Taylor Series: Notes for CSCI3656

Liz Bradley

Department of Computer Science
University of Colorado
Boulder, Colorado, USA 80309-0430

lizb@cs.colorado.edu

Research Report on Curricula and Teaching CU-CS-CT005-02

Note: these notes are meant to supplement (and mostly replace) appendix A of the CSCI3656 textbook, *Applied Numerical Analysis*, by Gerald and Wheatley.

1 Introduction

One way to deal with a nasty, analytically intractable function like $\tan(\log_\pi x^{37.3})$ is to approximate it with a benign one. There are lots of ways to do this, almost all of which[1] define "benign" as "polynomial." The reason for this is that polynomials are easy to work with: to write down, to take derivatives, to make more or less complex, etc. In general, polynomials look like this:

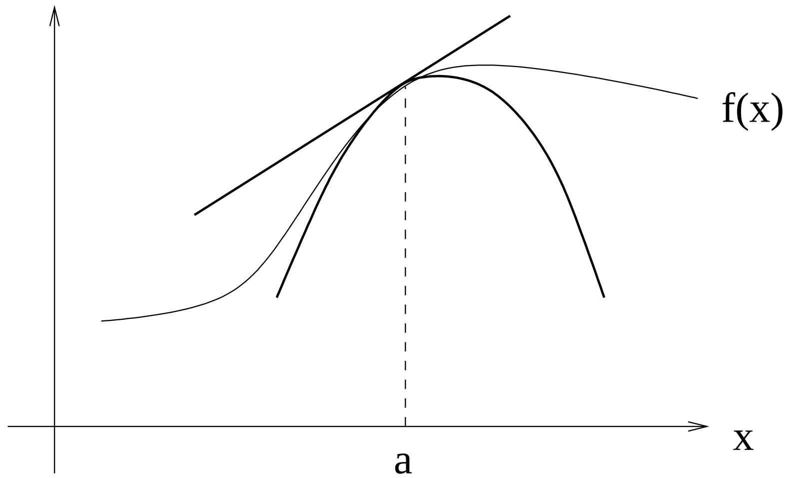
$$P_n(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_{n+1} x^n = \sum_{i=0}^{n+1} a_i x^i$$

where n is the "degree" of the polynomial: that is, the highest power of x that appears in it. (Polynomials can be in other variables besides x, of course, or even in multiple variables.)

At the top of the following page is a simple pictorial example of how one could use two simple polynomials — a line and a parabola — to approximate another, more-complicated function $f(x)$ in the region near a specific point a. Note that both of these fitting functions are close to $f(x)$ near the point a, that the quality of their fit to $f(x)$ degrades as one moves away from a, and that the higher-degree polynomial (the parabola) appears to do a

[1] One notable exception is Fourier decomposition, which uses combinations of sinusoids to approximate functions that are periodic in time.

better job across a wider range.



A **Taylor series** is a specific mathematical recipe for constructing a polynomial $P_n(x)$ of degree n that approximates a given function $f(x)$ near a point a. Here is the formula:

$$P_n(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \frac{(x-a)^3}{3!}f'''(a) + \ldots + \frac{(x-a)^n}{n!}f^{(n)}(a)$$

where $f'(x), f''(x) \ldots f^{(n)}(x)$ are the first, second, ... n^{th} derivatives of the function $f(x)$ with respect to x. This formula is on the CSCI 3656 formula sheet.

As an example, consider $f(x) = e^x$. We could fit a line to this function near some point $x = a$ by building a degree-one Taylor series like so:

$$P_1(x) = f(a) + (x-a)f'(a) = e^a + (x-a)e^a$$

(This makes use of the fact that the derivative of e^{kx} with respect to x is ke^{kx}.) If we wanted the line that fits e^x near $a = 0$, then, we'd simply plug in that value:

$$P_1(x) = e^0 + (x-0)e^0 = 1 + x$$

(Please draw a rough plot of e^x and $1+x$ so this makes sense to you.) If we wanted to fit a line to a different region of e^x instead — say, near $a = 1$ — we'd simply plug that a-value in:

$$P_1(x) = e^1 + (x-1)e^1 = 2.718 + 2.718(x-1) = 2.718x$$

(Again, please draw a picture of this for yourself.) If we wanted to fit a *parabola* to e^x near $a = 0$, we'd have to use one more term in the Taylor series:

$$P_2(x) = e^0 + (x-0)e^0 + \frac{(x-0)^2}{2!}e^0 = 1 + x + \frac{x^2}{2}$$

2 Why Bother

There are several reasons why you need to understand Taylor series and know how to build them:

- Because many, many numerical computation methods are based on these kinds of series.
- If you have a table of values of a function (e.g., e^x for $x = 0.1, 0.2, \ldots, 0.9$), you can use Taylor series to calculate its value at some in-between point (e.g., $e^{0.21}$).
- If working with a function would unnecessarily complicate your life and you can get away with something simpler, a Taylor series is often a good thing to try. In many graphics applications, for instance, the true effects of light falling on a complicated surface are both horrendously expensive to compute *and* effectively invisible to the human eye, so practitioners approximate those surfaces with simple curves instead.
- The notion of a series whose "goodness" increases with successive terms will help you understand error in numerical methods.

3 Taylor Series and Error

A Taylor-series approximation, in general, is good near the point where you built it. If I use $P_2(x) = 1 + x + x^2/2$ to obtain an estimate for e^0, for instance, my answer is perfect. (*Thought question: is this always true?*) As I move away from 0, the approximation gets worse:

$$\begin{array}{ll} e^0 = 1 & P_2(0) = 1 \\ e^{0.1} = 1.1052 & P_2(0.1) = 1.1050 \\ e^{0.5} = 1.6487 & P_2(0.5) = 1.6250 \\ e^2 = 7.389 & P_2(2) = 5 \\ e^5 = 148.4 & P_2(5) = 18.5 \end{array}$$

The inherent error in a Taylor-series approximation is also related to the match between the complexity of the approximation and the complexity of the underlying function. If $f(x)$ is a line and you fit a degree-one Taylor polynomial $P_1(x)$ to it, your answer will be exact — and not only at the point where you built the polynomial, but everywhere. If $f(x)$ is a *parabola*, you'll need a degree-*two* Taylor polynomial for a perfect global fit. *(If $f(x)$ is a line and you try to fit a degree-two Taylor polynomial to it, what do you think will happen? Please try this and see.)* In general, you need the degree n of $P_n(x)$ to be at least as large as the "degree" of $f(x)$ in order to get a perfect global fit, and the error in your results will depend on how much smaller n is than the "degree" of $f(x)$.

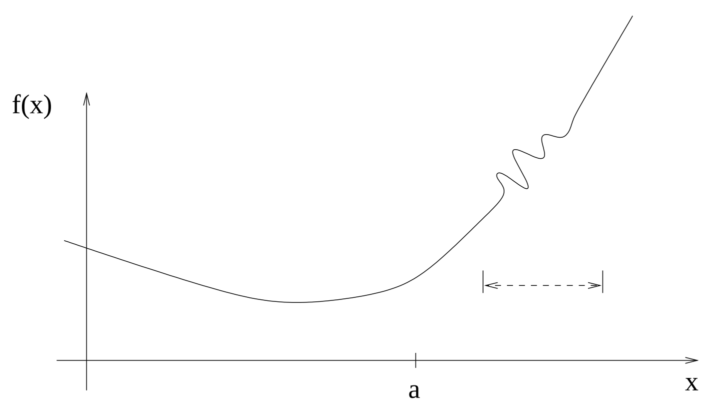
The quotes above are important; the word "degree" only makes sense for polynomials, and if $f(x)$ were a polynomial, we wouldn't be bothering with Taylor series at all. Nonetheless, the generalized notion of the degree of a function as a way to assess (and compare) complexity is useful in developing intuition about how all of this works. We'll talk more about this later and make the associated ideas more clear.

Here is a formula that encodes all of those concepts. The error in an n^{th} degree Taylor-series polynomial approximation to a function $f(x)$ is

$$R_n(x) = f(x) - P_n(x) \approx \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(c_x)$$

This formula is also on the CSCI 3656 formula sheet. With the exception of the c_x term — which is confusing and to which we'll come back to below — this is pretty easy to pick apart and understand. The first piece of that formula ($(x-a)^{n+1}/(n+1)!$) captures the "the fit is perfect close to a and degrades as you move away from a" idea. The second piece — the $n+1^{st}$ derivative of f — captures the "the degree of $P_n(x)$ has to be at least as large as the "degree" of $f(x)$" argument. (*Thought experiment: what is the $n+1^{st}$ derivative of a degree-n polynomial?*) And the whole thing looks suspiciously like the next term you'd have added to the series to get $P_{n+1}(x)$. This is a common heuristic in numerical methods: you can estimate the error in your results using the next term that you would have added to make things better. This is treated in more depth in the CSCI 3656 Error Notes (Research Report on Curricula and Teaching CU-CS-CT004-02).

Consider the task of fitting this function with a Taylor series:



If I wrote down a degree-two Taylor-series approximation to this function near the point a, the resulting $P_2(x)$ would be a good fit to $f(x)$ in some regions and a bad fit in others — specifically, in the wavy region near the right identified with the dashed line. Moreover, error estimates must always be pessimistic, so any calculations of the error in my $P_2(x)$ must be done using the worst possible conditions. That means that the $R_n(x)$ equation should be evaluated at the *worst possible point* – that is, the x that makes it largest.

The c_x term in the $R_n(x)$ formula captures these ideas. It is a "worst case" factor — a placeholder that means "find the x that makes this worst and plug it in." For example, the

Taylor-series fits to e^x near $x = 0$ that are given on the previous pages have the following error:

$$R_n(x) = \frac{(x-0)^{n+1}}{(n+1)!} e^{c_x}$$

These fits were *constructed* at $a = 0$, but they are *used* at some other x, so in order to evaluate the error that they may contain, we need to find the "worst-case" value in the interval $[0, x]$. In this case, the $R_n(x)$ function is monotonic upwards across this interval and the answer is relatively easy: the error is biggest at the right-hand end of the interval, so $c_x = x$ and

$$R_n(x) = \frac{(x-0)^{n+1}}{(n+1)!} e^{x}$$

In general, it's not always easy to look at an $R_n(x)$ function and see what x makes it biggest; a good general strategy is to evaluate it at the endpoints of the interval, check (using derivatives) to see if there's a maximum inside the interval and, if so, evaluate the $R_n(x)$ function there too, then take the largest value of the whole lot.

Of course, all of this is somewhat of an academic exercise, since these calculations presuppose that we know a fair bit about our answer. Nonetheless, this formula does have some practical utility. For instance, if we want to know how many terms to put into the Taylor series to obtain a particular level of accuracy, the $R_n(x)$ formula can help. If I wanted to build a Taylor series for e^x around $x = 0$ that was accurate to two decimal places out to $x = 1$, for instance, I would have to use enough terms for this to be true:

$$R_n(1) = \frac{2.718}{(n+1)!} \leq 0.01$$

which translates, in this case, to $n \geq 5$.

All of the material in this section is covered in much more detail in any calculus text.