The art of approximation

Where do we come from? What are we? Where are we heading? Initially, our motivation came from a desire to "understand the world around us" (economics, biology, space, perspective in art or any other topic we may be interested in). In trying to acquire such an understanding, we discovered a need to describe relationships between the various components of this "world". Our interest in mathematics therefore stemmed from its standing as a universal language, allowing us to understand complicated relationships and explain them in a simple manner.

Throughout the term we have been exploring different steps of such an understanding process. Starting from our "world" or "system" (e.g. economics) we proceeded to extract relationships (e.g. supply and demand equations) which we interpreted using functions. For the most part, these functions turned out to be differentiable and calculus was the tool we developed to analyze them:

$$\text{system} \to \text{relationships} \to \text{functions} \to \text{calculus}.$$

Highlighting this thought process over and over again is probably the best antidote to the student's (mis)conceptions of calculus as a collection of random topics to be mastered disjointly. The purpose of the last few lectures outlined below is to complete this line of thought by finally showing the students how easy functions such as polynomials can be used to understand general differentiable ones. The take-home message for the students should be that differential calculus is a weapon best wielded through Taylor approximation.

Linear Approximation. As we have seen by now, general functions are just too wild to be understood completely. Continuous functions are nicer but they can still be hard to tackle. This is what led us to the class of differentiable functions: those which we hoped to understand in terms of simpler functions such as polynomials. The reason why we are forced to consider easier functions instead of handling differentiable ones directly is that even the simplest non-polynomial ones are hard to understand on their own. For instance, consider the function $f(x) = \sqrt{x}$. *Is it differentiable? What is its value at $x = 4$? What about at $x = 4.1$?* Most students will be surprised to find out that even their computer cannot tell them the exact value of $\sqrt{4.1}$!

This is where the main idea of (differential) calculus kicks in again:

$$\text{knowing a little bit about } f'(x) \text{ tells you a lot about } f(x).$$

Where have we seen this idea before? Students should recall this as the key allowing us to sketch the graphs of differentiable functions. Nevertheless, for many of our practical applications we might need more detailed information than a rough sketch and actually have to crunch some numbers. This is done by approximating differentiable functions as polynomials through so-called Taylor expansions.

How would you try to approximate a differentiable function using a polynomial? Students should be in the habit by now to deem this question as too general and seek one that is more restrictive in scope (and therefore easier to answer). *How would you try to approximate a differentiable function using a linear one?* Having found a more tractable question, they should be guided back to the source: *What does it mean for a function to be differentiable at a point? What does the derivative of a function actually tell us?* To begin, it would be worthwhile to review our definitions as a class. We initially introduced $f'(a)$ as the slope of the tangent line to the graph of f at $(a, f(a))$ which was obtained as a limit of the slopes of secants through points $(a, f(a))$ and $(a+h, f(a+h))$ as $h \to 0$. In order to guide students towards a sensible linear approximation, it could then be remarked that writing "$a + h$" is just a way of describing a "number close to a". Calling this number x, i.e., letting $x := a + h$, we obtain another familiar expression for the derivative

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}. \tag{1}$$

(It should be kept in mind that seemingly easy transitions such as this one can be quite tricky for students to grasp on the fly and suitable pictures may make it easier to digest.) *Given that this limit exists, what do we know about the quantities $f'(a)$ and $\frac{f(x)-f(a)}{x-a}$?* Recalling the definition of a limit, students should articulate that if x is very close to a then

$$f'(a) \approx \frac{f(x) - f(a)}{x - a} \tag{2}$$

or, re-arranging the terms,

$$f(x) - f(a) \approx f'(a)(x - a). \tag{3}$$

What does this equation look like? The class should eventually realize that this is basically the point-slope form of the equation of a line! *How could we use this approximate equation to cook up a linear function?* Well, we know that $\approx$ is not actually an equality for $f(x)$. However, since $f(a)$, $f'(a)$ and a are all constants, we can always consider the line whose equation is given by

$$L(x) - f(a) = f'(a)(x - a) \tag{4}$$

for a linear function

$$L(x) := f(a) + f'(a)(x - a). \tag{5}$$

Once again, even though this step is fairly straightforward, one should be aware that it will be confusing to most students and they should be given a moment to digest it as a class. One way to gage whether they are still understanding what's going on is to ask the following question: *What does the graph of this linear function $L(x)$*

look like? Having correctly identified the tangent line, the crucial question remains: *Why is $L(x)$ useful?* Here, the answer students should get at is two-fold. On the one hand, many differentiable functions and their derivatives are easy to evaluate at certain fixed values (e.g. $\sqrt{x}$ at $x = 4$, $x = 9$, $x = 16$, etc.) and, given a, $f(a)$ and $f'(a)$, it is very easy to compute values of $L(x)$. On the other hand, the fact that f is differentiable at a *ensures* that when x is very close to a, then $f(x)$ is very close to $L(x)$. In other words, we can get a feeling for the values of $f(x)$ near a by computing the values of $L(x)$.

This is a great place to revisit the function $\sqrt{x}$ and investigate its behaviour near $x = 4$. This could be followed by a consideration of e^x near the point $x = 0$. By now, the inquisitive students are probably starting to wonder about the range of validity of such approximations. A good transition into the topic could be to ask: *Why is our approximation an overestimate for $\sqrt{x}$ and an underestimate for e^x?* Aided by graphs, the class should come to realize this is an issue of concavity. In regions where the second derivative is negative, the graph of the function is trapped below its tangents and vice-versa. This is a first hint that higher order derivatives are telling us something about the effectiveness of our linear approximations.

How can we quantify the difference between $L(x)$ and $f(x)$? What do we expect such a difference to look like? Students will be quick to point out that once x is far away from a the values of $f(x)$ and $L(x)$ may have little to do with each other. This observation should lead them to articulate that the difference between $f(x)$ and $L(x)$ must be a *function* of x. Indeed, we can already make this precise:

$$R(x) = f(x) - L(x) = f(x) - f(a) - f'(a)(x - a) \tag{6}$$

is clearly a function of x since it is a sum of functions of x (Students might need to be reminded that $f(a)$, $f'(a)$ and a are all constants!). *What would you expect the limit of $R(x)$ to be as $x \to a$? What is this saying about our approximation?* Intuitively, students should expect the answer to be zero and articulate the meaning of this. They could also be encouraged to actually compute the limit as a sanity check. *Is this expression for $R(x)$ useful?* In order to answer this question, students should recall that our initial motivation was to understand the behaviour of $f(x)$ at points where we had no idea what the value of $f(x)$ actually was. Unfortunately, our formula for $R(x)$ involves $f(x)$ and consequently cannot give us the desired insight.

How could we possibly get around this problem? Recalling their observations about over and under-estimates for $\sqrt{x}$ and e^x, the students should be led to articulate a possible relation between $R(x)$ and $f''(x)$. Indeed, since the second derivative tells us how fast the first derivative is increasing or decreasing it controls how quickly the graph of f can "escape away" from the graph of its tangent line. A good way to illustrate this is by considering the difference between the graph of x^2 (resp. $1000x^2$) and its linear approximation near $x = 0$. Combining this thought with the known dependence of $R(x)$ on the distance between x and a, the following estimate becomes plausible to the students (and the eager ones should try and prove it!):

If $|f''(c)| \leq M$ for all c between a and x, then the error in the linear approximation to $f(x)$ at a is bounded as $|R(x)| \leq \frac{M}{2}|x-a|^2$.

Moreover, its usefulness is easily illustrated by our run-on examples and it paves the way to an intuitive understanding of Taylor and Lagrange's Theorems.

Taylor Approximation. Having gained a somewhat satisfying understanding of linear approximation, we can now return to our initial question:

How would you try to approximate a differentiable function using a polynomial?

Unfortunately, our previous approach does not generalize to polynomials in an obvious way so we need to ask ourselves more questions. *Can you think of another way to interpret our linear approximation? What are the similarities between a function and its linear approximation?* The direction in which the students should be nudged here is to observe that if we are approximating a function f at a linearly with L, then $f(a) = L(a)$ and $f'(a) = L'(a)$. *Is this a feature we could generalize to polynomials?*

A class discussion should eventually converge upon the following sequence of questions and answers. *Given a smooth function f, can you find a simple polynomial*

(1) p_0 *which agrees with f at a?*

The simplest possible answer: take the constant polynomial $p_0(x) := f(a)$.

(2) p_1 *which agrees with f and f' at a?*

To find the simplest possible answer, the students should observe that the minimal degree of such a p_1 is one, i.e., that $p_1(x) = c_0 + c_1 x$ is linear. *Can we find c_0 and c_1?* Using our constraints, we have that

$$p_1(a) = c_0 + c_1 a = f(a) \text{ and } p_1'(a) = c_1 = f'(a)$$

so

$$c_1 = f'(a) \text{ and } c_0 = f(a) - f'(a) \cdot a.$$

Rearranging our expression for $p_1(x)$ we see that we've stumbled back upon our initial linear approximation $p_1(x) = L(x) = f(a) + f'(a)(x-a)$. *Was this a coincidence?*

(3) p_2 *which agrees with f, f' and f'' at a?*

To find the simplest possible answer, the students should observe that the minimal degree of such a p_2 is two, i.e., that $p_1(x) = c_0 + c_1 x + c_2 x^2$ is quadratic. Using our constraints and solving a similar system of equations we find that

$$p_2(x) = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2}$$

works! *Why is the third summand divided by two?*

From here, the students should be able to see that this procedure could be iterated as many times as we want (provided that the derivatives exist). This is best illustrated by an easy example such as approximating e^x near $x = 0$ where the procedure is particularly easy to carrie out. Thinking of these Taylor polynomial approximations of f at a as the simplest solutions to the problem of finding a polynomial $p_n(x)$ which agrees with the first n derivatives of f at a will help ground this alien concept in the student's minds. Nevertheless, a question remains: *In what sense is* $p_n(x)$ *an approximation of* $f(x)$*?* Introducing Taylor and Lagrange's Theorems as answers to this question should, once again, help ground this alien concept in the student's mind.