

# Mathematical Review

Carlos Góes  
IFC, World Bank Group

September 28, 2025

## 1 Functions and Notation

**Functions** A function is a mapping from some set  $X$  to some other set  $Y$ . You should think of it as some machine, which we call  $f$ , that transforms some input element  $x$  of the input set  $X$  ( $x \in X$ ) into some output element  $y$  of the output set  $Y$  ( $y \in Y$ ):

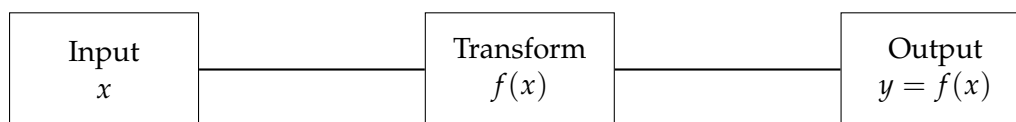
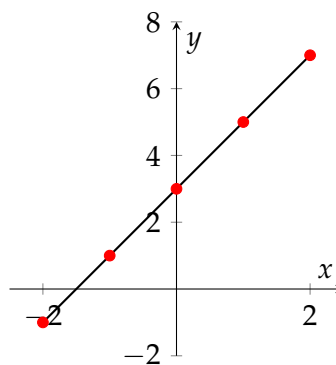


Figure 1: What is a function?

Consider, for instance, the function that represents a line on a plane  $f(x) = 2x + 3$ . As you know from high school mathematics, 3 is the *intercept* of this function (whenever  $x = 0$ ,  $y = 3$ ) whereas 2 is its *slope* (for every unit increase of  $x$ ,  $y$  increases by 2 units). How do we think of this as a mapping? Consider the following numbers from the input set:  $\{-2, -1, 0, 1, 2\}$ . Think of how each of these input  $x$ 's are associated with output  $y$ 's:

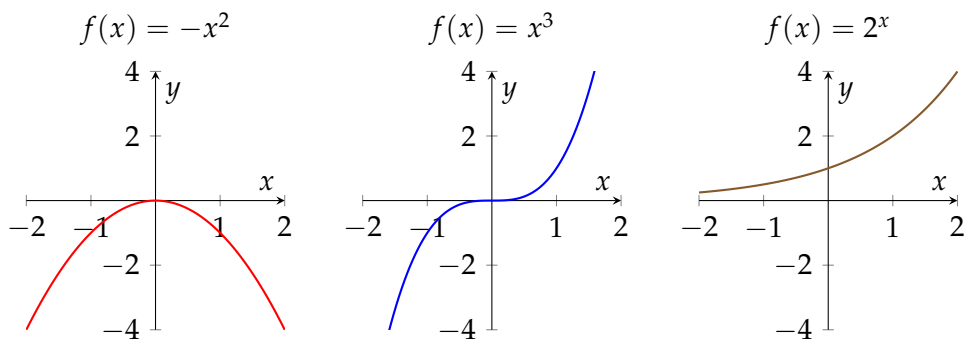
Input: $x$	Transform: $f(x)$	Output: $y = f(x)$
-2	$2(-2) + 3$	-1
-1	$2(-1) + 3$	1
0	$2(0) + 3$	3
1	$2(1) + 3$	5
2	$2(2) + 3$	7



This is what we mean when we say a function is a *mapping*: every  $x$  is associated with a  $y$ . The red dots on the Figure above represents how each  $x$  (in the horizontal axis) is associated with our calculated  $y$ 's. Now imagine that we could calculate  $f(x)$  for every real number between -2 and +2

– or, in math notation  $x \in (-2, 2)$ . As you probably remember from high school, there are infinitely many real numbers in the  $(-2, 2)$  interval, so it is very hard to compute by hand all of these numbers. However, a continuous linear function like  $f(x) = 2x + 3$  can be represented as a line on a plane. The black line on the figure above represents exactly that: it starts where  $x = -2$  and ends where  $x = 2$ . And, for every one of those  $x$ 's, the line maps it to a  $y$ .

Below there are two other examples of functions that you might have seen before. A quadratic function  $f(x) = -x^2$ ; a cubic function  $f(x) = x^3$  and an exponential function  $f(x) = 2^x$ . All of them are represented as curves taking values over the input set  $X = (-2, 2)$  and associating them with specific values in the  $Y = (-4, 4)$  range (our output set). If you wanted to, you could replicate a table like the one we did above to make the *mapping* clear in your mind.



## 2 Derivatives

### 2.1 Intuition

If you have never taken a calculus course, perhaps you have never seen or understood what a derivative is. However, as you will see below, you have already been working with them (by a different name) throughout your student life. Intuitively, a derivative represents *how much the function  $y = f(x)$  changes* whenever  $x$  changes a very small amount or:

$$\text{derivative of } f(x) \text{ with respect to } x = \frac{df(x)}{dx} = \frac{\text{change in } f(x)}{\text{change in } x} = \text{how much } f(x) \text{ changes whenever } x \text{ changes a very small amount}$$

Let  $\Delta x$  denote the *change* in  $x$ . So we are asking how much the value of our function  $f(x)$  changes whenever  $x$  increases to  $x + \Delta x$ . The change in  $x$  is  $(x + \Delta x) - x = \Delta x$ . The change in the value of our output is  $f(x + \Delta x) - f(x)$ . The definition of a derivative is what this value converges to as the change in  $x$  becomes smaller and smaller:  $\Delta x \rightarrow 0$ . Formally:

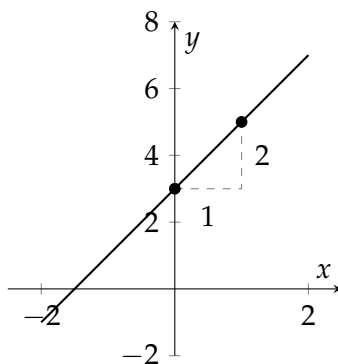
$$\frac{df(x)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

(As an aside, the operator  $\lim_{\Delta x \rightarrow 0}$  just denotes that we are evaluating whatever is on its right-hand side as  $\Delta x$  becomes closer and closer to zero (its *limit*). You don't have to worry about that for this course. However, if you want to go deeper, you can take a look at some of the resources I posted online to understand derivatives a little bit better.)

For simple linear functions such as  $f(x) = 2x + 3$ , we can easily calculate the derivative using the definition above:

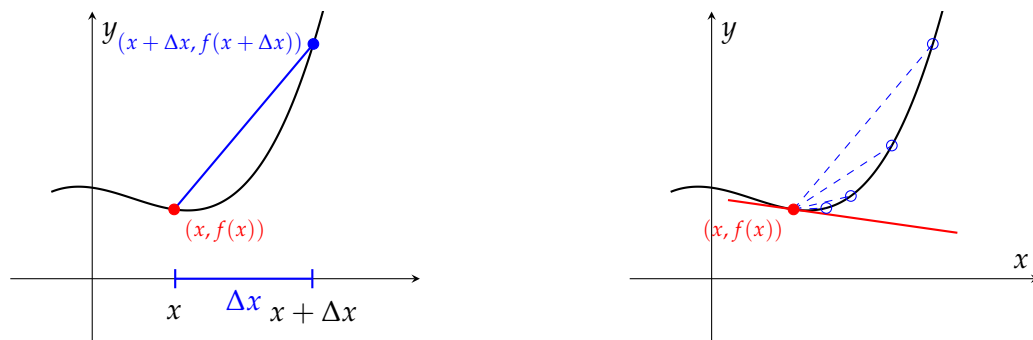
$$\begin{aligned} \frac{df(x)}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{[2(x + \Delta x) + 3] - [2(x) + 3]}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{2\Delta x}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} 2 \\ &= 2 \quad (\text{since 2 is a constant that does not depend on } \Delta x) \end{aligned}$$

For linear functions, then, the derivative of the function with respect to  $x$  is simply the slope of the curve. This is a good moment to test your intuition. If you intuitively understand what derivatives are, this result is not surprising. In linear functions, such as  $f(x) = 2x + 3$ ,  $y$  increases at a *constant rate* as  $x$  changes. Regardless of what  $x$  is, if you increase  $x$  by one unit,  $y$  will increase by two units. The slope of the function around  $x$  denotes exactly this rate of change.



For nonlinear functions, the derivative will be geometrically represented by the tangent of the function (i.e., its slope) around  $x$ . Why is that? Recall that the definition of the derivative is how much the function  $f(x)$  changes as  $x$  changes a tiny amount.

Let us try to illustrate what happens with these changes as  $\Delta x$  becomes smaller and smaller. In the Figure below, we first plot two points. In red,  $(x, f(x))$  and in blue  $(x + \Delta x, f(x + \Delta x))$ . The line connecting those two points represents the average change in  $y$  between  $f(x + \Delta x) - f(x)$  as per unit of change in  $x$ ,  $\Delta x$  (in high school, you might have learned this is a *secant* line).



Now look at the chart on the right and see what happens as we make the change in  $x$  smaller and smaller. First, you can see that the line segments connecting the points are becoming flatter. This shows that the average increase in  $y$  per unit of increase in  $x$  is becoming smaller. Geometrically, you can see that by observing that the slope of the function is changing and becoming flatter as we get closer to the red dot. So much so that, immediately around the red dot, our function is actually decreasing.

Second, we can see that our blue dashed lines seem to be “converging” to the red line. The red line is the *tangent* of the function around the point  $(x, f(x))$ . It represents how much our function increases  $f(x + \Delta x) - f(x)$  as  $\Delta x$  becomes closer and closer to zero. But wait... that is exactly our definition of a derivative. Geometrically, a derivative of a function at point  $x$  is represented by the *tangent* of that function around such a point.

Intuitively, the red line shows what is the rate of change of the function around point  $x$ . It will slope down if the function is decreasing and slope up if the function is increasing at point  $x$ . If the red line were flat, that would mean that  $y$  does not change with  $x$ . One example of this would be a function that is a horizontal line (with  $y$  unchanged for multiple values of  $x$ ).

## 2.2 Rules

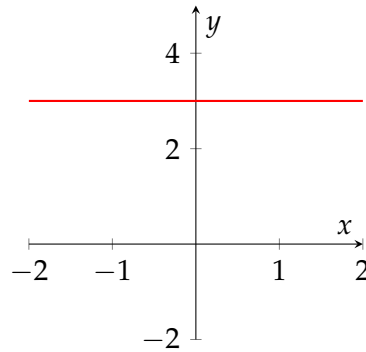
We have now seen what a derivative is. We will now apply some rules for taking derivatives for specific functional forms. This handout will not cover where those rules come from. We would need about half a semester to do that – that is what you do in Calc 1. We will just cover some of the rules and go over examples. We also introduce another notation, in which we express the derivative of single variable functions as:

$$f'(x) = \frac{df(x)}{dx}$$

**Constant** The derivative of a constant  $c$  with respect to some variable is zero:

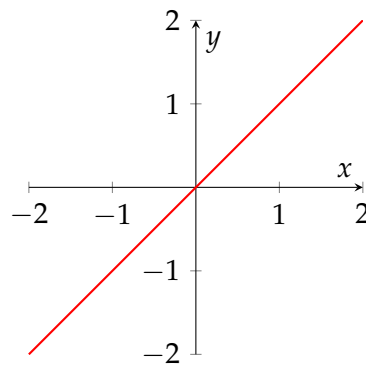
$$\frac{dc}{dx} = 0, \quad \text{for constant } c$$

which should be intuitive. A constant does not vary (otherwise it would not be a constant). Therefore, it cannot vary with some variable  $x$ . For instance, suppose that we express some function  $y = 3$ . How would we plot this? As a horizontal line:



and how much does  $y$  change as  $x$  changes? Well, it doesn't. So  $f'(x) = 0$

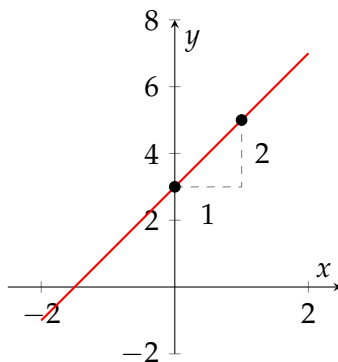
**Variable with respect to itself** How much does  $x$  change as  $x$  changes? Obviously, it changes as a one-to-one proportion. This would be like defining a function  $y = x$ . Therefore:  $f'(x) = dx/dx = 1$ . The chart would simply be a 45-degree line:



**Variable multiplying a constant** Whenever a variable is *multiplying* some constant  $c$ , then its derivative is just the multiplying constant:

$$\frac{d}{dx}(c \times x) = c \times \frac{dx}{dx} = c$$

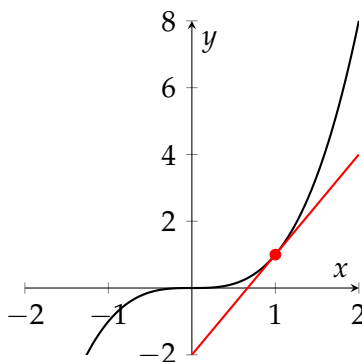
We showed this before, when we talked about the derivative of a linear function  $f(x) = 2x + 3$  being its slope 2 and, therefore,  $f'(x) = 2$ .



**Power Rule** Perhaps the most common derivative you will encounter is that of powers of  $x$ . If you have a function  $f(x) = x^n$ , then its derivative is:

$$\frac{d}{dx}x^n = nx^{n-1}, \quad \text{for integer or real } n.$$

Example: if  $f(x) = x^3$ , then  $f'(x) = 3x^2$ . This tells us that the slope is not constant, but grows with  $x$ .

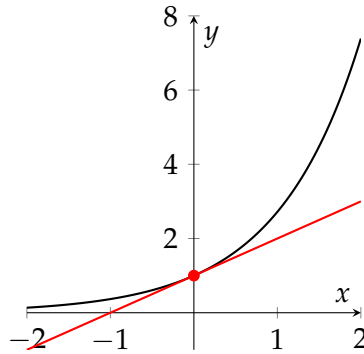


Notice: at  $x = 1$ , slope is  $3(1)^2 = 3$ , which matches the tangent drawn above.

**Exponential Rule** Exponential functions are special because they grow at a rate proportional to themselves:

$$\frac{d}{dx}e^x = e^x.$$

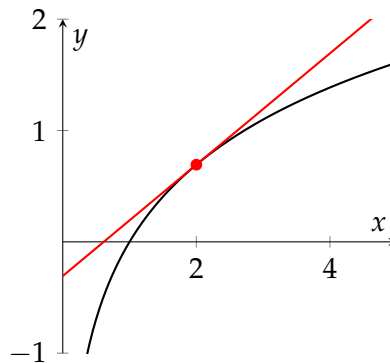
Example: at  $x = 0$ , the slope equals the height of the function (both are 1).



**Logarithmic Rule** The natural log is the inverse of the exponential. Its derivative is:

$$\frac{d}{dx} \ln(x) = \frac{1}{x}, \quad x > 0.$$

This means the slope becomes smaller as  $x$  grows.



**Product Rule** If we multiply two functions, the derivative distributes with a correction term:

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

Example:  $f(x) = x^2$ ,  $g(x) = e^x$ . Then

$$\frac{d}{dx}[x^2 e^x] = 2x e^x + x^2 e^x.$$

**Quotient Rule** If we divide two functions, we use:

$$\frac{d}{dx} \left[ \frac{f(x)}{g(x)} \right] = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}.$$

Example:  $\frac{d}{dx} \left( \frac{x}{x+1} \right) = \frac{(1)(x+1) - x(1)}{(x+1)^2} = \frac{1}{(x+1)^2}.$

**Chain Rule** Finally, if a function is nested inside another, we use:

$$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x).$$

Example:  $y = \ln(x^2)$ . Let  $u = x^2$ . Then:

$$\frac{dy}{du} = \frac{1}{u}, \quad \frac{du}{dx} = 2x, \quad \Rightarrow \quad \frac{dy}{dx} = \frac{1}{x^2} \cdot 2x = \frac{2}{x}.$$

## Summary Table

$\frac{d}{dx}c = 0,$	$\frac{d}{dx}x = 1,$
$\frac{d}{dx}x^n = nx^{n-1},$	$\frac{d}{dx}e^x = e^x,$
$\frac{d}{dx}\ln(x) = \frac{1}{x},$	$\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x),$
$\frac{d}{dx}[cf(x)] = cf'(x),$	$\frac{d}{dx}[fg] = f'(x)g(x) + f(x)g'(x),$
$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f(x)'g(x) - f(x)g'(x)}{g(x)^2},$	$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x).$

## 2.3 Second Derivatives and Concavity

**Second derivatives** Up to this point, we have learned that the first derivative  $f'(x)$  tells us the *slope* of the function at a point — how fast it is going up or down. But sometimes we also want to know how the slope itself is changing. This is what the *second derivative* measures:

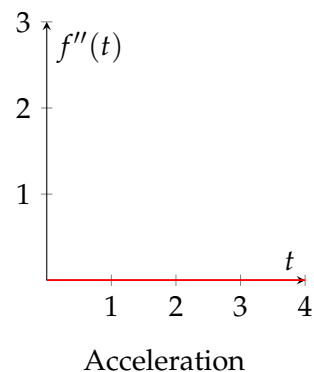
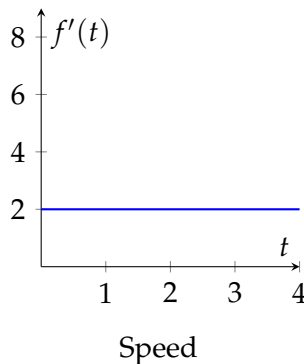
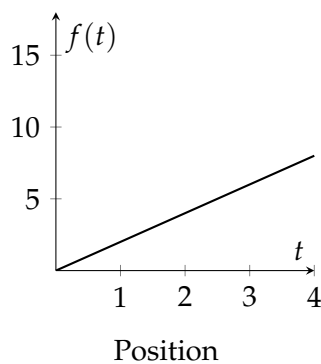
$$f''(x) = \frac{d}{dx}(f'(x)).$$

Suppose, for instance, that  $f(t)$  denotes the position of a car at moment  $t$  (how many miles it has traversed in  $t$  minutes). Then, intuitively:

- $f'(t)$  tells us the speed of the car – how fast it is going in miles/minute or the rate of change of  $f$  with respect to  $t$ .
- $f''(t)$  tells us the acceleration of the car – whether or not its speed is going up or down – or what is the rate of change of the slope.

Now, in particular, let us suppose that the car is moving at a constant rate of 120 miles per hour. In that case  $f(t) = 2t$ , because every minute we move 2 miles. What is the speed of the car. It is a constant rate of  $f'(t) = 2$  miles per minute (or, equivalently, 120 miles per hour). What is its acceleration? Since the car speed is not going up or down (it is constant!) its acceleration is exactly  $f''(t) = \frac{d}{dt}[f'(t)] = 0$



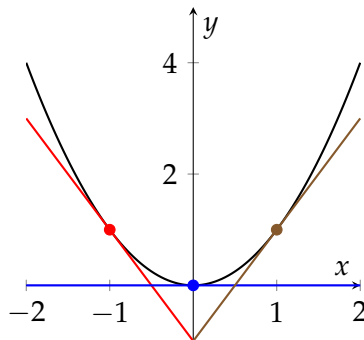


**Concavity.** The sign of the second derivative determines whether the curve is bending upwards or downwards:

- If  $f''(x) > 0$ , then  $f$  is **convex** (curves upward). The slope is increasing.
- If  $f''(x) < 0$ , then  $f$  is **concave** (curves downward). The slope is decreasing.

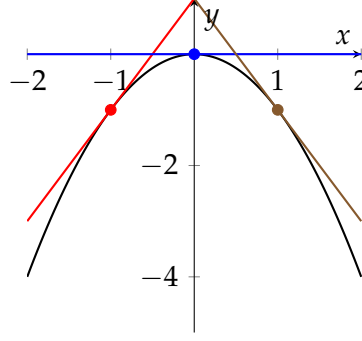
**Examples.**

1. **Convex function:**  $f(x) = x^2$ . Here,  $f'(x) = 2x$  and  $f''(x) = 2 > 0$ . The curve always bends upwards.



Notice how the slope goes from negative (red tangent at  $x = -1$ ), to flat (blue tangent at  $x = 0$ ), to positive (brown tangent at  $x = 1$ ). The slope is *increasing*, so the function curves upward.

2. **Concave function:**  $f(x) = -x^2$ . Here,  $f'(x) = -2x$  and  $f''(x) = -2 < 0$ . The curve always bends downward.



Here, the slope goes from positive (red tangent at  $x = -1$ ), to flat (blue tangent at  $x = 0$ ), to negative (brown tangent at  $x = 1$ ). The slope is *decreasing*, so the function curves downward.

### 3 Optimization

**Utility maximization** Consider the following utility maximization problem. Workers in each country  $i$  inelastically supply their labor and earn labor income  $w_i L_i$ . They have preferences over goods  $p \in \{C, R\}$ . They purchase quantities  $Q_{i,C}, Q_{i,R}$  for prices  $P_{i,C}, P_{i,R}$  and exhaust their labor income, maximizing:

$$\max_{\{Q_{i,C}, Q_{i,R}\}} U_i(Q_{i,C}, Q_{i,R}) \equiv Q_{i,C}^{\alpha_i} Q_{i,R}^{1-\alpha_i} \quad s.t. \quad P_{i,C} Q_{i,C} + P_{i,R} Q_{i,R} = w_i L_i$$

This is a simple constrained concave maximization problem, that you should know how to solve from your calculus classes. One way to solve it is to replace the constraint into the objective function. Note  $Q_{i,R} = \frac{w_i L_i}{P_{i,R}} - \frac{P_{i,C}}{P_{i,R}} Q_{i,C}$ , so the maximization problem below is equivalent to the original one:

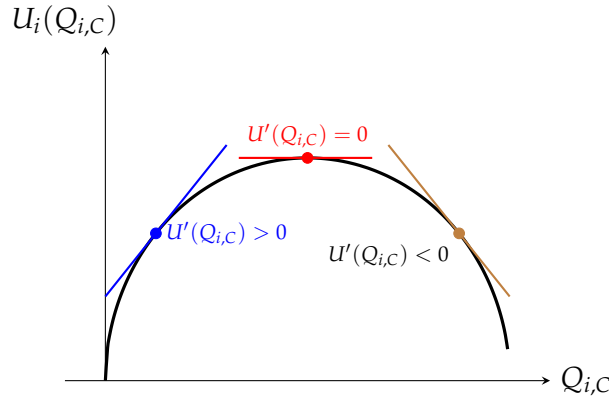
$$\max_{\{Q_{i,C}\}} U_i(Q_{i,C}, Q_{i,R}(Q_{i,C})) \equiv Q_{i,C}^{\alpha_i} \left( \frac{w_i L_i}{P_{i,R}} - \frac{P_{i,C}}{P_{i,R}} Q_{i,C} \right)^{1-\alpha_i}$$

How does this function look like? To simplify our life, let us pick some values for prices, such that  $w_i L_i = P_{i,R} = P_{i,C} = 1$ . (In practice, as we will see in class, we will solve for the prices but our goal here is simply to illustrate how a maximization problem looks like). This problem then becomes:

$$\max_{\{Q_{i,C}\}} U_i(Q_{i,C}) \equiv Q_{i,C}^{\alpha_i} (1 - Q_{i,C})^{1-\alpha_i}$$

If we plot the function  $U_i(Q_{i,C})$  for values of  $Q_{i,C} \in (0, 1)$ , we see a concave (belly shaped, bends downward) function like the one below. From the picture, three things stand out. First, if  $U'(Q_{i,C}) > 0$  at the consumer choice of  $Q_{i,C}$ , they can increase their consumption of computers (decreasing their consumption of roses) to attain higher utility. This is represented by the blue point in the chart below. Second, if  $U'(Q_{i,C}) < 0$  at the consumer choice of  $Q_{i,C}$ , they can decrease their consumption of computers (and increase their consumption of roses) to attain higher utility. This is represented by the

brown point in the chart below. Finally, if  $U'(Q_{i,C}) = 0$  at the consumer choice of  $Q_{i,C}$ , the consumer cannot improve its utility and has achieved the maximum point. This is represented by the dark red point in the chart below.



Therefore, utility is maximized at the point where the derivative of the utility function with respect to the choice variables is zero. How can we find the point. Well, we take the derivative of the utility function, set it equal to zero, and solve for  $Q_{i,C}$ . How would we do that? We need to apply first the product rule then the chain rule<sup>1</sup>.

$$U'_i(Q_{i,C}) = \alpha_i \times Q_{i,C}^{\alpha_i-1} (1 - Q_{i,C})^{1-\alpha_i} + (1 - \alpha_i) \times Q_{i,C}^{\alpha_i} (1 - Q_{i,C})^{-\alpha_i} \times (-1)$$

Setting this equal to zero, and solving for  $Q_{i,C}$  you can show:

$$\begin{aligned} \alpha_i \times Q_{i,C}^{\alpha_i-1} (1 - Q_{i,C})^{1-\alpha_i} + (1 - \alpha_i) \times Q_{i,C}^{\alpha_i} (1 - Q_{i,C})^{-\alpha_i} \times (-1) &= 0 \\ \alpha_i \times Q_{i,C}^{\alpha_i-1} (1 - Q_{i,C})^{1-\alpha_i} &= (1 - \alpha_i) \times Q_{i,C}^{\alpha_i} (1 - Q_{i,C})^{-\alpha_i} \\ \alpha_i (1 - Q_{i,C}) &= (1 - \alpha_i) \times Q_{i,C} \\ \alpha_i - \alpha_i Q_{i,C} &= (1 - \alpha_i) \times Q_{i,C} \\ \alpha_i &= Q_{i,C} \end{aligned}$$

At prices  $w_i L_i = P_{i,R} = P_{i,C} = 1$ , the optimal choice of computers is  $Q_{i,C} = \alpha_i$ . Using the budget constraints, we can derive the optimal choice of roses as:

---

<sup>1</sup>To make that explicit, let us write use the following auxiliary functions:  $f(x) = x^{\alpha_i}$ ,  $g(x) = x^{1-\alpha_i}$ ,  $h(x) = 1 - x$ . Then we can restate this function as:

$$\max_{\{Q_{i,C}\}} U_i(Q_{i,C}) \equiv f(Q_{i,C}) \times g(h(Q_{i,C}))$$

Then:

$$U'_i(Q_{i,C}) = f'(Q_{i,C}) \times g(h(Q_{i,C})) + f(Q_{i,C}) \times g'(h(Q_{i,C})) \times h'(Q_{i,C})$$

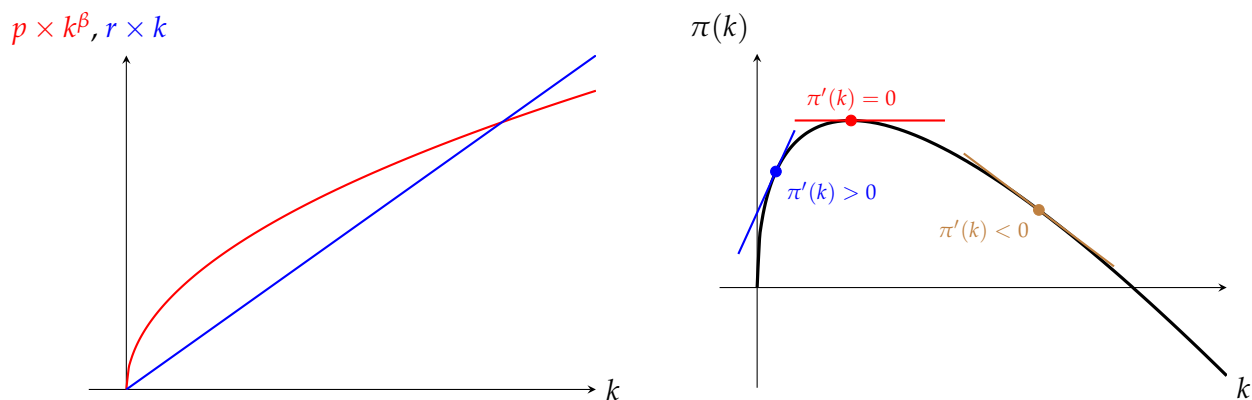
$$Q_{i,R} = \frac{w_i L_i}{P_{i,R}} - \frac{P_{i,C}}{P_{i,R}} Q_{i,C} = 1 - Q_{i,C} = 1 - \alpha_i$$

So, given prices  $w_i L_i = P_{i,R} = P_{i,C} = 1$ , optimal choice of computers and roses, respectively, will be their preference shares  $(\alpha_i, 1 - \alpha_i)$ . In Handout 1, I show you how to solve for this problem with generic prices (and, in fact, how to solve for equilibrium prices). This handout serves as an illustration for why we use derivatives to solve maximization problems.

**Profit maximization** Consider the following profit maximization problem. A firm uses capital  $k$  to produce an output good  $y$  using production function  $y = f(k) = k^\beta$ , for  $0 < \beta < 1$ . They sell their product for price  $p$  in competitive markets, taking prices as given. Each period, they rent their capital at rate  $r$  per unit of  $k$ . They maximize profits according to:

$$\max_k \pi(k) = p \times k^\beta - r \times k$$

Again, to make our lives easier, let us suppose that  $p = r = 1$ . How does this profit maximization function look like? We can actually plot each part of this profit function independently. On the left chart, we plot in red revenues  $p \times k^\beta$  and in blue costs  $r \times k$ . Total profits, which we plot on the right, are the differences between costs and revenues  $\pi(k) = p \times k^\beta - r \times k$ , which is the difference between the two curves on the left. You can see that the difference between the red and blue curves initially increases, reaches a maximum, then starts to decrease, and eventually turns negative – when the blue line lies above the red curve.



We plot that difference on the right-hand side chart. We see that the profit function  $\pi(k)$  is a concave (belly shaped, bends downward) function. From the picture, three things stand out. First, if  $\pi'(k) > 0$  at the firm's choice of  $k$ , they can increase their profits by contracting *more* capital and *increasing* total production. This is represented by the blue point in the chart above. Second, if  $\pi'(k) < 0$  at the firm's choice of  $k$ , they can increase their profits by contracting *less* capital and *decreasing* total production. This is represented by the brown point in the chart above. Finally, if  $\pi'(k) = 0$  at the firm's choice of  $k$ , they cannot increase their profits by changing their use of capital or production and have attained maximum profits. This is represented by the red point in the chart above.