.0918)^3}{cos(-0.0918) - 3(-0.0918)^2} = 0.0019 \tag{19}$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 0.0019 - \frac{sin(0.0019) - (0.0019)^3}{cos(0.0019) - 3(0.0019)^2} = -0.0000 \tag{20}$$

$$\tag{21}$$

It is easy to see that $x = 0$ is a true solution for where equation 15 equals zero. However, if we select a different initial guess $x_0 = 0.8$, the successive approximations are as follows:

$$x_0 = 0.8 \tag{22}$$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0.8 - \frac{sin(0.8) - (0.8)^3}{cos(0.8) - 3(0.8)^2} = 0.9679 \tag{23}$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0.9679 - \frac{sin(0.9679) - (0.9679)^3}{cos(0.9679) - 3(0.9679)^2} = 0.9309 \tag{24}$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 0.9309 - \frac{sin(0.9309) - (0.9309)^3}{cos(0.9309) - 3(0.9309)^2} = 0.9286 \tag{25}$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 0.9286 - \frac{sin(0.9286) - (0.9286)^3}{cos(0.9286) - 3(0.9286)^2} = 0.9286 \tag{26}$$

$$\tag{27}$$

The value $x = 0.9286$ is also an approximate solution for where equation 15 equals zero.

This illustrates one of the concerns with iterative methods: they are sensitive to the initial guess. In general, the guess should be close to the sought-after solution so that the method converges. Depending on the shape of the function, it is possible for the method to get into an unending cycle or follow a diverging path. In some cases, where there is more than one local minimum (as in the last example), the solution found may differ depending on the choice of initial guess.

Another issue with this particular iterative method is that the function must be continuous and differentiable. Other methods such as the secant method and bisection method avoid this restriction by using different approximations.

In practice, the choice of which iterative method to use depends somewhat upon the problem and shape of the function.