Lecture 33

Newton's Method

As you know, the three big ideas I want you to get out of calculus are:

- Limits,
- Derivatives, and
- Integrals.

However, there are some underlying *themes* in calculus, and the biggest theme you've seen over and over again is the theme of *approximation*, or estimation.

For example,

1. To define a derivative: We recognized that the difference quotient was an approximation for the slope of a tangent line. So we took the limit as $h \to 0$ to define the derivative.
2. To define an integral: We recognized that Riemann sums approximated areas under a curve; so we took the limit as $n \to \infty$ to define the integral.
3. We saw a big application of derivatives, second derivatives, third derivatives, and so forth, by taking Taylor polynomials. This allows us to approximate values of complicated functions like sin using easy functions like polynomials.[1]

Today, we are going to learn another method of approximation, called Newton's Method.

[1]By the way, in good situations, if you take the limit of these Taylor polynomials of degree d as $d \to \infty$, you would recover the original function! You may see this in your next semester of calculus.

While Taylor polynomials allowed us to approximate the value of $f(x)$ given x, Newton's Method tries to reverse the process: Given a value of f that we want, can we find an x so that $f(x)$ equals that value?

Example 33.0.1. Let's say you want to find the cube root of 9. There are two ways you could try to compute this:

- There is a function called $g(x) = \sqrt[3]{x}$. The cube root of 9 is given by $g(3)$. So we could try to use Taylor polynomials to compute $g(3)$.

- There is a function called $h(x) = x^3$. We could ask, *where* does $h(x)$ equal 9? Newton's Method allows us to try and approximate the answer.

Of course, the end answers ought to be identical. But as you'll see, the geometry that comes out of these two different questions are quite different.

33.1 The idea of Newton's Method

Here is Newton's idea.

Preliminary step: Write f

(To reduce the question to one of finding roots.) First, we can always rephrase our question about "where does a function h equal a particular value c" to a simpler question: "where does a function f equal zero?" at the expense of changing our function. Let me illustrate this by example:

Example 33.1.1.

1. If you want to know where $h(x) = x^3$ achieves the value 2, you can ask where the function $f(x) = x^3 - 2$ achieves the value 0.

2. If you want to know where $h(x) = \sin(\cos(x))$ achieves the value 0.2, you can ask where the function $f(x) = \sin(\cos(x)) - 0.2$ achieves the value 0.

3. If you want to know where $h(x) = x^2 + 1$ achieves the value 9, you can ask where the function $f(x) = x^2 - 8$ achieves the value 0.

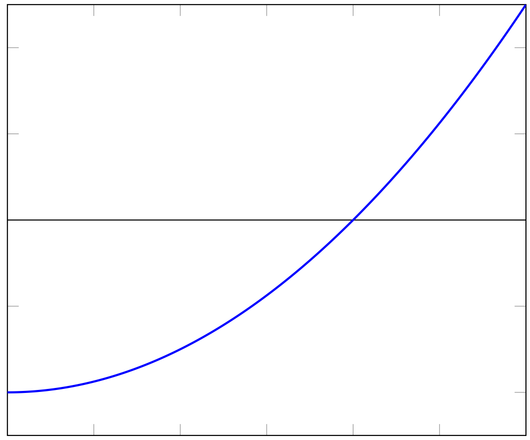
Remark 33.1.2 (Roots)**.** Remember that if $f(x_0) = 0$, then x_0 is called a *root* of f. What the preliminary step tells us is that if we can find roots of every function f, then for any function h and any value c, we can find an x_0 so that $h(x_0) = c$.

So, the preliminary step is to change our question:

$$\text{When does } h \text{ equal } c? \implies \text{When does } f \text{ equal } 0?$$

The trick is to find the appropriate f given h and c. And this is, in fact, easy: We always take $f(x) = h(x) - c$.

Here is what the graph of f might look like. We are interested in finding a root of f—that is, finding a place where f intersects the x-axis.



For example, if this were the graph of $f(x) = x^2 - 5$, the root of f would be $\sqrt{5}$.

Step 0. Choose a seed x_0.

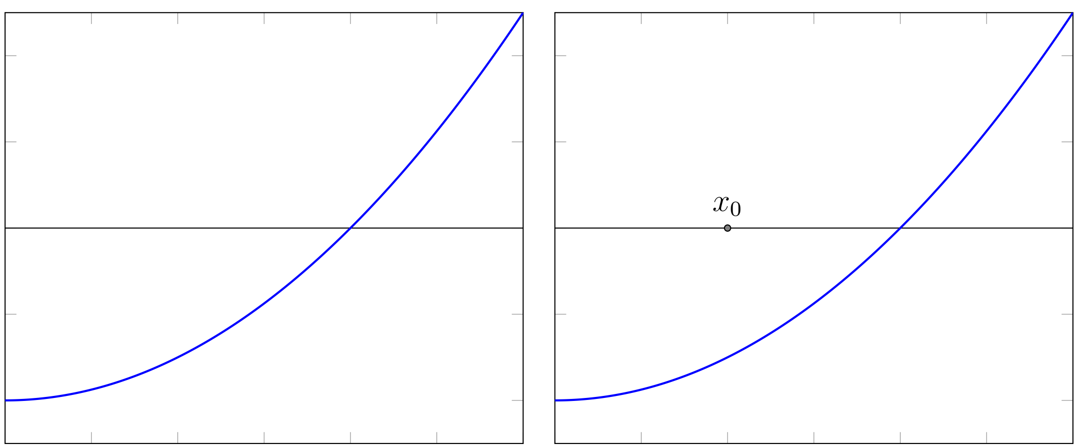
So suppose we are given a function f, and we want to find a root. (That is, we want to find a value of x for which f equals zero.) We first take a guess, and we call that guess x_0. This initial guess is called a *seed*.

Now, it may very well be that $f(x_0)$ does *not* equal zero. That's okay. x_0 might be thought of more as a point *near* the root, or an approximation to the root, rather than an actual guess that's meant to be correct. The closer x_0 is to the actual root, the better Newton's Method will work, but let's ignore that for now.[2]

So, Step zero is:

$$\text{Choose } x_0.$$

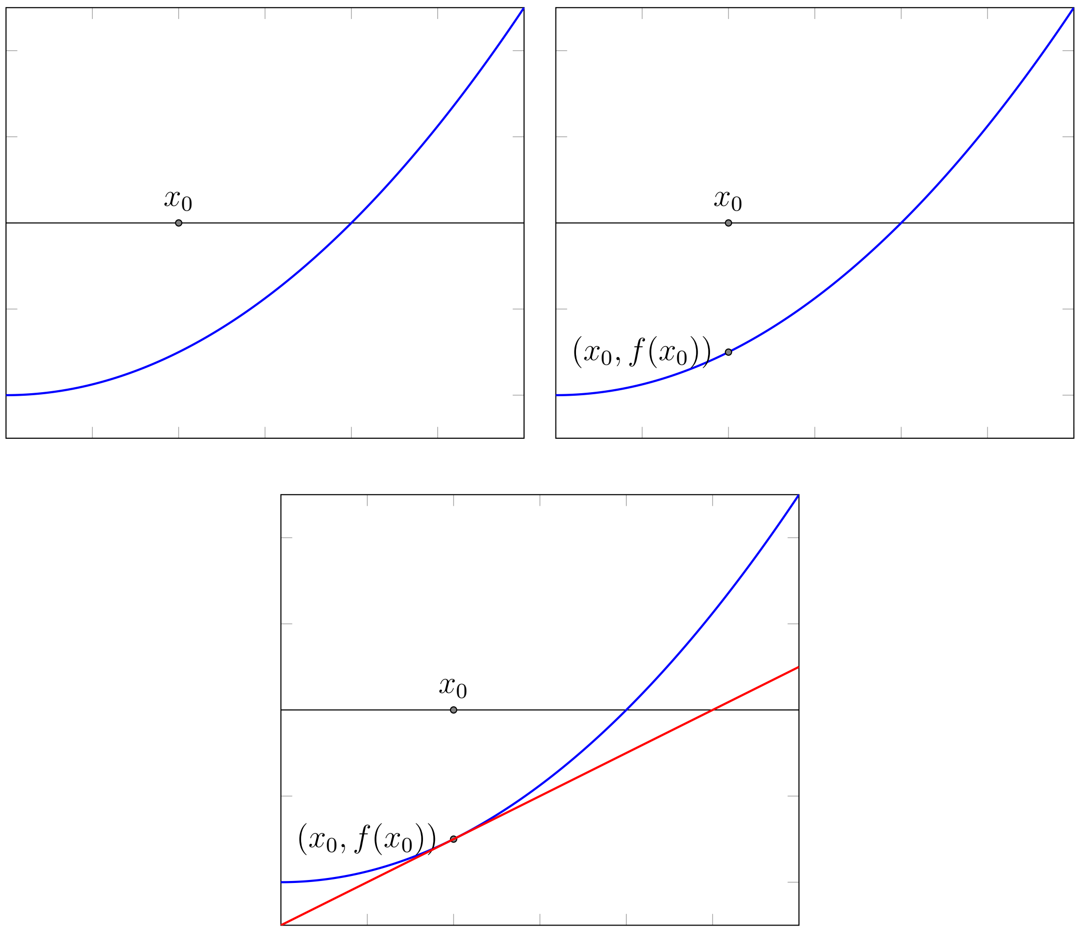
We emphasize that there is no "correct" choice here; it's up to you. There are better choices of x_0 than others (a better x_0 is one that happens to be closer to the actual root), but it may be hard to know what choices are good a priori.



[2]In fact, most calculus classes—including ours, unfortunately—will not describe *why* Newton's Method works, or why better guesses of x_0 result in better implementations of Newton's Method. For those of you who have felt like this class has already gone by quite quickly, you might sympathize with the sad reality that there just isn't enough time to cover (nor learn) everything.

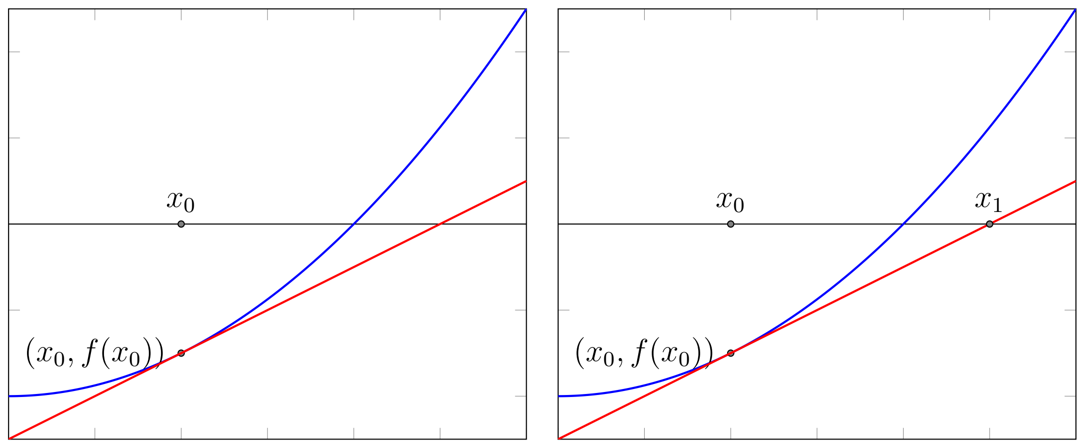
Step 1(a). Draw the tangent line at x_0.

The next step of Newton's idea is to compute $f(x_0)$, then draw the line tangent to f at the point $(x_0, f(x_0))$. Here is a picture:



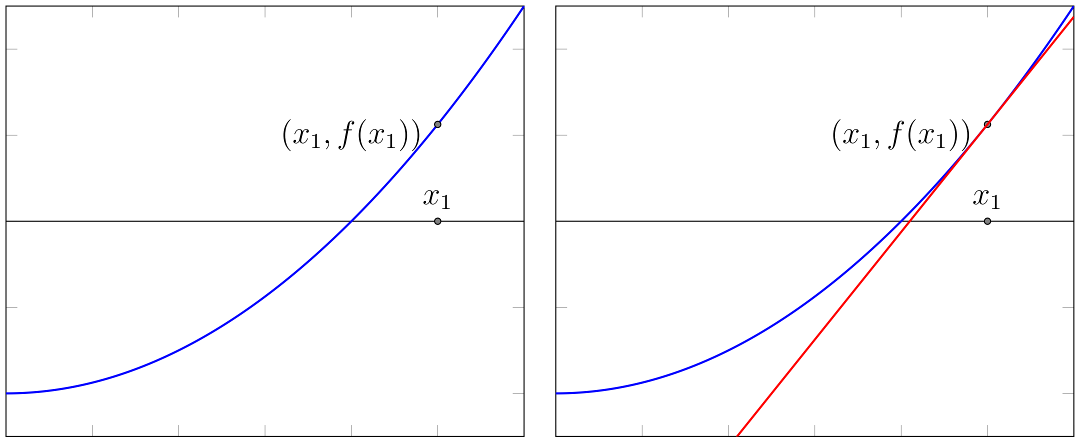
We have drawn in red the tangent line.

33.1.0 Step 1(b). Find x_1: where your tangent line meets the x-axis.



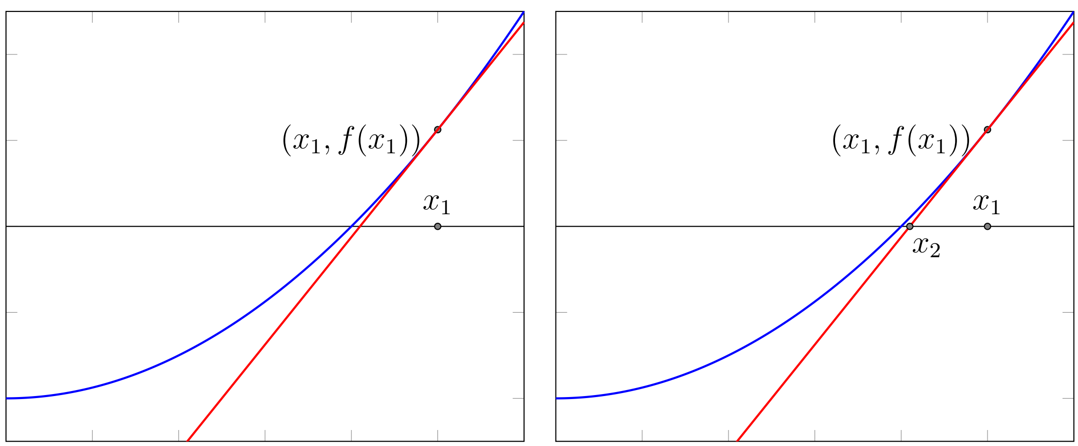
We have drawn in red the tangent line at $(x_0, f(x_0))$, and we have written as x_1 the x-coordinate at which the red line intersects the x-axis.

33.1.1 Step 2(a). Draw the tangent line at x_1.



Since we ascertained a number called x_1 in the previous step, we can now plot the line (drawn in red) that is tangent to the graph of f at the x-coordinate x_1.

33.1.2 Step 2(b). Find x_2: Where the tangent line to x_1 intersects zero.



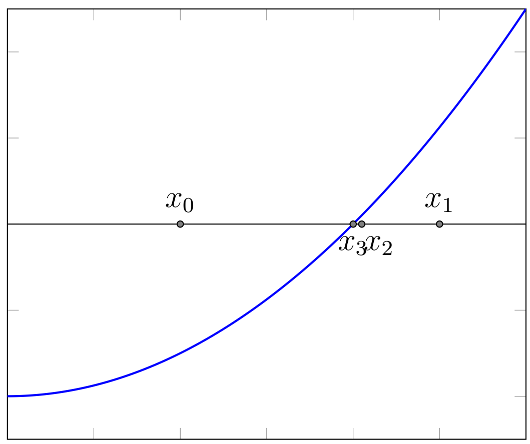
33.1.3 Step n(a): Draw the tangent line at x_{n-1}.

33.1.4 Step n(b): Find x_n, where the tangent line to x_{n-1} intersects zero.

33.2 What Newton's Method is supposed to do

As you can see, Newton's Method doesn't "end" after some number of steps. At the nth step, you find a new number called x_n, and you should think of x_n as the *nth approximation* to the root you seek. It is rare that x_n actually equals the root you seek, but in good situations, x_n is very close to the root you seek so long as n is large enough (that is, so long as you do enough steps).

For example, here is a drawing of x_0, x_1, x_2, x_3 for the example I drew out in the previous pages:



The take-away is that the numbers do get closer and closer to where the blue curve (the function) actually intersects the x-axis. In fact, at the scale of our sheet of paper, x_3 probably looks like it's exactly the root of f! My promise to you is, that in this example, x_4 is even closer to the root, x_5 is even closer, and so forth. This is how Newton's Method behaves in the best scenarios—you can get more and more accurate with each step, and in fact, you can get as accurate as you want.[3]

[3]Same caveat as before: We won't discuss how we know these facts, and we also won't discuss how close we know we are after n steps. This is a bit frustrating, because being able to say with confidence: "After n steps, we are within 100 decimal places" would be very convenient. It is possible to do this, but we won't discuss how in this class.

33.3 Practice

Let's use Newton's Method to approximate the square root of 4. (Yes, yes, I know that the square root of 4 is 2. Let's just see whether Newton's Method can get us close to this answer!)

Exercise 33.3.1. Let's begin with a (pretty bad) seed of $x_0 = 1$. (I say this is pretty bad because we know that the square root of 4 is 2; but we are trying out a number fairly far away from 2. A seed like 1.9 may have been far better.) Find x_1.

Note that to get started, we need to

1. Identify f. In this case, it's $f(x) = x^2 - 4$, as this intersects the x-axis when $x^2 = 4$.
2. Find $f(x_0)$. In this case, $f(x_0) = x_0^2 - 4 = 1^2 - 4 = -3$.
3. Find $f'(x_0)$. In this case, $f'(x) = 2x$, so $f'(x_0) = 2 \cdot 1 = 2$.
4. Find the intersection of the tangent line with the x-axis. In this case, the equation of the tangent line in point-slope form is
$$(y - f(x_0)) = f'(x_0)(x - x_0).$$
So when $y = 0$, we find
$$x = x_0 + \frac{f(x_0)}{f'(x_0)}.$$
5. Thus
$$\begin{aligned} x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} && (33.1)\\ &= 1 - \frac{f(1)}{f'(1)} && (33.2)\\ &= 1 - \frac{-3}{2} && (33.3)\\ &= \frac{5}{2}. && (33.4) \end{aligned}$$

Exercise 33.3.2. Continue and find x_2.

Exercise 33.3.3. Continue and find x_3. How close are your x_i to the actual root of 2?

33.4 Discussion of Newton's Method

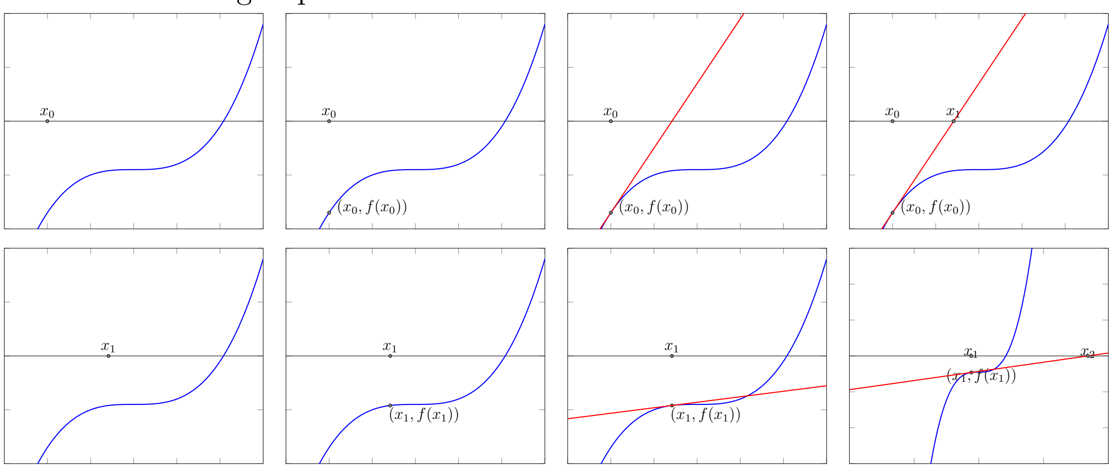
33.4.1 Newton's Method is recursive

Newton's Method is an example of a *recursive* algorithm. If you know what a for loop in computer programming is, you have seen recursive algorithms before. *Recursion* is a process by which we use a result from the previous step to compute something in the next step. Often, recursive algorithms repeat the same procedure every step; the only difference in each repetition is that the input may differ for each step.

For us, step n requires us to know x_{n-1} (that is, we need to know the output of the $(n-1)$st step). But the geometric process for taking x_{n-1} and outputting x_n is identical to all other steps of the method—you draw a tangent line, and find where it intersects the x-axis.

33.4.2 Newton's Method may fail

Consider the following implementation of Newton's Method:



In the last square, we had to *zoom out* to find where the red line intersected the x-axis. Note that we don't see how the blue curve (the graph of f) behaves at $f(x_2)$. What if the blue curve exhibits the same kind of behavior as at x_1, forcing the tangent line to look kind of flat, so that our x_3 ends up even farther away from the root? And what if this happens over and over again, so that each x_n just gets us farther away?

The reality is that *this could happen!* Indeed, for some functions, many choices of x_0 may force you into a terrible sequence of bad approximations. So, Newton's Method may fail. Keep this in mind.

Regardless, it turns out that *so long as one of your approximations x_i is close enough* to the actual root, then eventually, your approximations will become as close as you want to the actual root. This is the result of a theorem that I won't mention any further, but the proof of the theorem uses something called the *contraction principle*, which you'll learn about if you ever take an advanced differential equations class or an advanced numerical analysis class. And, how close does one of your x_i need to be to the root? That depends on the values of the derivative of f at points nearby the root. In practice, it may be that you have a lot of information on these derivative values, or it may be that you have to fly blind.

33.5 The fastest way to do Newton's Method

You may have noticed during your exercises that it's quite a hassle to write an equation for a tangent line and compute the x-intercept over and over again. Thankfully, there's an easy formula to get x_n based on x_{n-1}.

Note that the equation of a tangent line is (using point-slope form):

$$y - f(x_{n-1}) = f'(x_{n-1})(x - x_{n-1}).$$

So if $y = 0$, we have

$$-f(x_{n-1}) = f'(x_{n-1})(x - x_{n-1}) \tag{33.5}$$

$$-\frac{f(x_{n-1})}{f'(x_{n-1})} = (x - x_{n-1}) \tag{33.6}$$

$$x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})} = x. \tag{33.7}$$

So the point at which the tangent line intersects the x-axis (i.e., intersects the line $y = 0$) is given by

$$x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.$$

On the other hand, this x-value is, by definition, x_n! Hence we find the formula:

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}.$$

Memorizing this formula removes the geometric legwork of writing down a line and its equation and its x-intercept each iteration. However, you should also internalize the geometric algorithm, because it is very easy to misremember this formula. The geometry is forever.

33.6 Preparation for next time

33.6.1

(a) Approximate the cube root of 7 using Newton's Method, starting with a seed x_0, and computing all the way up to x_5. (You may use a calculator to find the x_i. Note that you should *never* have to use the cube root function on your calculator for this problem.)

(b) Compare your x_5 to what a calculator tells you the cube root of 7 is. (For the latter, you may use the cube root function.) How does your x_5 compare?

33.6.2 (Plus One) Basic word problems

A rectangular pen (which has four walls) will have three of its walls made of brick, and one of its walls made of glass. Brick walls cost 500 dollars per meter to make, while glass walls cost 700 dollars per meter to make. If the rectangular pen has length l meters and width w meters, and if the glass wall has length l meters, write a function (in terms of l and w) representing the total cost of a wall of length l and width w, in dollars.