MAT1193 – 9a Approximating functions with polynomials

Most of the applications of the derivative we have been considering use that fact that 'locally' around a base point, a very simple function - the tangent line - provides a good approximation to the potentially very complex original function. For example, using the tangent line approximation allowed us to predict the behavior of a discrete time dynamical system near and equilibrium state.

While the tangent line can be a good approximation very near a base point, the function will eventually curve away, and the tangent line won't be such a good approximation. Is there any way to construct an approximation that remains pretty good even further away from the base point? The answer is yes, but there is a cost. The better you want your approximation to be, the more complicated the approximating function becomes. The common way to manage this tradeoff between the accuracy of approximation and the complication of the function is to consider **polynomial functions**. These are functions consist of sums of increasing powers of x where each power is multiplied by a constant. These constants are known as the **coefficients** of the polynomial and the maximal power included is known as the **order** of the polynomial. For example, the polynomial

\[ f(x) = 7 - 3x + 14x^2 - \pi x^3 \]

is a third order polynomial (\(x^3\) is the highest power of x) with coefficients 7, -3, 14, and \(\pi\). First order polynomials are **linear polynomials** and have graphs that are lines. For example, \(g(a) = 3a - 13\). Second order polynomials are known as **quadratic polynomials**, and third order polynomials are **cubic polynomials**. An example of a cubic polynomial is \(r(z) = -z^3 + 23z + 1000\). (The term 'quadratic' traces back to the Latin word for a square, which has four sides.)

Lower order polynomials are simpler and can be specified with fewer coefficients. However, they can only approximate functions over a narrow range of values. Higher order polynomials are more complicated functions, but can do a better job of approximating the curvature of complex functions. Here is one example showing successively higher order approximations (dashed) to the main function (solid black). A zoomed in version is shown at the right.
Now suppose we are given a function and we want to approximate that function near some base point that we are interested in (e.g. it is an equilibrium value for a DTDS or for some other reason). To be specific, suppose we want to approximate the function \( f(x) = \ln(x) \) near the point \( x_0=2 \). (Note that in the textbook they let \( a \) be the value of the base point.) Suppose also that we have chosen the complexity of the polynomial that we are going to use for our approximation, say a cubic approximation. In this section we’ll write the function that approximates \( f \), by putting a ‘hat’ on \( f \): \( \hat{f} \). (We read that by just saying ‘f hat.’)

\[
\hat{f}(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3
\]

How do we determine the coefficients \( c_0, c_1, c_2, \) and \( c_3 \) that give us the best approximation to the function near that base point? Our overall strategy is pretty simple:

Find a polynomial of the appropriate form that matches the first, second, third, etc. derivatives evaluated at that base point, for as many derivatives as you can.

Before jumping in to an example, I’ll introduce one trick that will make our lives much easier. Instead of writing \( \hat{f}(x) \) as a normal polynomial in powers of \( x \),

\[
\hat{f}(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3
\]

we will write \( \hat{f}(x) \) in powers of \( \Delta x \), where \( \Delta x \) is the distance from our base point:

\[
\Delta x = x - x_0
\]

\[
\hat{f}(x) = c_0 + c_1 \Delta x + c_2 \Delta x^2 + c_3 \Delta x^3 = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + c_3 (x - x_0)^3
\]

In doing this, we will get different values of the coefficients, \( c_0, c_1, c_2, \) and \( c_3 \). But if you had to, you could do the algebra to multiply out all the terms like \( (x-x_0)^3 \) and derive the coefficients for the polynomial in \( x \) from the new coefficients in the polynomial in \( \Delta x \). But that whole process would get pretty ugly, so we’ll just worry about what happens when we write

\[
\hat{f}(x) = c_0 + c_1 (x - x_0) + c_2 (x - x_0)^2 + c_3 (x - x_0)^3
\]

For our specific example, we would write

\[
\hat{f}(x) = c_0 + c_1 (x - 1) + c_2 (x - 1)^2 + c_3 (x - 1)^3
\]

So we want to match this approximating function to our original function \( f(x) = \ln(x) \) near the base point \( x_0=1 \). The most basic thing we can do is to make sure that the approximating function is exactly equal to the original function exactly at the base point: \( \hat{f}(x_0) = f(x_0) \). But if \( x = x_0 \) then \( \Delta x=0 \) and

\[
\hat{f}(x_0) = c_0 + c_1 (x_0 - x_0) + c_2 (x_0 - x_0)^2 + c_3 (x_0 - x_0)^3 = c_0 \rightarrow c_0 = \hat{f}(x_0)
\]
so now we can write

\[ \hat{f}(x) = f(x_0) + c_1(x - l) + c_2(x - l)^2 + c_3(x - l)^3 \]

For our specific example

\[ \hat{f}(l) = c_0 + c_1(l - 1) + c_2(l - 1)^2 + c_3(l - 1)^3 = c_0 = \ln(l) = 0 \quad \text{so} \]
\[ \hat{f}(x) = 0 + c_1(x - l) + c_2(x - l)^2 + c_3(x - l)^3 = c_1(x - l) + c_2(x - l)^2 + c_3(x - l)^3 \]

Now let’s make sure that the first derivative, evaluated at our base point, is the same for our original function and our approximating function. For our original function,

\[ f(x) = \ln(x) \rightarrow f'(x) = 1/x \quad \text{and so} \quad f'(x_0) = f'(1) = 1/1 = 1. \]

Now let’s take the derivative of our approximating function. Using the sum rule, the power rule, the constant product rule, and the chain rule we find

\[ \hat{f}(x) = f(x_0) + c_1(x - x_0) + c_2(x - x_0)^2 + c_3(x - x_0)^3 \]
\[ \hat{f}'(x) = c_1 + 2c_2(x - x_0) + 3c_3(x - x_0)^2 \]
\[ \hat{f}'(x_0) = c_1 + 2c_2(x_0 - x_0) + 3c_3(x_0 - x_0)^2 = c_1 \]

For our specific example \( f(x) = \ln(x) \), this works out to be

\[ \hat{f}(x) = f(x_0) + c_1(x - l) + c_2(x - l)^2 + c_3(x - l)^3 \]
\[ \hat{f}'(x) = c_1 + 2c_2(x - l) + 3c_3(x - l)^2 \]
\[ \hat{f}'(l) = c_1 + 2c_2(1 - l) + 3c_3(1 - l)^2 = c_1 \]

Since we want the first derivative to match the original function we have

\[ \hat{f}'(x_0) = c_1 = f'(x_0), \quad \text{and for} \quad f(x) = \ln(x) \quad \text{we have} \quad \hat{f}'(l) = c_1 = f'(l) = 1. \]

We’re making progress and can write our approximation as

\[ \hat{f}(x) = f(x_0) + f'(x_0)(x - x_0) + c_2(x - x_0)^2 + c_3(x - x_0)^3 \quad \text{(in general)} \]
\[ \hat{f}(x) = f(x - l) + c_2(x - l)^2 + c_3(x - l)^3 \quad \text{(specific example)} \]

It is very tempting to just see the pattern here and think that \( c_0 = f(x_0), \ c_1 = f'(x_0), \ c_2 = f''(x_0), \ c_3 = f'''(x_0), \) etc. But the would be WRONG! To see this we need to keep taking derivatives.

The second derivative is just the derivative of the derivative, so

\[ \hat{f}''(x) = 2c_2(x - x_0) + 3 \cdot c_3(x - x_0)^2 \]
\[ \hat{f}''(x) = 2c_2 + 3 \cdot 2 \cdot c_3(x - x_0) \quad \text{(general)} \]
\[ \hat{f}''(x_0) = 2c_2 \]
To make the second derivative match between our approximation and the original function, we need \(2c_2 = f'(x_0)\) or \(c_2 = f''(x_0)/2\). In our specific example \(f(x) = \ln(x)\) we had \(f'(x) = 1/x = x^{-1}\). Applying the power rule, \(f''(x) = -1 \cdot x^{-2} = -1/x^2\). Plugging in \(x_0 = 1\), we have \(f'(x_0) = f'(1) = -1/1 = -1\). So the approximation is now

\[
\hat{f}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + c_3(x - x_0)^3
\]

(in general)

\[
\hat{f}(x) = (x - 1) - (x - 1)^2 + c_3(x - 1)^3
\]

(specific example)

We just have to find \(c_3\). So taking another derivative (the third), we have

\[
\hat{f}'(x) = 2\left(\frac{f''(x_0)}{2}\right) + 3 \cdot 2 \cdot c_3(x - x_0)
\]

\[
\hat{f}''(x) = 3 \cdot 2 \cdot c_3
\]

To make the third derivative match between our approximation and the original function, we need \(3 \cdot 2 \cdot c_3 = f'''(x_0)\) or \(c_3 = f'''(x_0)/(3 \cdot 2)\). In terms of our specific example, we had \(f'(x) = 1/x^2 = x^{-2}\), so \(f'''(x) = (-2) \cdot x^{-3} = 2/x^3\). Plugging in \(x_0 = 1\), we have \(f'''(x_0) = f'''(1) = 2/1 = 2\). So the final approximation is

\[
\hat{f}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f'''(x_0)}{3 \cdot 2}(x - x_0)^3
\]

(in general)

\[
\hat{f}(x) = (x - 1) - (x - 1)^2 + \frac{2}{3 \cdot 2} (x - 1)^3 = (x - 1) - (x - 1)^2 + \frac{1}{3} (x - 1)^3
\]

(specific example)

Notice that I didn’t multiply out \(3 \cdot 2 = 6\). That was to emphasize what the true pattern is in terms of finding the coefficients for our polynomial approximation. First, notice that when you plug in \(x = x_0\) all the other terms dropped out, leaving \(c_3\) to be determined by the first derivative, \(c_2\) to be determined by the second derivative, etc. Second, each time you take a derivative, the power rule brings down the power out front and then reduces the power by 1. So when you get to the third derivative, you’ve multiplied by \(3 \cdot 2\). If we wanted to do a fourth order polynomial approximation, for \(c_4\) we need the fourth derivative of the term \((x-x_0)^4\) and we’d have pulled down the power of 4, the power of 3 and the power of 2, so that \(c_4 = f^4(x_0)/(4 \cdot 3 \cdot 2)\).

Therefore, our final formula for approximating a function near a base point \(x_0\) with a polynomial function of order \(n\) is

\[
\hat{f}(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \ldots + \frac{f^{(n)}(x_0)}{n! \cdot (n-1)! \cdot \ldots \cdot 2 \cdot 1} (x - x_0)^n
\]

The approximating polynomial of power \(n\) is called the **nth order Taylor polynomial**, named after 18th century mathematician Brook Taylor.
One thing to point out is that the 1\text{st} order approximation is

\[ \hat{f}(x) = f(x_0) + f'(x_0)(x - x_0) \]

This is just the point slope form of the tangent line at $x_0$, a line with slope $f'(x_0)$ passing through the point $(x_0, f(x_0))$.

Let's check out how good our approximations are for $f(x) = \ln(x)$ near $x=1$. In particular let's check out the value of the approximation evaluated at $x = 1.101$, 1.1, and 1.5 and 2. We'll do this for 1\text{st} (tangent line), 2\text{nd}, 3\text{rd}, and 4\text{th} order approximations.

\begin{tabular}{|c|c|c|c|c|c|}
\hline
 & x & 1 & 1.01 & 1.1 & 1.5 & 2 \\
\hline
ln(x) & 0 & .0100 & 0.0953 & 0.4055 & 0.6931 & \\
\hline
1\text{st order} & $\hat{f}(x)$ & 0 & .0100 & 0.1000 & 0.5000 & 1.0000 \\
\hline
2\text{nd order} & $\hat{f}(x)$ & 0 & .0100 & 0.0950 & 0.3750 & 0.5000 \\
\hline
3\text{rd order} & $\hat{f}(x)$ & 0 & .0100 & 0.0953 & 0.4167 & 0.8333 \\
\hline
4\text{th order} & $\hat{f}(x)$ & 0 & .0100 & 0.0953 & 0.4010 & 0.5833 \\
\hline
\end{tabular}