Class 3 Notes: Single Variable Derivatives

Myles D. Garvey, Ph.D

Summer, 2019

Contents

1 Introduction

In the last lecture, we discussed the idea of limits. As we mentioned there, almost all of limits rest on the idea of understanding what happens to a function as we approach, or "move", the input variable closer to another value. Limits serve as the foundation of calculus. However, another concept that also serves as a foundation of calculus is the concept known as the derivative. We showed in the motivation of calculus section that derivatives are highly useful for finding "optimal" values. This makes it especially attractive for business applications, since many businesses

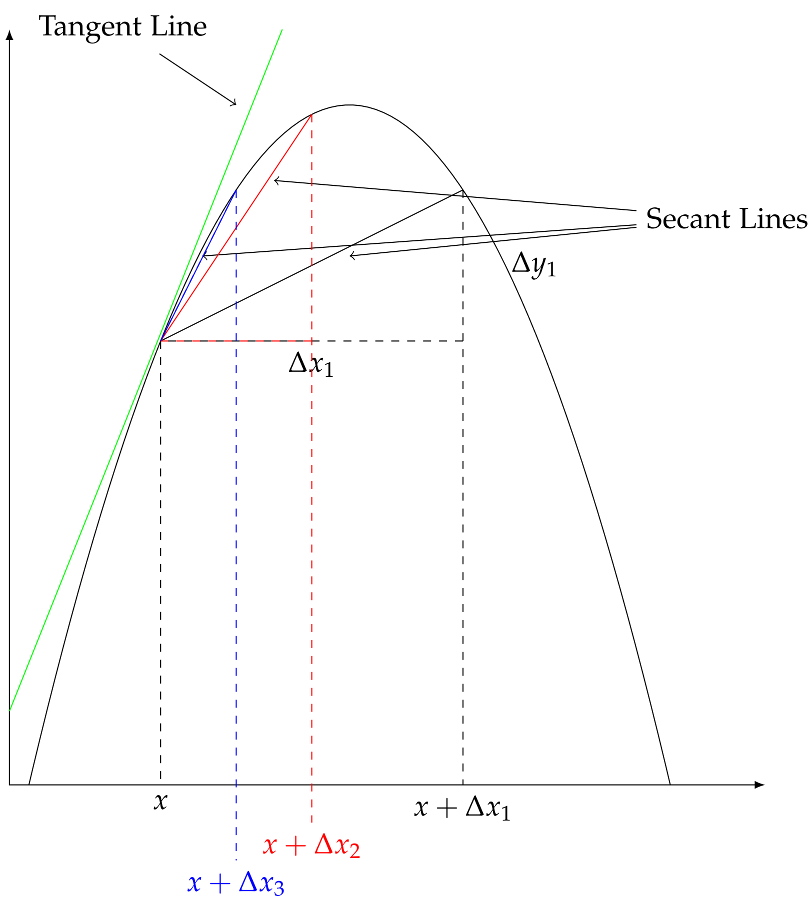

Figure 1: The Approximation of the Tangent Line (the green line) at $x = 2$ using secant lines (black, red, and blue) of $f(x) = -\frac{1}{2}(x - 2.5)^2 + 2x + 2$

are often trying to act "optimally". With the right tool box, we can apply these concepts to solve applied business problems.

In this class, we will cover the concept of the derivative in much greater detail than before. We will begin with it's basic definition and interpretation as a rate of change. We will then understand the differences between a derivative at a point versus as a function. After, we will review through the differentiation rules, which makes it easy for us to compute derivatives without resorting to evaluating limits. Next, we will discuss different properties of differentiation, as well as talk about multiple derivatives, yet another important concept that has many applications. We will close our lecture with profiling the toolbox that we use to solve optimization problems, as well as understand the opposite of derivatives: antiderivatives.

2 Definition of the Derivative

2.1 Rates of Change

The concept of rate of change rests on the idea that one quantity changes by a certain amount as a result of a change in a different quantity. A standard mathematical function only represents a state of something, not the change in state of something. For example, we may have a function

$f(x)$ which represents the cost of some decision level x. In this instance, the function itself represents the cost of a decision. We may be curious to know the extent of the change in cost if we change our decision. Suppose x represents our decision, and $f(x)$ represents the cost of making decision the x. If we increase our decision by a quantity Δx, then our new decision is $x + \Delta x$, where $\Delta x = x_2 - x_1$. However, we know may want to know how much the corresponding cost has changed. We can compute a new cost $f(x + \Delta x)$ and compare it to our old cost $f(x)$. The difference would represent the change in cost as a result of the change in the decision.

Definition 1 *Let $f(x)$ be a function. Generally, the rate of change of a function is the ratio between the level of change in $f(x)$ and the level of change in x. The average rate of change of $f(x)$ over an interval $x \in [a, b]$ is $\frac{f(b)-f(a)}{b-a}$.*

If we take the ratio of the change in cost to the change in decision, then we obtain a *rate of change*, which would give us the change in cost for every unit change in decision $\frac{f(x+\Delta x)-f(x)}{(x+\Delta x)-(x)} = \frac{f(x+\Delta x)-f(x)}{\Delta x}$. For example, if we decided to order $x = 5$ units, suppose this cost us a total of $f(x) = f(5) = 10$ dollars. If we decide to order an *additional* 2 units, $\Delta x = 2$, then our new decision is $x + \Delta x = 5 + 2 = 7$. Suppose this new decision will cost us $f(x + \Delta x) = f(5 + 2) = f(7) = 12$. Then the change in cost has been $f(x + \Delta x) - f(x) = 12 - 10 = 2$ dollars. Therefore, the *rate of change* in cost for every additional unit ordered would be $\frac{f(x+\Delta x)-f(x)}{\Delta x} = \frac{12-10}{2} = 1$. This says that for every *additional* unit we order, we can expect an increase in the cost (i.e. a change in the cost) of $1.

You learn in algebra about *linear functions*, and these are easy to compute the rates of changes. Pick any two points on the line, take the difference in y-values, Δy, take the difference in x-values Δx, and find the slope, $\frac{\Delta y}{\Delta x}$. This rate of change is the same for any decision x. In practice, however, cost functions and other functions are commonly *non-linear*. This means that the rate of change at different decisions may be different. Consider driving a car. If we want to know the rate of change in position and we drive exactly at 25 mph, then we can easily find our rate of change in position (that is, the number of miles we can expect to change for every change in time by one hour). This is of course a linear function, since we are driving at the same speed of 25 mph. If, on the other hand, we are driving at different speeds, say first at 25mph, then at 30mph, then at 20 mph, etc, it becomes difficult to estimate our rate of change. This is due to the fact that unlike driving at a steady 25 mph where the rate of change itself does not change over time, when we change speeds, our rate of change in position is itself changing. At one moment in time, we are experiencing greater changes in position as a result of change in time, due to driving "faster". At a different moment in time, we are experiencing lower changes in position as a result of change in time, due to driving "slower". As such, we need a new mechanism to determine the rate of change of a function at a given value for the independent variable.

2.2 The Derivative of a Function at a Point

When we want to find the rate of change of a function *at a given point*, we are essentially trying to find the line that is tangent to the function at the point. We can see this illustrated in Figure 1, where we seek to find the rate of change of $f(x) = -\frac{1}{2}(x - 2.5)^2 + 2x + 2$ at the point x. The idea is we would like to know the rate of change exactly at the point x. The green line is a line which has *tangent* to the curve at x, and the slope of this green line represents the rate of change at the point x. But how would we compute such a value?

The answer lies in computing an *approximation* of the tangent line with *secant lines*. We notice that it is difficult to exactly find the tangent line using only algebra. The reason is simple: there is only one point at which the tangent line and the curve $f(x)$ meet. You learn in algebra that in order to find the equation of a line, we need at least two points. (Theoretically, one could use geometric and trigonmetric arguments to derive the tangent line, but this is too cumbersome, and frankly, a waste of time). However, we can *approximate* the tangent line.

Pick another point, say $x + \Delta x$, where $\Delta x > 0$. Then we have two points, and we can find a line between them. This line is referred to as a *secant line*. From this, we can compute the slope of the line. We can do this as follows. Suppose we want to find the rate of change of $f(x)$ at $x = 2$. Pick $\Delta x = 4$ (this is an arbitrary decision, you can pick any quantity, even negative, that you wish). We now have another point which is at $x + \Delta x = 2 + 4 = 6$. So we can find two points, namely $(2, f(2) = -\frac{1}{2}((2) - 2.5)^2 + 2(2) + 2 = 23.875)$ and $(6, f(6) = -\frac{1}{2}((6) - 2.5)^2 + 2(6) + 2 = 7.875)$. To compute the slope of the secant line, think back to algebra:

$$\begin{aligned}
\text{Slope} &= \frac{\Delta y}{\Delta x} \\
&= \frac{f(x + \Delta x) - f(x)}{(x + \Delta x) - x} \\
&= \frac{f(6) - f(2)}{6 - 2} \\
&= \frac{7.875 - 5.875}{4} \\
&= \frac{2}{4} \\
&= 0.5
\end{aligned}$$

Now suppose we "move" the point $x + \Delta x$ closer to the point x by lowering the value of Δx. Initially, we set $\Delta x = 4$, suppose we now set this equal to $\Delta x = 2$. Then we get a different secant line (shown in the figure as the red line) which is "closer" to the actual tangent line. If we find the slope of this line, we will get a "better" approximation of the slope of the tangent line:

$$\begin{aligned}
\text{Slope} &= \frac{\Delta y}{\Delta x} \\
&= \frac{f(x+\Delta x) - f(x)}{(x+\Delta x) - x} \\
&= \frac{f(2+2) - f(2)}{(2+2) - 2} \\
&= \frac{f(4) - f(2)}{(4) - 2} \\
&= \frac{[-\frac{1}{2}((4) - 2.5)^2 + 2(4) + 2] - [-\frac{1}{2}((2) - 2.5)^2 + 2(2) + 2]}{(4) - 2} \\
&= \frac{8.875 - 5.875}{2} \\
&= \frac{3}{2} \\
&= 1.5
\end{aligned}$$

Now suppose we "move" the point $x + \Delta x$ even closer to the point x by lowering the value of Δx yet again. Let's now set this equal to $\Delta x = 1$. Then we get a different secant line (shown in the figure as the blue line) which is even "closer" to the actual tangent line than the red or black line. If we find the slope of this line, we will get an even "better" approximation of the slope of the tangent line:

$$\begin{aligned}
\text{Slope} &= \frac{\Delta y}{\Delta x} \\
&= \frac{f(x+\Delta x) - f(x)}{(x+\Delta x) - x} \\
&= \frac{f(2+1) - f(2)}{(2+1) - 2} \\
&= \frac{f(3) - f(2)}{(3) - 2} \\
&= \frac{[-\frac{1}{2}((3) - 2.5)^2 + 2(3) + 2] - [-\frac{1}{2}((2) - 2.5)^2 + 2(2) + 2]}{(3) - 2} \\
&= \frac{7.875 - 5.875}{1} \\
&= \frac{2}{1} \\
&= 2
\end{aligned}$$

Here is the point, as we take $\Delta x \to 0$, we see that the slope of the secant line $\frac{f(x+\Delta x)-f(x)}{(x+\Delta x)-x}$ will eventually "converge" to the slope of the real tangent line. The slope of the real tangent line at x for a function $f(x)$ is called the *derivative* of $f(x)$ at point x. If we are looking for the derivative at point $x = a$, then we denote the slope of this tangent line at $x = a$ as $f'(a)$. Another way to write this is $f'(a) = \frac{df}{dx}\Big|_{x=a}$. Since this number is found by taking the limit of the slope of the secant line as the distance between the points (Δx) goes to 0, we can define: $f'(a) = \frac{df}{dx}\Big|_{x=a} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x)-f(x)}{(x+\Delta x)-x}$.

Definition 2 *The derivative of a function $f(x)$ at a point a is defined as the limit of the slope of the secant line between a point x and a and point x approaches point a. That is, the derivative of a function at a point is defined as*

$$
\begin{aligned}
f'(a) = \frac{df}{dx}\Big|_{x=a} &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)-f(x)}{(x+\Delta x)-x} \\
&= \lim_{\Delta x \to 0} \frac{f(x+\Delta x)-f(x)}{\Delta x}
\end{aligned}
$$

We can illustrate the limit by taking smaller and smaller values of Δx closer to zero of our example function to determine what happens to the slope of the secant line. We illustrate this in the table below for our function $f(x) = -\frac{1}{2}(x-2.5)^2 + 2x + 2$ at $a = 2$:

Δx	$\frac{f(a+\Delta x)-f(a)}{(a+\Delta x)-a}$
4.000	0.5000
2.000	1.5000
1.000	2.0000
0.500	2.2500
0.100	2.4500
0.050	2.4750
0.005	2.4975

We notice that for our derivative, it appears to be approaching 2.5 (that is, the slope of the line that is tangent to $f(x)$ at $x = a$ is 2.5). But can we find the exact value of this limit? Let's try by using the algebra approach to finding limits that we discussed earlier:

$$
\begin{aligned}
f'(2) &= \lim_{\Delta x \to 0} \frac{f(2+\Delta x)-f(2)}{(2+\Delta x)-2} \\
&= \lim_{\Delta x \to 0} \frac{[-\frac{1}{2}(2+\Delta x-2.5)^2 + 2(2+\Delta x)+2] - [-\frac{1}{2}(2-2.5)^2 + 2(2)+2]}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{[-\frac{1}{2}(\Delta x-0.5)^2 + 2\Delta x + 6] - 5.875}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{-.5(\Delta x)^2 + .5(\Delta x) - 0.125 + 2\Delta x + 6 - 5.875}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{-.5(\Delta x)^2 + 2.5\Delta x}{\Delta x} \\
&= \lim_{\Delta x \to 0} -.5(\Delta x) + \lim_{\Delta x \to 0} 2.5 \\
&= 0.5(0) + 2.5 \\
&= 2.5
\end{aligned}
$$

2.3 The Derivative of a Function at Any Point

We saw earlier that the derivative of a function at a point a can be easily found by manipulating our equations and using the limit laws to find the derivative at a given point a. We can also do this, however, on any generic point a. The result is not a number, but rather a function. That is, we can define the *derivative of a function* as such:

Definition 3 *The derivative of a function $f(x)$ at any point x is a function that is defined as:*

$$f'(x) = \frac{df}{dx} = \frac{d}{dx}[f(x)] = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{1}$$

Example 1 *Find $f'(x)$ when $f(x) = -\frac{1}{2}(x-2.5)^2 + 2x + 2$. We can do that by using the definition of the derivative:*

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(-\frac{1}{2}((x+\Delta x)-2.5)^2 + 2(x+\Delta x)+2) - (-\frac{1}{2}(x-2.5)^2 + 2x + 2)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(-\frac{1}{2}((x+\Delta x)-2.5)^2 + 2\Delta x) - (-\frac{1}{2}(x-2.5)^2)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(-\frac{1}{2}((x-2.5)+\Delta x)^2 + 2\Delta x) - (-\frac{1}{2}(x-2.5)^2)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(-\frac{1}{2}(x-2.5)^2 + -\frac{1}{2}(2(x-2.5)\Delta x + \Delta x^2 + 2\Delta x) - (-\frac{1}{2}(x-2.5)^2)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(-\frac{1}{2}(2(x-2.5)\Delta x + \Delta x^2 + 2\Delta x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} (-\frac{1}{2}(2(x-2.5) + \Delta x + 2) \\
&= -\frac{1}{2}(2(\lim_{\Delta x \to 0} x - \lim_{\Delta x \to 0} 2.5) + \lim_{\Delta x \to 0} \Delta x + \lim_{\Delta x \to 0} 2 \\
&= -\frac{1}{2}(2(x-2.5) + 0 + 2 \\
&= -(x-2.5) + 2 \\
&= 4.5 - x
\end{aligned}$$

As we can see, we now have a new function as a result of manipulating $f(x)$ in the way defined by the limit. This becomes useful, since we do not need a specific point to find the derivative. Instead, we can use the approach above to find a new function which we can use to easily compute the derivative of the function at any given point. Earlier we found $f'(2)$. We can see that $f'(2) = 4.5 - x = 4.5 - 2 = 2.5$, which we can see is the same value we found before.

3 Differentiation Rules

As we can see, we do not need to resort to approximation techniques to find derivatives of functions, but rather, we can use our newfound tool of limits to help us compute our derivatives.

However, even this is quite laborious, and so we rarely if ever use the definition of the derivative to compute it. Instead, most of the effort in taking derivatives is often ignored in favor of leveraging simple *derivative rules*. If our function follows a general format, then we can leverage these rules to easily find derivatives without having to resort to finding limits. Let us illustrate one such example.

Suppose we have a function that looks like this: $f(x) = x^n$. Let us see if we can derive a "rule" that we can use to easily find the derivative of any function that looks like this. First, let us use the definition of the derivative:

$$\begin{aligned} f'(x) &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{(x+\Delta x)^n - x^n}{\Delta x} \end{aligned}$$

Now, if you remember from Algebra, we had something called the *binomial theorem*, which stated that given any two variables a and b, we can expand the term $(a+b)^n$ as follows:

$$\begin{aligned} (a+b)^n &= \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i} \\ &= \binom{n}{0} a^0 b^n + \binom{n}{1} a^1 b^{n-1} + \cdots + \binom{n}{n-1} a^{n-1} b^1 + \binom{n}{n} a^n b^0 \\ &= b^n + \binom{n}{1} a^1 b^{n-1} + \cdots + \binom{n}{n-1} a^{n-1} b^1 + a^n \end{aligned}$$

Where $\binom{n}{i} = \frac{n!}{i!(n-i)!}$, and where $n! = n(n-1)(n-2)\ldots(3)(2)(1)$ (Note that $0! = 1$ by definition). So, we can see then that the term $(x+\Delta x)^n$ expands to:
$(x+\Delta x)^n = (\Delta x)^n + \binom{n}{1} x^1 (\Delta x)^{n-1} + \cdots + \binom{n}{n-1} x^{n-1} (\Delta x)^1 + x^n$. Plugging this into our derivation, we have (I have demonstrated the use of multiple notations to represent the derivative, these are only a handful, that is, $\frac{df}{dx}, \frac{d}{dx}[x^n], f'(x)$ are just simply saying "take the derivative of the function with respect to the variable x, where when we say "with respect to x", we are just simply saying that the variable x is the variable that we are changing by taking changes in x, by bringing Δx, very close to 0):

$$
\begin{aligned}
\frac{df}{dx} = \frac{d}{dx}[x^n] = f'(x) &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(x+\Delta x)^n - x^n}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{[(\Delta x)^n + \binom{n}{1} x^1 (\Delta x)^{n-1} + \cdots + \binom{n}{n-1} x^{n-1} (\Delta x)^1 + x^n] - x^n}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(\Delta x)^n + \binom{n}{1} x^1 (\Delta x)^{n-1} + \cdots + \binom{n}{n-1} x^{n-1} (\Delta x)^1 + x^n - x^n}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(\Delta x)^n + \binom{n}{1} x^1 (\Delta x)^{n-1} + \cdots + \binom{n}{n-1} x^{n-1} (\Delta x)^1}{\Delta x} \\
&= \lim_{\Delta x \to 0} (\Delta x)^{n-1} + \binom{n}{1} x^1 (\Delta x)^{n-2} + \cdots + \binom{n}{n-2} x^{n-2} (\Delta x)^1 + \binom{n}{n-1} x^{n-1} \\
&= \lim_{\Delta x \to 0} (\Delta x)^{n-1} + \lim_{\Delta x \to 0} \binom{n}{1} x^1 (\Delta x)^{n-2} + \cdots + \lim_{\Delta x \to 0} \binom{n}{n-2} x^{n-2} (\Delta x)^1 + \lim_{\Delta x \to 0} \binom{n}{n-1} x^{n-1} \\
&= 0 + 0 + \cdots + 0 + \lim_{\Delta x \to 0} \binom{n}{n-1} x^{n-1}
\end{aligned}
$$

$$
\begin{aligned}
&= \lim_{\Delta x \to 0} \frac{n!}{(n-1)!(n-(n-1))!} x^{n-1} \\
&= \lim_{\Delta x \to 0} \frac{n(n-1)!}{(n-1)!} x^{n-1} \\
&= \lim_{\Delta x \to 0} n x^{n-1} \\
&= n x^{n-1}
\end{aligned}
$$

So, if we have a function that looks like $f(x) = x^n$, then we can very easily compute it's derivative by using the equation $f'(x) = nx^{n-1}$. For example, suppose we would like to find the derivative of $f(x) = x$. Then we can use the equation, since $f(x) = x = x^1$. So $n = 1$, and therefore, the derivative would be $f'(x) = nx^(n-1) = (1)x^{(1)-1} = x^0 = 1$. If we wanted to find the derivative of the function $f(x) = x^2$, we would use the same equation but with $n = 2$: $f'(x) = nx^{n-1} = (2)x^{(2)-1} = 2x^1 = 2x$. So, if we wanted to know what the slope of the function x^2 at $x = 5$, we can use the equation of the derivative $f'(x) = 2x$, so that $f'(5) = 2(5) = 10$ (Test this by approximating the limit just as before but with the function $f(x) = x^2$ and $a = 5$)

We have many other rules that we could derive as well. I will leave it as an exercise for you to show the following is true:

Theorem 1 *Let $f(x)$ and $g(x)$ be functions. In addition, let $a \in \mathbb{R}$ be a constant. Then the following is true:*

$$\frac{d}{dx}[af(x)] = a\frac{d}{dx}[f(x)], a \in \mathbb{R} \tag{2}$$

$$\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx}[f(x)] + \frac{d}{dx}[g(x)] \tag{3}$$

$$\frac{d}{dx}[f(x) - g(x)] = \frac{d}{dx}[f(x)] - \frac{d}{dx}[g(x)] \tag{4}$$

$$\frac{d}{dx}[f(x)g(x)] = g(x)\frac{d}{dx}[f(x)] + f(x)\frac{d}{dx}[g(x)] \tag{5}$$

$$\frac{d}{dx}[a] = 0, a \in \mathbb{R} \tag{6}$$

$$\frac{d}{dx}[f(g(x))] = \frac{df}{dg}\frac{dg}{dx} = f'(g(x))g'(x) \tag{7}$$

$$\frac{d}{dx}[x^n] = nx^{n-1} \tag{8}$$

$$\frac{d}{dx}[a^x] = a^x ln(a) \tag{9}$$

$$\frac{d}{dx}[\ln(x)] = \frac{1}{x} \tag{10}$$

$$\frac{d}{dx}[\log_a(x)] = \frac{1}{x\ln(a)} \tag{11}$$

$$\frac{d}{dx}[e^x] = e^x \tag{12}$$

The derivative rules can be used to find the derivatives of functions without having to resort to using the limit definition of the derivative.

Example 2 *Let $h(x) = (3x^3 + 2x^2 - 3x + 5)(2x^2)$. Find $h'(x)$. We know from rule (5) that the derivative of the multiplication of two functions is the sum of the product of the first function and the derivative of the second function, and the second function and the derivative of the first function. If we let $f(x) = 3x^3 + 2x^2 - 3x + 5$ and $g(x) = (2x^2)$, then to find the derivative of $h(x)$, we need to find the derivative of $f(x)$ and $g(x)$. Let's first work on $f(x)$. From rule (3) it can be shown that $\frac{d}{dx}[(3x^3 + 2x^2 - 3x + 5)] = \frac{d}{dx}[3x^3] + \frac{d}{dx}[2x^2] - \frac{d}{dx}[3x] + \frac{d}{dx}[5]$. So let us work on each term individually. First we see that $\frac{d}{dx}[3x^3] = 3\frac{d}{dx}[x^3]$, according to rule (2). We also know that $\frac{d}{dx}[x^3] = 3x^2$ by rule (8). Therefore, the derivative of the first term is: $\frac{d}{dx}[3x^3] = 3(3x^2) = 9x^2$. Following this same logic and rule numbers, we can show that $\frac{d}{dx}[2x^2] = 2(2x) = 4x$ and $\frac{d}{dx}[3x] = 3(1x^0) = 3(1)(1) = 3$. Finally, by rule (6), we have $\frac{d}{dx}[5] = 0$. Putting this all together, we have: $f'(x) = \frac{d}{dx}[(3x^3 + 2x^2 - 3x + 5)] = \frac{d}{dx}[3x^3] + \frac{d}{dx}[2x^2] - \frac{d}{dx}[3x] + \frac{d}{dx}[5] = 9x^2 + 4x - 3 + 0 = 9x^2 + 4x - 3$. Now we need to find $g'(x)$. We see that applying rule (2), we have $g'(x) = \frac{d}{dx}[2x^2] = 2\frac{d}{dx}[x^2]$. Applying rule (8) we have $\frac{d}{dx}[x^2] = 2x$.*

Therefore, we have $g'(x) = \frac{d}{dx}[2x^2] = (2)(2x) = 4x$*. Now we can use rule (7) to find* $h'(x)$*:*

$$\begin{aligned}
h'(x) &= \frac{d}{dx}[(3x^3+2x^2-3x+5)(2x^2)] \\
&= \frac{d}{dx}[f(x)g(x)] \\
&= g(x)f'(x) + f(x)g'(x) \\
&= (2x^2)(9x^2+4x-3) + (3x^3+2x^2-3x+5)(4x) \\
&= 18x^4 + 8x^3 - 6x^2 + 12x^4 + 8x^3 - 12x^2 + 20x \\
&= 30x^4 + 16x^3 - 18x^2 + 20x
\end{aligned}$$

The last rule is called the *chain rule*, and it helps us find the derivative of more complicated types of functions, namely when other functions are "composed" of other function. Suppose that $f(x) = (3x-2)^2$. We see that this function is the composition of two functions, namely $h(y) = y^2$ and $y = g(x) = 3x - 2$. Therefore, we can easily see that $h(g(x)) = (3x-2)^2 = f(x)$. In order to take this derivative, we need to first find $h'(y)$, that is, the derivative of the function y^2 with respect to y. According to our derivative rules (the very first one we discussed), we know that this is equal to $h'(y) = 2y$ (set $n = 2$ and then use the equation to find the derivative). Now we plug back in the y: $h'(y) = h'(g(x)) = 2(g(x)) = 2(3x-2)$. However, we are not done, as we need to find the derivative of $g'(x)$. We see that this function is a sum of terms, so using our rules above, we can break up the derivative of the entire function into smaller derivatives:

$$\begin{aligned}
g'(x) &= \frac{d}{dx}[3x-2] \\
&= \frac{d}{dx}3x - \frac{d}{dx}2 \\
&= 3\frac{d}{dx}x - \frac{d}{dx}2 \\
&= 3(1) - (0) \\
&= 3
\end{aligned}$$

Now we have the expression $h'(g(x)) = 2(3x-2)$ and $g'(x) = 3$, following the equation for the chain rule, we thus have: $f'(x) = \frac{d}{dx}[(3x-2)^2] = h'(g(x))g'(x) = (2(3x-2))(3) = 18x - 12$

Example 3 *Now that we have a tool box for derivatives, let us see if our earlier approximation techniques match with the eaiser to use tools. Recall that we had the function* $f(x) = -\frac{1}{2}(x-2.5)^2 + 2x + 2$*. The derivative of the entire function with by the sum of the derivatives. Hence, we need to find* $\frac{d}{dx}[-\frac{1}{2}(x-2.5)^2]$, $\frac{d}{dx}2x$ + $\frac{d}{dx}2$*. Let us start with the last one. This is easy, and according to rule (13), the derivative of a number is 0. Hence,* $\frac{d}{dx}2 = 0$*. For the second term, we see that we can first apply rule (9) and then our exponent rule discussed just before that (you know, the* $\frac{d}{dx}[x^n] = nx^{n-1}$ *rule).:* $\frac{d}{dx}2x = 2\frac{d}{dx}x = 2(1)x^{1-1} = 2(1)x^0 = 2$*. Therefore, the second term is just 2. Now the tricky part, but it should not be too bad since we have the chain rule on our side. The trick is to break up the function* $-\frac{1}{2}(x-2.5)^2$ *into smaller functions. We notice we can do this with two functions. If we let* $h(y) = -\frac{1}{2}(y)^2$ *and we let* $y = g(x) = x - 2.5$*, then it is easy to see that* $-\frac{1}{2}(x-2.5)^2 = h(y) = h(g(x))$*. Now we have a function that matches our rules, namely* $h(g(x))$*. We know that the derivative of this function is* $h'(y)g'(x)$*., where* $y = g(x)$*. So, let us first find* $h'(y)$*. Again, we see that* $h(y) = \frac{1}{2}y^2$*. We can first apply rule (9) and then* *the power rule with $n = 2$. So we have $\frac{d}{dy}[-\frac{1}{2}y^2] = -\frac{1}{2}\frac{d}{dy}y^2 = -\frac{1}{2}[2y^1] = y$. So, $h'(y) = -y$. Now for $g'(x)$. Following our rules, we find that $\frac{d}{dx}[g(x)] = \frac{d}{dx}[x - 2.5] = \frac{d}{dx}[x] - \frac{d}{dx}2.5] = (1) - (0) = 1$. We now have found that $h'(y) = -y = -g(x) = -(x - 2.5)$ and that $g'(x) = 1$. Plugging these in, we have $\frac{d}{dx}[-\frac{1}{2}(x - 2.5)^2] = \frac{d}{dx}[h(g(x))] = h'(y)g'(x) = -(x - 2.5)(1) = -(x - 2.5)$. So, in sum, we have:*

$$\begin{aligned}
\frac{d}{dx}[f(x)] &= \frac{d}{dx}[-\frac{1}{2}(x-2.5)^2 + 2x + 2] \\
&= \frac{d}{dx}[-\frac{1}{2}(x-2.5)^2] + \frac{d}{dx}[2x] + \frac{d}{dx}[2] \\
&= -(x - 2.5) + 2 + 0 \\
&= -x + 2.5 + 2 \\
&= 4.5 - x
\end{aligned}$$

We now have an equation to use to very easily find the rate of change of the function at any point. Our point from earlier was $x = 2$, plugging this into the equation for the derivative $f'(x) = 4.5 - x$, we see that, $f'(2) = 4.5 - 2 = 2.5$, which matches our approximation approach (except, this approach is much, much easier!)

4 Types of Differentiation

4.1 Implicit Differentiation

Functions are not always defined in an explicit manner, like $y = f(x)$. Sometimes, functions are defined *implicitly*.

Definition 4 *Let x be an independent variable and $f(x)$ a function of x. Then the function $f(x)$ is said to be defined implicitly if the implicit relation $R(x, f(x)) = 0$ defines the function $f(x)$.*

Example 4 *We can define a function between x and $f(x)$ by using the relation $x^2 + [f(x)]^2 = 10$. This equation can define one of two functions: a positive circle and a negative circle.*

If functions are defined implicitly by equations, then we can also find the derivatives of those functions by performing *implicit differentiation*:

Definition 5 *Let $f(x)$ be defined by the equation $R(x, f(x)) = 0$. Then we can find $\frac{df}{dx}$ by finding $\frac{dR}{dx}$ and solving for $\frac{df}{dx}$*

Example 5 *We can find $\frac{df}{dx}$ of the function defined by the equation $x^2 + y^2 = 10$ by taking the derivative of both sides.:*

$$\begin{aligned}
\frac{d}{dx}[x^2 + y^2] &= \frac{d}{dx}[10] \\
\frac{d}{dx}[x^2] + \frac{d}{dx}[y^2] &= \frac{d}{dx}[10] \\
2x + \frac{d}{dx}[y^2] &= 0
\end{aligned}$$

Now, since y is a function of x, we must apply the chain rule to $\frac{d}{dx}[y^2] = 2y\frac{dy}{dx}$. So, we must have:

$$\begin{aligned} 2x + 2y\frac{dy}{dx} &= 0 \\ 2y\frac{dy}{dx} &= -2x \\ \frac{dy}{dx} &= \frac{-x}{y} \end{aligned}$$

4.2 Multiple Derivatives

Recall that when we take the derivative of a function $f(x)$, we obtain a new function, $f'(x)$. Given that $f'(x)$ is itself also a function, we can also take the derivative of this function. Doing so leads us to the *second derivative* of a function.

Definition 6 *The second derivative $f''(x) = \frac{d^2}{dx^x}$ of a function $f(x)$ is the derivative of $f(x)$'s derivative. That is, the second derivative is found by first taking the derivative of $f(x)$ and then finding the derivative of the resulting function.*

Example 6 *Let $f(x) = 3e^{2x}$. We first find the derivative of this function by leveraging the chain rule, and breaking up the function into $y = g(x) = 2x$ and $h(y) = 3e^y$. We can see that leveraging the derivative rules that $h'(y) = 3e^y$ and $g'(x) = 2$. Hence, we have*

$$\begin{aligned} f'(x) &= h'(g(x))g'(x) \\ &= (3e^{2x})(2) \\ &= 6e^{2x} \end{aligned}$$

Now we can find the second derivative of $f(x)$ by finding the derivative of $f'(x)$. Again, we leverage the chain rule. Let $y = g(x) = 2x$ and $h(y) = 6e^y$. Then $g'(x) = 2$ and $h'(y) = 6e^y$. Then:

$$\begin{aligned} \frac{d^2}{dx^2} = f''(x) &= h'(g(x))g'(x) \\ &= (6e^{2x})(2) \\ &= 12e^{2x} \end{aligned}$$

Technically, we do not need to stop at the second derivative, and we will see during our series and sequences lecture that in fact a function's n-th derivative will be very useful to us:

Definition 7 *The n-th derivative of a function $f(x)$ is the resulting function after taking the derivative n times of $f(x)$. We denote this quantity as $f^{(n)}(x) = \frac{d^n f}{dx^n}$*

Example 7 *Taking our example from earlier, we can easily find the n-th derivative of $f(x) = 3e^{2x}$ by recognizing the following pattern:*

n	$f^{(n)}(x)$
1	$(2)(3e^{2x})$
2	$(2)(2)3e^{2x}$
$\vdots$	$\vdots$
n	$2^n 3e^{2x}$

Therefore, the n-th derivative of $f(x)$ is $f^{(n)}(x) = \frac{d^n f}{dx^n} = 2^n 3e^{2x}$

5 Unconstrained Optimization

5.1 Properties of Derivatives and Functions

We saw in an earlier lecture, when we motivated the idea of the concept of the derivative, that the derivative affects the shape of the curve that the function defines. Recall that we said that when $f'(x) = 0$ at a point x, then the tangent line is flat, which indicates that the function may or may not have a maximum or minimum value at x. This type of behavior of functions, based on information about the function's derivative, can be exploited to create a toolbox of methods for finding minimum and maximum values of functions. Especially in business applications, we care about finding the minimum or maximum of a function since most of the time our function is modeling some type of objective as a result of making some type of decision.

5.1.1 Increasing and Decreasing Functions

The type of information that we can obtain from the derivative becomes useful in characterizing what our function looks like. For example we know:

Definition 8 *An increasing function on the interval (a, b) is a function $f(x)$ such that for any $x_1, x_2 \in (a, b)$ where $x_1 < x_2$, then $f(x_1) < f(x_2)$.*

Definition 9 *A decreasing function on the interval (a, b) is a function $f(x)$ such that for any $x_1, x_2 \in (a, b)$ where $x_1 < x_2$, then $f(x_1) > f(x_2)$.*

Theorem 2 *If $f'(x) > 0\ \forall x \in (a, b)$, then $f(x)$ is an increasing function on (a, b).*
Likewise, if $f'(x) < 0\ \forall x \in (a, b)$, then $f(x)$ is a decreasing function on (a, b)

5.1.2 Concavity and Convexity

Functions can be described by how much they are increasing or decreasing. That is, if a function is increasing, the rate at which it increases could change.

Definition 10 *A function is said to be concave on the interval (a, b) if the function's rate of change is itself changing at a decreasing rate.*

Definition 11 *A function is said to be convex on the interval (a, b) if the function's rate of change is itself changing at an increasing rate.*

Observe Figure 3, which is the plot $f(x) = 1000 - (x - 10)^2$. The derivative of this function is $f'(x) = -2(x - 10) = 20 - 2x$. If we take the second derivative, we obtain $f''(x) = -2$. This says that the rate of change of the slope at any point x is itself decreasing. We can observe this in the plot. Near the value of -20, we see that the slope is positive and fair high, namely $f'(-20) = 60$. Notice that as we increase the value of x, the slope decreases, and eventually turns negative:

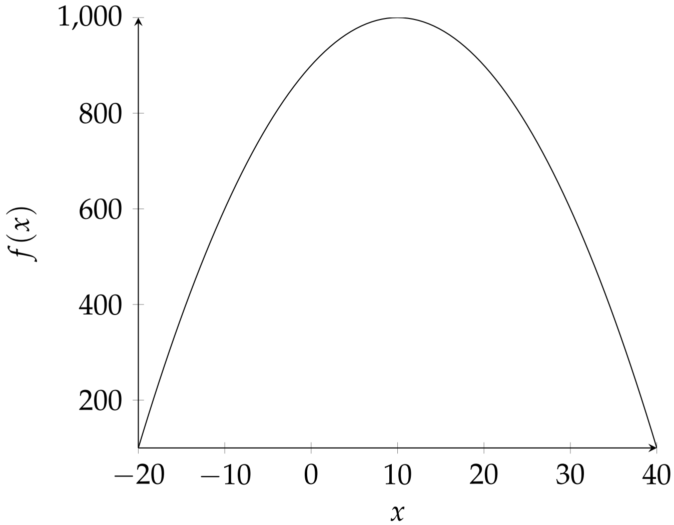

Figure 2: The function $f(x) = 1000 - (x - 10)^2$

x	$f'(x)$
-20	60
-10	40
0	20
10	0
20	-20

We can see that this makes sense, since as we get closer to the value $x = 10$, we notice that the corresponding $f(x)$ values are increasing, but not by as much as before (as evidenced by the values of the derivative at different values of x). These types of functions are referred to as concave, and as we can see, we can easily identify them by computing the second derivative. Likewise, if we observed the opposite, that is, if we observed that the slopes are increasing as we increase the value of x, then these functions are referred to as convex. Again, we can easily identify them by computing the second derivative and identifying if it is positive or negative:

Theorem 3 *(Convexity and Concavity of Functions)*
If $f''(x) < 0$ for all $x \in (a, b)$, then f is a concave function on (a, b).
If $f''(x) > 0$ for all $x \in (a, b)$, then f is a convex function on (a, b)

Sometimes, functions change from being concave to convex, and vice versa. If may first be concave on the interval (a, b), and then change to be convex on the interval (b, c). In such a situation, then we can expect the second derivative to be negative on the interval (a, b) and positive on the interval (b, c). This would imply that the second derivative would be equal to 0 at point b. Such a point is often referred to as an inflection point:

Definition 12 *An inflection point b in an interval (a, c) is a point where a function $f(x)$ turns from convex to concave, or, from concave to convex. If b is an inflection point, then $f''(b) = 0$.*

Example 8 *Consider the function $f(x) = \frac{1}{2}(x-2)^3 + 2$. Looking at the plot of the function, we can see that the function is concave in $(0, 2)$ and convex in $(2, 4)$. If we take the second derivative, we find that*

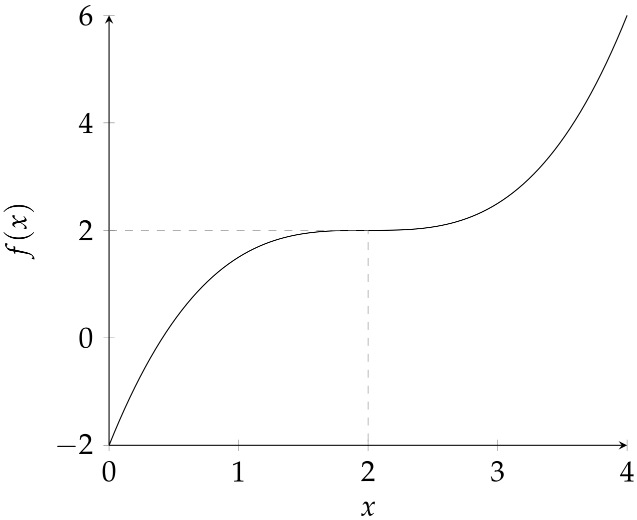

Figure 3: The function $f(x) = \frac{1}{2}(x-2)^3 + 2$

$f''(x) = 3(x-2)$. We notice that in order for $f''(x) < 0$, we have $3(x-2) < 0 \rightarrow (x-2) < 0 \rightarrow x < 2$. Likewise, we can see that in order for $f''(x) > 0$, we have $3(x-2) > 0 \rightarrow (x-2) > 0 \rightarrow x > 2$. Obviously, we can also see that $f''(2) = 0$. Hence, we can clearly see that the function is concave on the interval $(-\infty, 2)$ since $f''(x) < 0$ on the interval. Likewise, we can see that the function is convex on the interval $(2, \infty)$ since $f''(x) > 0$. Hence, $x = 2$ is an inflection point.

5.1.3 Extreme Values

Not only can derivatives be used to explain whether or not functions are increasing or decreasing and to determine the convexity or concavity properties, we can also use derivatives to find extreme values in intervals.

Definition 13 *A function $f : A \rightarrow \mathbb{R}$ is said to have a global maximum at x^* if $f(x^*) \geq f(x)$, $\forall x \in A$. In this instance, we write $x^* = \arg\max_{x \in A} f(x)$. That is, x^* is the value of the independent variable where the maximum value of $f(x)$ lies.*

Definition 14 *A function $f : A \rightarrow \mathbb{R}$ is said to have a global minimum at x^* if $f(x^*) \leq f(x)$, $\forall x \in A$. In this instance, we write $x^* = \arg\min_{x \in A} f(x)$. That is, x^* is the value of the independent variable where the maximum value of $f(x)$ lies.*

Theorem 4 *(Extreme Value Theorem)*
Let $f : A \rightarrow \mathbb{R}$ be continuous on a closed interval $[a, b]$. Then there exists x^*_{min} and x^*_{max} in $[a, b]$ such that $x^*_{min} = \arg\min_{x \in [a,b]} f(x)$ and $x^*_{max} = \arg\max_{x \in [a,b]} f(x)$. The points $f(x^*_{min})$ and $f(x^*_{max})$ are called extreme values.

Another important point (no pun intended) to mention is that in a closed interval, we can find the extreme values by checking the edges of the interval, as well as finding the values within the interval such that $f'(x) = 0$, since:

Theorem 5 *If x^* is an extreme value for a continuous function $f(x)$ on the interval $[a, b]$, then at least one of the following is true:*

1. $x^* = a$
2. $x^* = b$
3. $f'(x^*) - 0$

If $f'(x^*) = 0$, then x^* is called a critical number.

5.2 Unconstrained Optimization

Using the theorems and definitions presented in the previous section, we can formulate a toolbox of methods to use to find the extreme values of functions. As we mentioned previously, this task is of the utmost importance in applying method of calculus to business problems. Typically, when a business makes a decision, it needs to "score" that decision so as to judge if a specific decision is "good" or "bad". Decisions are often modeled as mathematical variables, and commonly are positive (although not always) quantities. A business can judge whether or not a decision is "good" or "bad" based on it's objective. An objective in a business a measurement of the decision. This can be anything from time, distance, cost, profit, revenue, risk, number of locations, etc.

What may be a "good" decision for one firm may be a "bad" decision for another, since the two firms may have differing objectives. For example, if we need to choose between two routes to send our trucks, two firms may consider one route better than the other. One firm may prefer to have their trucks arrive on time but at an extra cost, and the other firm may prefer to have their trucks arrive in such as way that is the lowest cost, despite the possibility of not arriving on time. Hence, when we model business problems, our first step is to understand what the objective of the firm is, and, to understand what decisions will be modeled.

Once we know this, we can construct a mathematical function that will take the decision(s) as input(s) to the function, and the function will "score" the decision that was inputted. Sometimes, the firm wants to decide which input it should select so as to minimize the objective, while other may want to select an input that will maximize the objective. For example, one firm may want to minimize delivery time, while another may want to maximize profit. Once we know what the decisions are, the objective, and how we should select a decision based on minimizing or maximizing the objective, we then can employ the tools of calculus to solve the model. That is, we can leverage calculus to help a firm figure out which is the "best" input into the "objective".

In the real world, decisions have what are called *constraints*. For example, if a firm is trying to decide how much product it should manufacture, we know that this decision is *constrained* to only be positive values or 0. It makes no sense in the real world for the variable "production quantity" to be assigned to values that are negative, and so, we must "constrain" our decision making by optimizing for values that are only positive. When we optimize without considering constraints, then it is said that we are conducting *unconstrained optimization*. On the other hand, when we consider constraints on the input variables, constraints which reflect real world constraints on what can be a possible decision the firm can produce, then it is said that we are conducting *constrained optimization*.

In this section, we will only consider constrained optimization. We will return back to this later in our course once we have completed our lectures on multivariate calculus. For now, let us leverage the tools we developed earlier to formalize a framework for finding optimal decisions when we have no constraints on our variables.

The first tool we will leverage is called the *first order conditions* of a function. Whenever we optimize a function, we always strive to first find the first order conditions. These are mathematical statements that must hold true about our optimal solution. These conditions help us derive relations between variables so that we can *characterize*, and in some cases, solve for, the optimal solution.

Definition 15 *The first order conditions of a function $f(x)$ is the relation:*
$f'(x) = 0$

We must proceed with caution. Recall from the previous section that it is possible for a second derivative to equal 0. If the derivative itself equals 0, it is possible for the second derivative to equal 0. That is, if x^* is a critical point, then this point could be where the extreme values are, or, where an inflection point is. However, recall that inflection points are not necessarily extreme values. We saw this with our example function $f(x) = \frac{1}{2}(x-2)^3 + 2$. We notice that in the interval $\mathbb{R}$, there is no "minimum" or "maximum" that is real-valued. Yet, there is a place where the derivative is equal to 0 (at $x = 2$). Thus, we need a second condition which will help us determine if the critical point is in fact telling us if a minimum or maximum is there.

Definition 16 *The second order conditions of a function $f(x)$ is the relation:*
$f''(x) < 0$ or $f''(x) > 0$, depending on if we are maximizing or minimizing, respectively.

So, we essentially first find the first order conditions, solve for x^*, find the second order conditions, and determine if the value of x^* is indeed a minimum or a maximum.

Example 9 *Optimize the function $f(x) = \frac{1}{2}(x-2)^3 + 2$. Let us first find the first order conditions. We do so by first computing the derivative, which is $f'(x) = \frac{3}{2}(x-2)^2$. Now we state the conditions and solve:*

$$\begin{aligned} \frac{3}{2}(x^* - 2)^2 &= 0 \\ (x^* - 2)^2 &= 0 \\ x^* - 2 &= 0 \\ x^* &= 2 \end{aligned}$$

Now we find the second order condition by first finding the second derivative, which is $f''(x) = 3(x-2)$. The conditions for maximizing are:

$$\begin{aligned} 3(x^* - 2) &< 0 \\ x^* &< 2 \end{aligned}$$

and for minimizing are:

$$\begin{aligned} 3(x^* - 2) &> 0 \\ x^* &> 2 \end{aligned}$$

This is problematic, because our only critical point is at $x = 2$. That is, $x^* = 2$, and we can see that $f''(x^*) = f''(2) = 0$. Therefore, this is an inflection point, and not an extreme value. It is not an extreme value because the second order conditions fail for it being an extreme value (namely that it is strictly less than 2 or strictly greater than 2). Therefore, the only extreme values are $\pm\infty$, respectively.

Example 10 *Suppose that a company experiences a constant weekly demand of λ units per week. If the firm manufactures products, it costs them a fixed cost of K dollars, and a variable cost of c dollars per unit. Additionally, if costs the firm h dollars per unit per week to hold it's inventory. We will assume that when the firm manufactures the product, it instantaneously manufactures all units, and they are immediately moved into inventory. It is assumed that the firm manufactures it's units once it's inventory level reaches 0, at which point, it immediately manufactures and replenishes the inventory. Over time, the inventory linearly depleates, and it can be shown that at time t, the inventory level is $Q - \lambda t$, where Q is the number of units the firm decides to manufacture. It can also be shown that T is the point in time where the inventory level is 0. Last, it can be shown, which we will do in a later lecture, that the average weekly cost of manufacturing and holding Q units is $C(Q) = \frac{K\lambda}{Q} + \frac{hQ}{2} + c\lambda$. Find the first and second order conditions, and find the optimal quantity to order so that it minimizes the average weekly cost.*

First we see that the first order conditions are found by first computing $C'(Q) = \frac{dC}{dQ} = -\frac{K\lambda}{Q^2} + \frac{h}{2}$. The first order conditions are therefore:

$$\begin{aligned}
\frac{dC}{dQ} &= 0 \\
-\frac{K\lambda}{Q^2} + \frac{h}{2} &= 0 \\
\frac{h}{2} &= \frac{K\lambda}{Q^2} \\
\frac{Q^2}{2} &= \frac{K\lambda}{h} \\
Q^2 &= \frac{2K\lambda}{h} \\
Q &= \sqrt{\frac{2K\lambda}{h}}
\end{aligned}$$

Next is to find the second order conditions. Since we are minimizing, the second order conditions are $C''(Q) > 0$. Finding this, we obtain:

$$\begin{aligned}
\frac{d^2C}{dQ^2} &> 0 \\
\frac{d}{dQ}\left[-\frac{K\lambda}{Q^2} + \frac{h}{2}\right] &> 0 \\
\frac{2K\lambda}{Q^3} &> 0
\end{aligned}$$

Since $K > 0$, $\lambda > 0$, and since $Q > 0$, then it must be true that $\frac{2K\lambda}{Q^3}$, and so, our second order conditions are satisfied, since $Q^* = \sqrt{\frac{2K\lambda}{h}} > 0$

We therefore can conclude that the optimal quantity to order is $\arg\min_{Q\geq 0} C(Q) = Q* = \sqrt{\frac{2K\lambda}{h}}$.

6 Antiderivatives

Up to this point, we have discussed the use of derivatives in applications for optimization, as well as characterizing the shape of functions by using the derivatives. What if, however, we wanted to find the opposite? That is, suppose we started with an equation for $f'(x) = \frac{df}{dx}$ and wanted to find the equation for $f(x)$?

Definition 17 *Suppose that $f : A \to \mathbb{R}$ is a function. Then the function $F(x)$ such that $F'(x) = f(x)$ is called the antiderivative of $f(x)$ for all $x \in A$.*

In otherwords, antiderivatives are the opposite of derivatives. Rather than starting with a function and finding the derivative, we are instead starting with a derivative and hope to find the original function. This should be pretty straightforward to find, since all we need to do is work backwards. For example, suppose we have the function $f(x) = x^n$. Can we find a function $F(x)$ such that $F'(x) = f(x) = x^n$? We know from the power rule that $f'(x) = nx^{n-1}$. Therefore, if we work backwards, we should have the function $F(x)$ "look like" once again a power function. That is, it should look something like ax^b. Remember, however, that the derivative of a number is equal to 0, so we need to take this into account. Therefore, the function $F(x)$ may also include some number C, and hence may look something like $F(x) = ax^b + C$. If we take the derivative of this function, we obtain abx^{b-1}. But we know that this must equal x^n by definition of our function $F(x)$. The only way this happens is if $b - 1 = n$, since a, b, n do not involve the term x in them, and so, we must have $b = n + 1$. Plugging this in, we have $a(n+1)x^n = x^n$, which means, we must have $a = \frac{1}{n+1}$. Plugging these in to our "guess" of $F(x) = ax^b + C$, we have $F(x) = \frac{1}{n+1}x^{n+1} + C$. We can of course double check that this is indeed correct:

$$\begin{aligned}
F'(x) &= \frac{d}{dx}\left[\frac{1}{n+1}x^{n+1} + C\right] \\
&= \frac{d}{dx}\left[\frac{1}{n+1}x^{n+1}\right] + \frac{d}{dx}[C] \\
&= \frac{1}{n+1}(n+1)x^{n+1-1} + 0 \\
&= x^n
\end{aligned}$$

If we take a similar approach, we can find antiderivatives for different types of functions:

Theorem 6 *Let $f(x)$ be the functions, respectively, in the first column. Then the antiderivative of these functions are shown in the second column:*

$f(x)$	$F(x)$
x^n	$\frac{1}{n+1}x^{n+1} + C$
$\frac{1}{x}$	$\ln\lvert x\rvert + C$
e^x	$e^x + C$
a^x	$\frac{a^x}{\ln a} + C$