Our next (and final!) topic for the course is optimization. This is all about finding where a function obtains its largest and smallest values. A central player in this subject is the degree 2 approximation. But first we will start with a review of degree 1 approximations as warmup.

Remember that the derivative of $f(x)$ at $x = a$ is defined to be the slope of the best linear approximation of $f(x)$ near $x = a$, or in other words we can write

$$f(x) = f(a) + f'(a)(x - a) + error$$

where the error is "very small" or "smaller than linear" (in a way which can be made precise by a limit).

This means that the function $f(x)$ should be very near to the line $f(a) + f'(a)(x - a)$, at least for x close enough to a.

So, for instance, if $f'(a)$ is positive, then that line has positive slope, hence is increasing (from left to right) and so the function $f(x)$ should also be increasing (near $x = a$). Similarly, if $f'(a)$ is negative, then $f(x)$ is well-approximated by a decreasing line, and hence should itself be decreasing.

To connect this to optimization: if $f(x)$ has a maximum at some point $x = a$, then it should be neither increasing at $x = a$ (because then we could step slightly to the right and get a bigger value!!) nor decreasing at $x = a$ (else we could step slightly left and find a bigger value). So at a maximum, $f'(a)$ should be zero. For more or less the same reasons, $f'(a)$ also has to be zero at a minimum.

There is one exception: if we can't step to the left or right then there is no issue. This happens if the domain of $f(x)$ is restricted so that it has right or left endpoints, in which case those endpoints also become possible maxima or minima.

For example, let's look at the function $f(x) = x^2 - 2x + 2$ restricted to x between -10 and 10. The derivative is $f'(x) = 2x - 2$, which is zero only if $x = 1$, and positive if $x > 1$ and negative if $x < 1$. So $f(x)$ is decreasing between -10 and 1, then increasing between 1 and 10. Its minima or maxima occur at either $x = 1$, where $f'(x) = 0$, or at the endpoints $x = -10, 10$. Plugging these into $f(x)$ we get:

$$f(1) = 1$$

$$f(-10) = 122$$

$$f(10) = 82$$

So the maximum of $f(x)$ is 122 at $x = -10$, and the minimum is 1 at $x = 1$, for x between -10 and 10. At $x = 10$, the value of $f(x)$ is larger than any nearby value (because they lie to the left, while the slope is positive). We call this a **local maximum** (and likewise for local minima: they are smaller than anything nearby, but maybe not the smallest value when we look at the entire function).

It would be helpful for us to be able to guess ahead of time whether a point is a max or min. This is where quadratic approximations enter: we understand this question very well for quadratic curves (parabolas). Namely, if the leading coefficient is positive, then the curve opens upward and it has a minimum; if the leading coefficient is negative, then the curve points down and has a maximum instead.

If we approximate a function to quadratic accuracy, we can tell whether it looks like an upward or downward parabola, and that tells us if we have a (local) min or max. Of course, there is a certain ambiguity that arises if the degree 2 coefficient of the degree 2 approximation is zero.

[worksheet examples]

A precise statement of our observations above is as follows:

The (local) mins and maxes of a function $f(x)$ occur at values of $x = a$ where $f'(a) = 0$, or at the endpoints of the domain.

Moreover, if the quadratic approximation of $f(x)$ has nonzero degree 2 coefficient, then: if that coefficient is positive, then the point is a local min, likewise if that coefficient is negative then that point is a local max.

It is helpful, sometimes, to know that we can put this statement in terms of multiple derivatives, rather than approximations, because sometimes the derivative rules make it easier to quickly figure out a derivative than to solve for an approximation.

This is the **second derivative test**. Suppose $f'(a) = 0$: if $f''(a) > 0$ then the point is a local min, and if $f''(a) < 0$ then the point is a local max. If $f''(a) = 0$ then we get no information. Let's show how this is connected to approximations. Take the best degree 2 approximation to $f(x)$ at $x = a$:

$$f(x) = A + B(x - a) + C(x - a)^2 + (error)$$

Since $f'(a) = 0$, we see that $B = 0$, and as always $A = f(a)$. So really we have

$$f(x) = A + C(x - a)^2 + (error)$$

Taking two derivatives, we get

$$f''(x) = 2C + (error)$$

Therefore $f''(a) = 2C$, and so if C is positive then so is $f''(a)$, hence the function has a local max. Likewise for negative and a min. If $f''(a) = 0$ then locally the function looks "very flat" and we get no information!! ****a better approximation could probably tell us whether it's a min or a max, though!!****