

# Supportive Quantitative Methods

## Table of Contents

- [1. Sets](#)
  - [1.1. Cartesian Products](#)
  - [1.2. Convex Combinations and Convex Sets](#)
- [2. Functions](#)
  - [2.1. Graphs](#)
  - [2.2. One to One and Onto Functions](#)
  - [2.3. Functions of multiple variables](#)
  - [2.4. Linear Functions](#)
  - [2.5. Monotonic Functions](#)
  - [2.6. Inverse Function](#)
- [3. Differentiation](#)
  - [3.1. The Secant](#)
  - [3.2. First Derivative](#)
  - [3.3. Second Derivative](#)
  - [3.4. Derivatives of Higher Order](#)
  - [3.5. Continuous and Smooth Functions](#)
  - [3.6. Taylor Expansion](#)
  - [3.7. Product and Quotient Rule](#)
  - [3.8. Chain Rule](#)
  - [3.9. Partial Derivatives](#)
- [4. Optimization](#)
  - [4.1. Unconstrained Optimization](#)
  - [4.2. Constrained Optimization](#)
- [5. A Hitchhiker's Guide to Mathematics for Economic Courses](#)
- [6. References](#)

## 1. Sets

A **set** is a collection of different things. The things contained in a set are called **elements** or **members**. To denote the membership of  $a$  to a set  $A$  we write  $a \in A$  and read "a belongs in A" or "a is in A". If we want to indicate that  $a$  is not a member of  $A$ , we write  $a \notin A$  and read "a does not belong in A" or "a is not in A". A set without elements is called the empty set, and it is denoted by  $\emptyset$ .

### Examples:

- A collection of four natural numbers

$$A = \{4, 2, 1, 3\}$$

- A collection of colors

$$B = \{\text{blue, white, red}\}$$

- The set of natural numbers

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

- The set of integers

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

- The set of rational numbers

$$\mathbb{Q} = \left\{ \frac{z}{n} : z \in \mathbb{Z}, n \in \mathbb{N} \right\}$$

- The set of real numbers

$\mathbb{R} = \dots$ needs more concepts to be described.

- The set of solutions of the parabola  $2x^2 - 6x + 4 = 0$

$$S = \{x \in \mathbb{R} : 2x^2 - 6x + 4 = 0\} = \{1, 2\}$$

Inclusion operators can be defined for sets. If every element of a set  $A$  is a member of a set  $B$ , we write  $A \subseteq B$  and say that "A is a subset of B". We can also say that "B is a superset of A". Two sets  $A$  and  $B$  are equal if they have exactly the same elements, i.e.,  $A \subseteq B$  and  $B \subseteq A$ . We then write  $A = B$ .

## 1.1. Cartesian Products

The **Cartesian product** of two sets  $A$  and  $B$  is a set with elements pairs combining an element of  $A$  and an element of  $B$ . Specifically, we write

$$A \times B = \{(a, b) : a \in A, b \in B\}.$$

Elements of the Cartesian product are denoted by  $(a, b) \in A \times B$ . Cartesian products are analogously defined of any finite collection of  $n$  sets. For example,

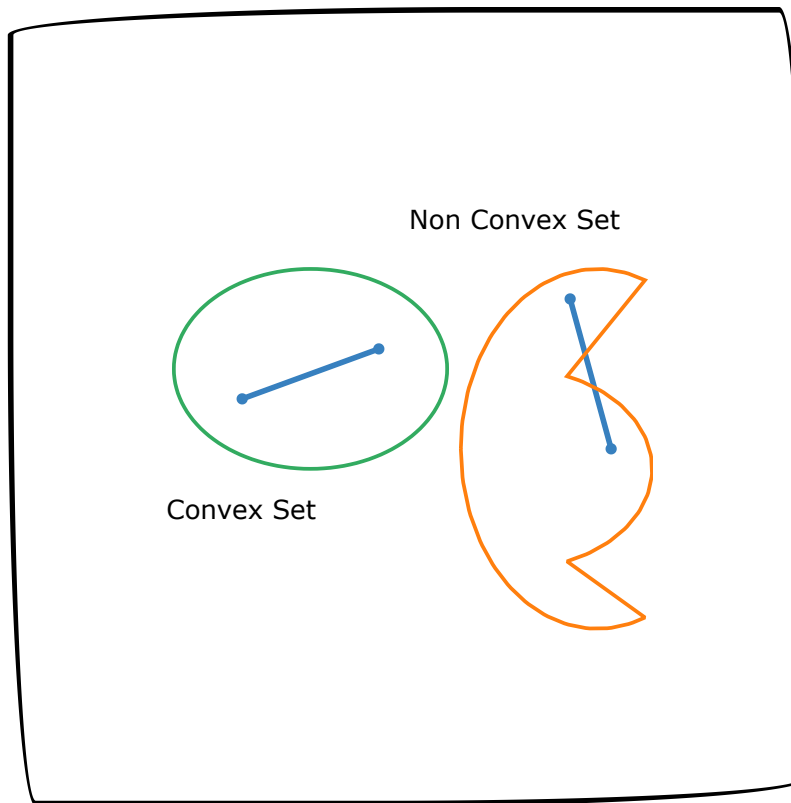
$$\times_{i=1}^n X_i = \{(x_1, \dots, x_n) : x_1 \in X_1, \dots, x_n \in X_n\}.$$

### Example:

Cartesian products are frequently used in economics and finance. For example, consider a situation where one would like to choose the amount of money invested in stocks and bonds. For simplicity, let  $X_1 = \mathbb{R}_{\geq 0}$  be the set of potential stock investments in Euros, and  $X_2 = \mathbb{R}_{\geq 0}$  be the set of potential bond investments in Euros (short sales are excluded). The Cartesian product  $X_1 \times X_2$  is known as the *investment opportunity set* in finance, i.e., the set containing all investment choices available to an economic entity.

## 1.2. Convex Combinations and Convex Sets

It is rare for economic choices to be made in isolation. When a person visits a retail store, she purchases a variety of products instead of a single one on most occasions. In production settings, entrepreneurs combine labor, capital, and other production factors to produce the desired output. In addition, entrepreneurs have to consider not only a particular pair of labor and capital but rather how this pair compares to other feasible pairs of capital and labor they could employ in production. A commonly used way (for reasons going beyond this introduction's scope) to mathematically describe such collections of choices is via their convex combinations. For any real number  $\alpha \in [0, 1]$ , and any two points  $x_1, x_2 \in X$ , we say that  $x = \alpha x_1 + (1 - \alpha)x_2$  is a **convex combination** of  $x_1$  and  $x_2$ .



We say that a set  $X$  is convex if it contains all the convex combinations of its elements. Namely,  $X$  is a **convex set** if for every  $\alpha \in [0, 1]$ , and every  $x_1, x_2 \in X$ , we have  $\alpha x_1 + (1 - \alpha)x_2 \in X$ .

## 2. Functions

A **function** is a rule that maps each element of a set  $X$  to exactly one element of a set  $Y$ . The set  $X$  is called the **domain** of the function, and the set  $Y$  is called the **codomain** of the function. A function that maps  $X$  to  $Y$  is usually denoted by  $f: X \rightarrow Y$ . For each element  $x \in X$ , we write  $f(x)$  and read "f of x" to denote the element of  $Y$  to which  $x$  is mapped. A function with codomain the set of real numbers  $\mathbb{R}$  is called a **real-valued function**. A function having the set of real numbers as its domain is called a **function of a real variable**. It is usual to omit these specializations whenever they are understood from context, and simply refer to functions  $f: \mathbb{R} \rightarrow \mathbb{R}$  as functions instead of real-valued functions of a real variable.

### Examples:

- The *identity function*  $f: \mathbb{R} \rightarrow \mathbb{R}$  maps all real numbers to themselves, i.e.,  $f(x) = x$  for all  $x \in \mathbb{R}$ .
- The *logarithmic function*  $f: \mathbb{R}_{>0} \rightarrow \mathbb{R}$  mapping positive real numbers to their logarithm, i.e.,  $f(x) = \log x$  for all  $x \in \mathbb{R}_{>0}$ .
- The *exponential function*  $f: \mathbb{R} \rightarrow \mathbb{R}$  mapping real numbers to their exponentials, i.e.,  $f(x) = e^x$  for all  $x \in \mathbb{R}$ .

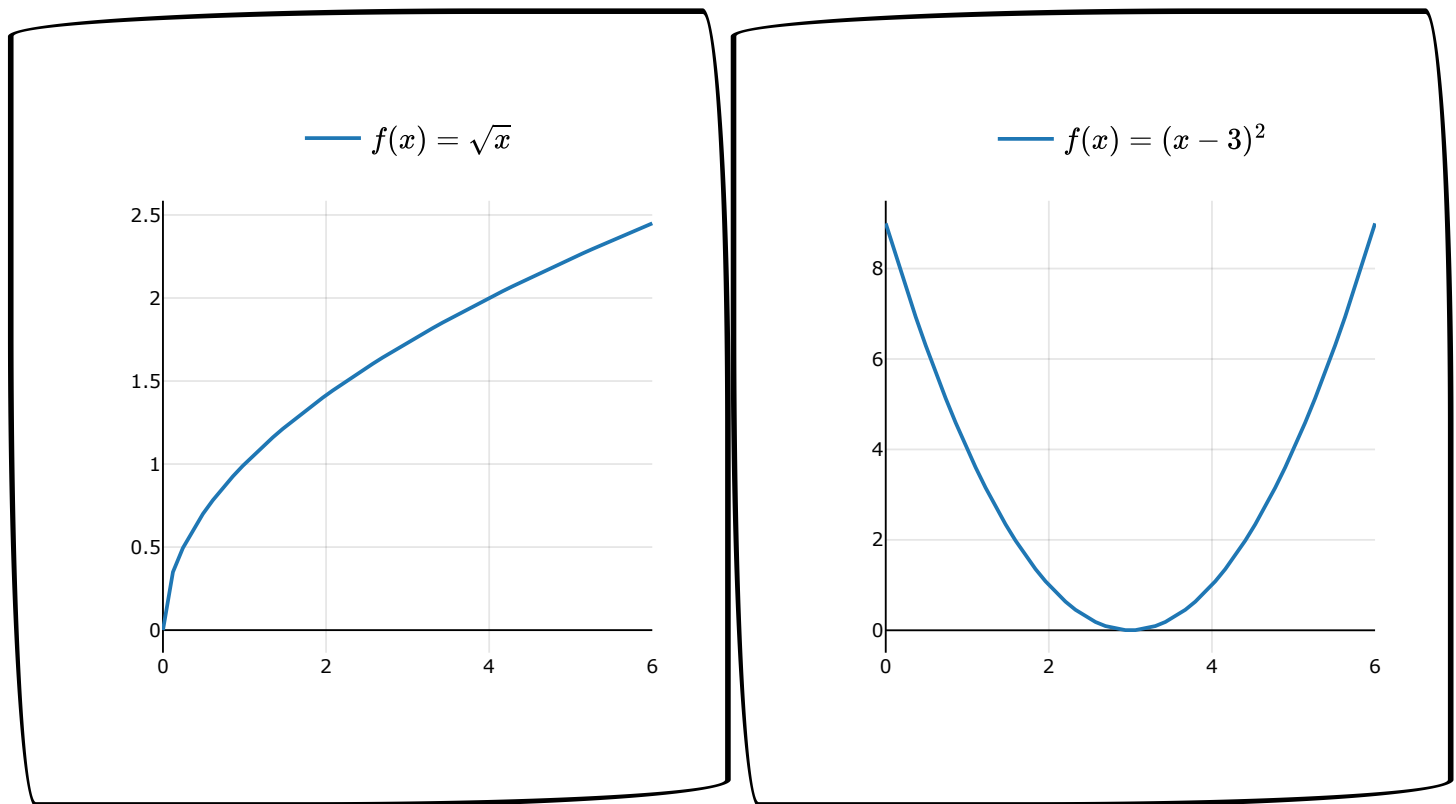
### 2.1. Graphs

The functions' graphs are among the most common and easily communicated ways to represent functions and depict their properties. As a mathematical object, the graph of a function is defined as a set. The **graph** of a function  $f: X \rightarrow Y$  is the subset of the Cartesian product  $X \times Y$  defined by

$$\mathcal{G}(f) = \{(x, f(x)) \in X \times Y: x \in X\}.$$

Graphs take familiar geometric representations of collections of points in the Cartesian coordinates for real-valued functions of one or two real variables.

The graphs of functions of two real variables are also called **surfaces**.



## 2.2. One to One and Onto Functions

Although each  $x \in X$  is uniquely mapped to  $f(x)$ , different elements of  $X$  can be mapped to the same element of  $Y$ . A function that maps each element of  $X$  to a unique element of  $Y$  is called **one to one** or **injection**.

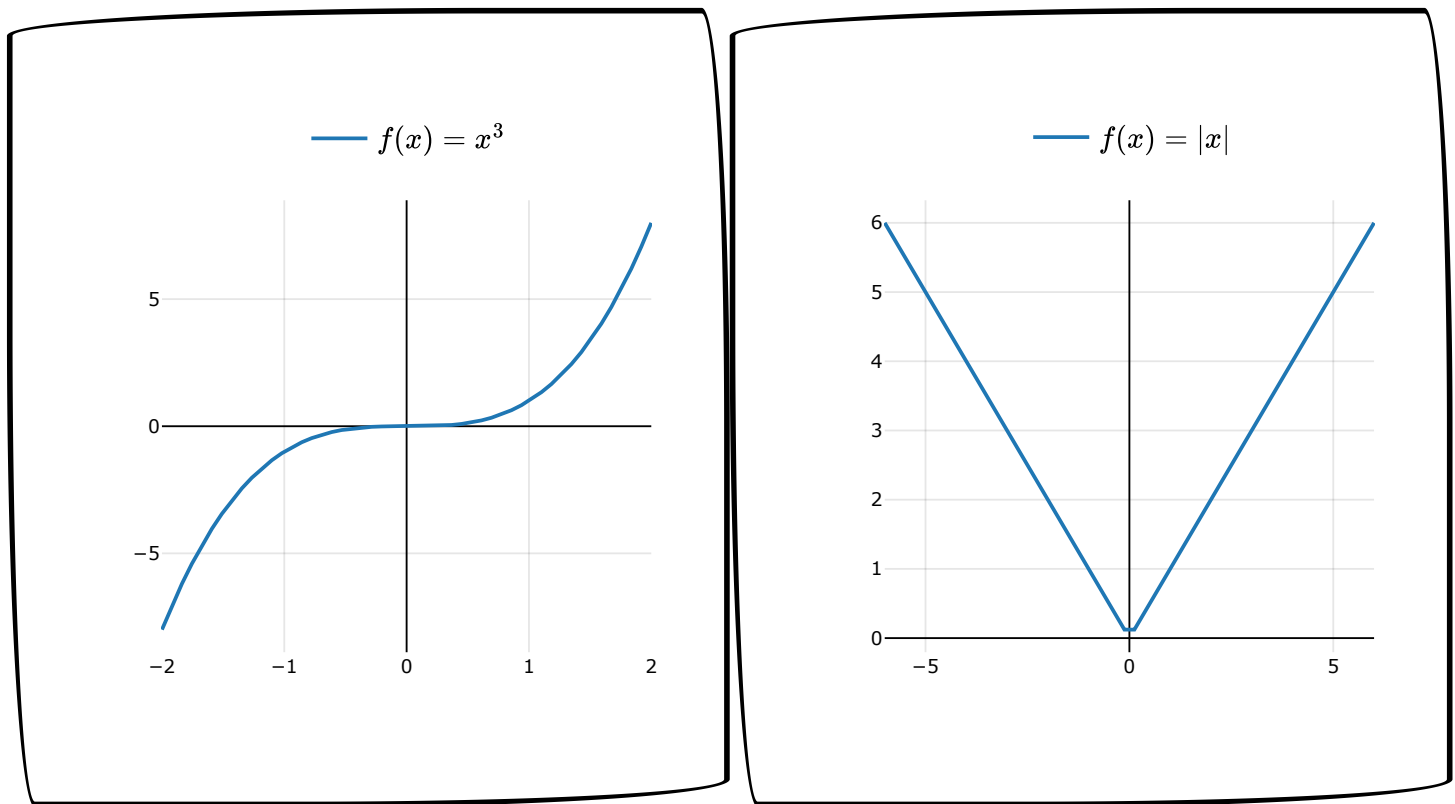
### Examples:

- The identity function  $f(x) = x$  is an injection.
- The constant function  $f: \mathbb{R} \rightarrow \mathbb{R}$  mapping all real numbers to the constant real number 5, i.e.,  $f(x) = 5$  for all  $x \in \mathbb{R}$ , is not an injection.
- The logarithm and exponential functions are injections.

Not all elements of  $Y$  have to be necessarily used by the association rule of the function. The subset of the codomain  $Y$  used by the function's rule is called **range**, and it is denoted by  $f(X)$ . In general, we have  $f(X) \subseteq Y$ , and if the last relation holds with equality, i.e.,  $f(X) = Y$ , we say that  $f$  is **onto** or a **surjection**. A function that is simultaneously an injection and a surjection is called a **bijection**.

### Examples:

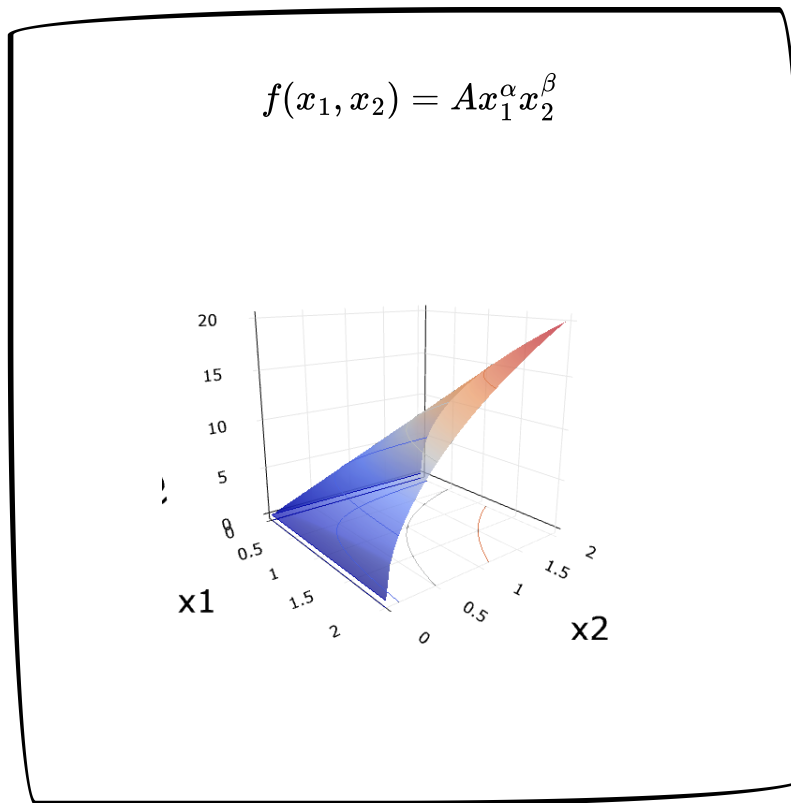
- The identity function  $f(x) = x$  is an injection and a surjection, namely a bijection.
- The constant function  $f(x) = 5$  is neither an injection nor a surjection.
- The  $f: \mathbb{R} \rightarrow \mathbb{R}$  that maps all real numbers to their absolute value (distance from zero), i.e.,  $f(x) = |x|$  for all  $x \in \mathbb{R}$ , is not a surjection.
- The logarithm is a surjection, but the exponential function is not a surjection.



### 2.3. Functions of Multiple Variables

It is fruitful to define rules for associating things from multiple sets to a codomain set on some occasions. For example, sweet recipes use different combinations of essential ingredient quantities to prepare various desserts. Chocolate cakes use chocolate and sugar but no vanilla, while vanilla cakes use sugar and vanilla but not chocolate. If we let  $\mathbf{X}_1$ ,  $\mathbf{X}_2$ , and  $\mathbf{X}_3$  be the set of potential quantities for chocolate, sugar, and vanilla, and  $\mathbf{Y}$  be the set of desserts, we could conceive recipes as functions of the form  $f : \mathbf{X}_1 \times \mathbf{X}_2 \times \mathbf{X}_3 \rightarrow \mathbf{Y}$ .

A mapping  $f: \times_{i=1}^n \mathbf{X}_i \rightarrow \mathbf{Y}$  is called a **function of multiple variables** (in this case  $n$  variables). We write  $f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \in \mathbf{Y}$  to denote the value of the an element  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) \in \times_{i=1}^n \mathbf{X}_i$ . Sometimes, it is more convenient to think of the elements of the Cartesian product  $\times_{i=1}^n \mathbf{X}_i$  as vectors and shorten the notation to  $f(\mathbf{x})$  with the implicit convention that  $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ .



### Example:

The most commonly used function of multiple variables used in economics, finance, and business is the **Cobb-Douglas** function  $f: \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ , which has the form

$$f(x_1, x_2) = Ax_1^\alpha x_2^\beta,$$

where  $A$ ,  $\alpha$ , and  $\beta$  are positive constants. In consumption theory,  $x_1$  and  $x_2$  denote consumption quantities and  $f(x_1, x_2)$  denotes their associated utility. In production theory,  $x_1$  and  $x_2$  denote production factor quantities and  $f(x_1, x_2)$  denotes their associated production output. In macroeconomics,  $x_1$  denotes labor quantities,  $x_2$  capital quantities, and  $f(x_1, x_2)$  the total output of the economy.

## 2.4. Linear Functions

Linear functions are special functions that preserve addition and (vector) multiplication. This property makes linear functions very useful in sciences because they allow us to calculate results of addition and multiplication in the codomain from the corresponding operations in the domain and vice-versa, which can significantly reduce the calculation complexities.

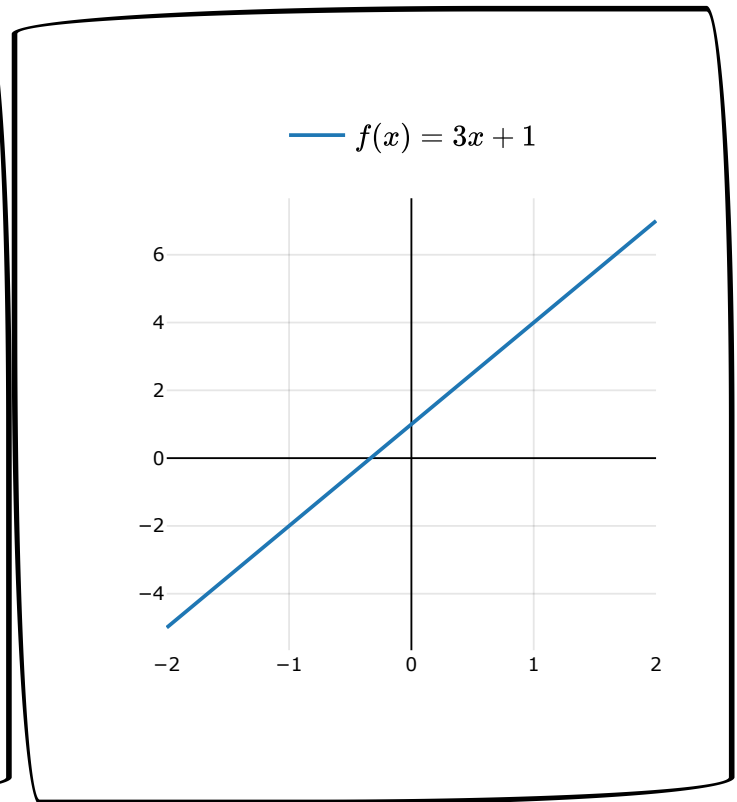
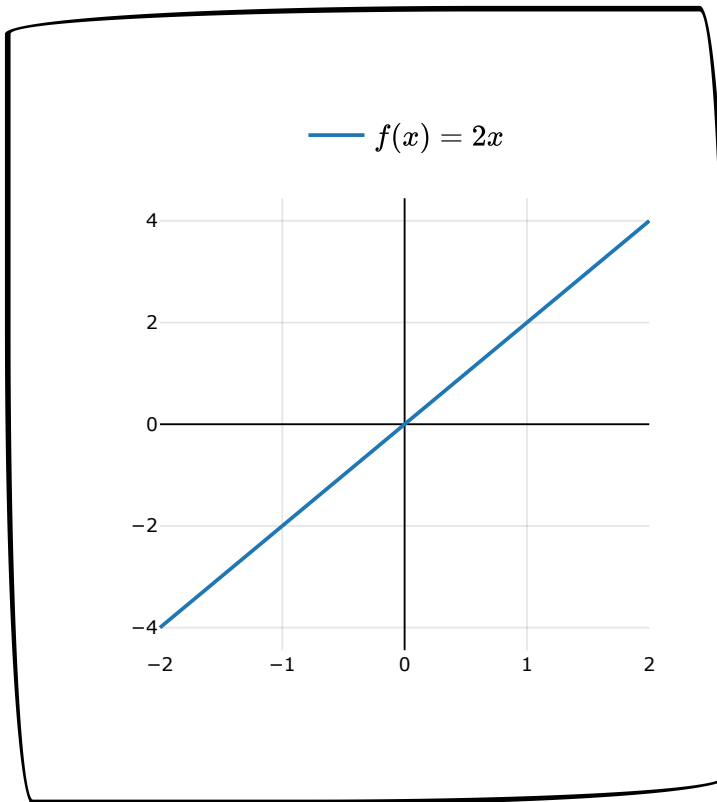
A function  $f: X \rightarrow Y$  is **linear** if (and only if)

$$f(x_1 + \alpha x_2) = f(x_1) + \alpha f(x_2)$$

for all  $x_1, x_2 \in X$  and  $\alpha \in \mathbb{R}$ . Linear functions necessarily satisfy the condition  $f(0) = 0$ . Geometrically, linear functions of a single variable have line graphs passing through the origin of the Cartesian coordinates. Formally, not all lines are linear functions (although this misconception becomes increasingly more mainstream). Functions with line graphs that do not pass through the origin of the Cartesian coordinates are called affine. A function is **affine** if (and only if) it preserves convex combinations, i.e.

$$f(\alpha x_1 + (1 - \alpha)x_2) = \alpha f(x_1) + (1 - \alpha)f(x_2),$$

for all  $x_1, x_2 \in X$  and  $\alpha \in [0, 1]$ .



### Example:

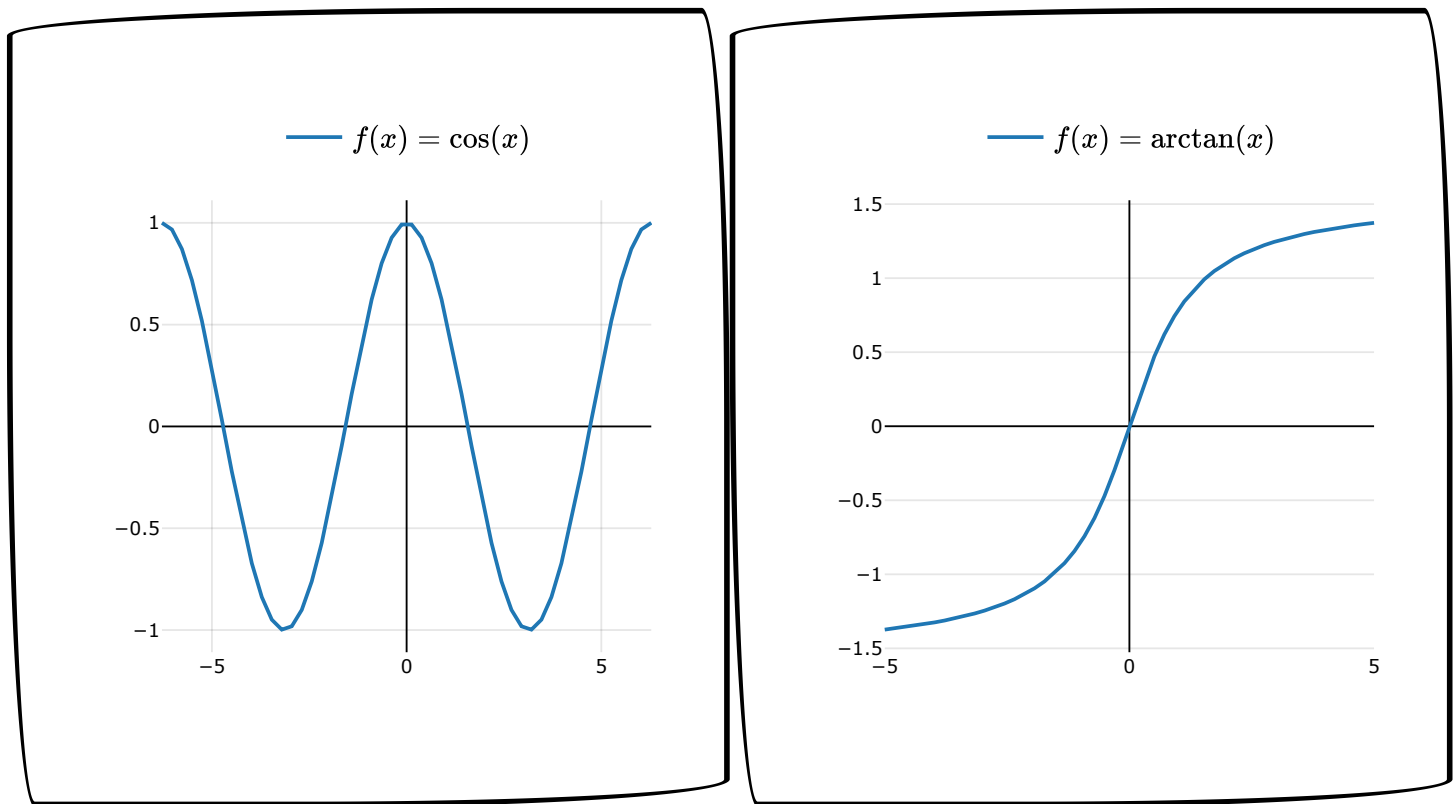
A commonly used production function in general equilibrium finance and macroeconomic models is the "Ak" function postulating that output is analogous to capital. The "Ak" function has the form  $f(x) = Ax$ , where  $A$  is interpreted as the technological constant and  $x$  is the capital used in production. Typically  $k$  is used to denote the input variable instead of  $x$ , which is why this function is known as "Ak". We keep the  $x$  notation here to be consistent with previous examples.

## 2.5. Monotonic Functions

In general, the value of a function can erratically change when it is given different domain values. Monotonic functions constitute a special class of functions with more predictable value changes restricted in particular directions. There are two main types of monotonic functions, namely increasing and decreasing functions. Although each type can be specialized using stricter monotonicity concepts, these two types adequately (for this introduction) describe the basic idea of monotonicity. A real valued function  $f$  is **increasing** if for all  $x_1, x_2 \in \mathbb{R}$  such that  $x_1 \geq x_2$ , we have  $f(x_1) \geq f(x_2)$ . A real valued function  $f$  is said to be **decreasing** if for all  $x_1, x_2 \in \mathbb{R}$  such that  $x_1 \geq x_2$ , we have  $f(x_1) \leq f(x_2)$ .

### Examples:

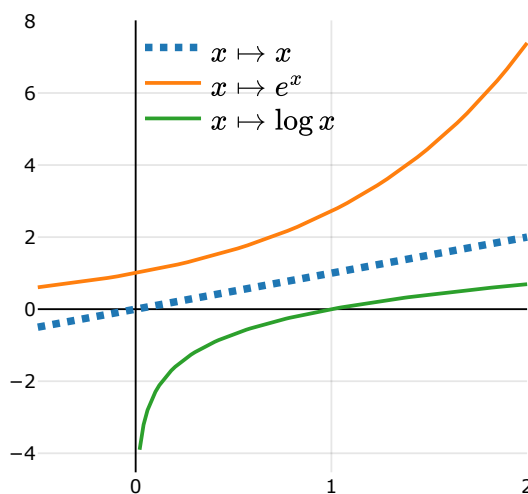
- The identity function  $f(x) = x$  is (strictly) increasing.
- The exponential function  $f(x) = e^x$  is (strictly) increasing.
- The function  $f(x) = e^{-x}$  is (strictly) decreasing.
- The logarithmic function  $f(x) = \log x$  is (strictly) increasing.
- The function  $f(x) = \sqrt{x}$  is (strictly) increasing.
- The function  $f(x) = \frac{1}{\sqrt{x}}$  is (strictly) decreasing.
- The constant function  $f(x) = 5$  is both increasing and decreasing.



## 2.6. Inverse Function

Functions are well-defined rules mapping elements of a set  $\mathbf{X}$  to elements of a set  $\mathbf{Y}$ . What happens, however, if we are interested in examining how elements of  $\mathbf{Y}$  are associated with elements of  $\mathbf{X}$  according to a given function? For example, suppose that  $\mathbf{c}$  is a consumption policy function (typically found in macroeconomics and finance) associating wealth (measured in Euros) with optimal consumption choices (measured in Euros). Thus, for each wealth level  $w$ ,  $\mathbf{c}(w)$  gives the optimal spending allocated to consumption commodities and services. Sometimes it is also relevant to inquire about the required wealth for which a particular consumption level is optimal. Such an inquiry associates consumption spending with wealth levels, which goes in the inverse direction of the association that  $\mathbf{c}$  describes.

## Logarithmic and Exponential Functions



This idea of inversion generalizes in mathematics via the concept of the inverse function. However, function inversion is not always possible. Suppose that we are given a function  $f: X \rightarrow Y$ . To have a well-defined inverse function one has to be able to associate each element of  $Y$  with exactly one element of  $X$ . If for some  $y \in Y$ , there exist two  $x_1, x_2 \in X$  such that  $f(x_1) = y = f(x_2)$  one cannot unambiguously define an association from  $Y$  to  $X$  based on  $f$  (which value of  $X$  should be chosen?  $x_1$  or  $x_2$ ?). Thankfully, we do not encounter such problems if we are given a one-to-one function because such functions guarantee that each  $x \in X$  is mapped to a unique element of  $Y$ . The **inverse function** of a function  $f: X \rightarrow Y$ , if it exists, is a function undoing the operation of  $f$ . Namely, it is a function  $f^{-1}: f(X) \rightarrow X$  such that  $f^{-1}(f(x)) = x$  for all  $x \in X$ .

### Exercises:

1. Find the inverse of the function  $f(x) = 3x$
2. Find the inverse of the function  $f(x) = e^{2x}$
3. Does the function  $f(x) = \sqrt{x}$  have an inverse?

## 3. Differentiation

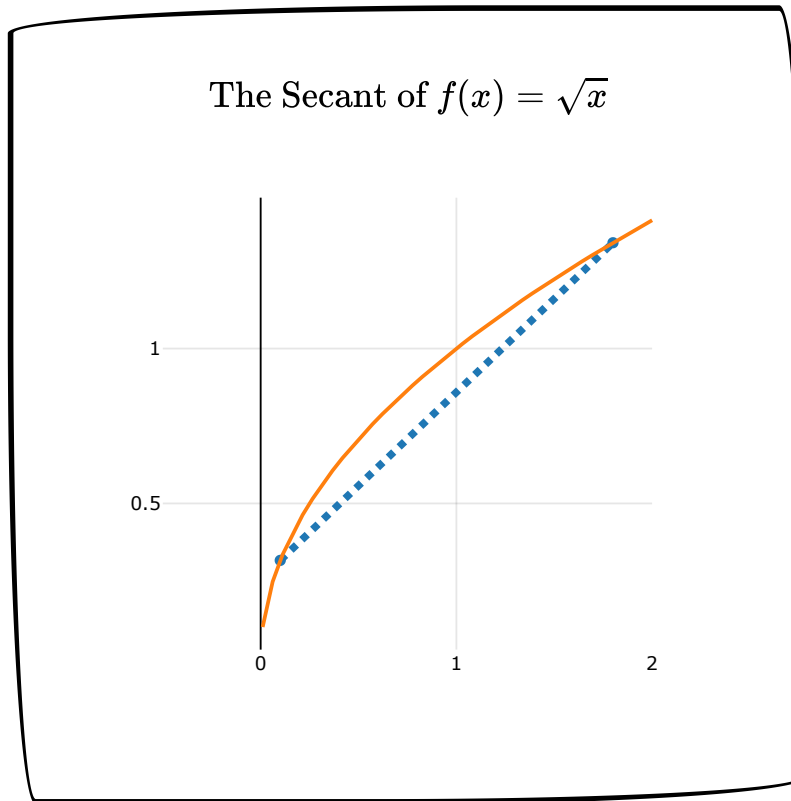
A function's rate of change conveys very useful information about the nature of the rule that associates the domain and the codomain of the function. Suppose, for example, that  $C$  is a function describing the cost of a production process. For each desired output quantity  $q \in \mathbb{R}_{\geq 0}$ ,  $C(q)$  gives the minimum production cost for  $q$  (we call such functions *cost functions* in economics). Knowledge of  $C$  allows one to find the minimum cost for producing an output quantity of, say, 5. What if we want to calculate the incremental cost of "slightly increasing production" by a small amount? We can calculate the additional cost by examining the rate of change of the cost function (this is known in economics as the *marginal cost function*).

So far, so good, but when it comes to the details of calculating the rate of change of  $C$ , one quickly realizes that it is not very clear what is meant by "slightly increasing production". Should we use production increments of one unit? Should we use some other arbitrary increment  $\Delta q = q_2 - q_1$ ? Well, if the cost function  $C$  is affine or constant, then unit increments work fine. As a matter of fact, any arbitrary choice of increment  $\Delta q$  works equally well. However, our luck is exhausted with such simple functions. For functions with non-trivial curvatures, the change that we calculate depends crucially on the length of the increment  $\Delta q$  that we use.

### 3.1. The Secant

A **secant** is a line with endpoints two distinct points of a function's graph. The secant's slope measures the rate of change of a function for a given increment. For a real function  $f$  and two points  $x_1, x_2$  in its domain, the secant's slope is given by the ratio of changes in the range and the domain of the function, i.e.,

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$



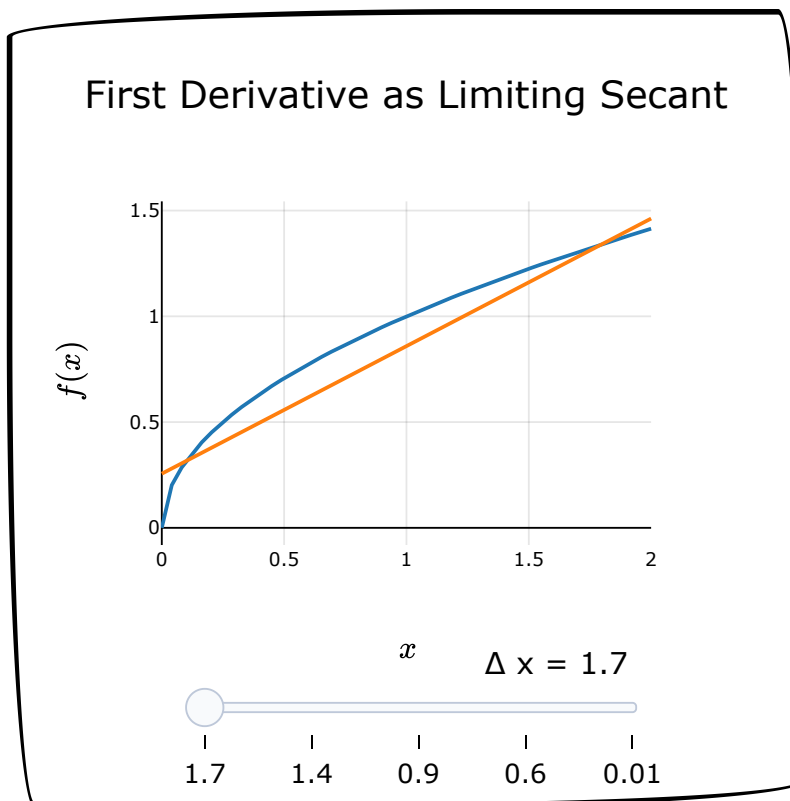
For highly curved functions, the slope of the secant can change significantly as the increment  $\Delta x$  changes.

### 3.2. First Derivative

The first derivative of a function gives us the rate of change of a function by formalizing what is meant by small changes in a way that avoids arbitrary choices of increments. The **first derivative** (or simply derivative if first is understood from context) of a function  $f$ , if exists, is the limit of the secant of  $f$  as the domain change increments go to zero, i.e.,

$$\frac{df(x)}{dx} = \lim_{x_2 \rightarrow x_1} \frac{f(x_2) - f(x_1)}{x_2 - x_1} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

The derivative is a (new) function giving the rate of change of  $f$  for infinitesimal changes in its domain variable. The derivative of  $f$  is often denoted by  $f'$ .



If a function's derivative is negative, the function is decreasing. If it is positive, the function is increasing.

#### Exercise:

The cost of producing  $q$  units of a commodity is given by  $C(q) = q^2 + q + 100$ .

1. Compute the rate of change of  $C$  for a unit increment at  $q = 100$ .
2. Compute the incremental cost  $C(q + 1) - C(q)$  and explain in words its meaning.
3. Compute the marginal cost  $C'(q)$  and  $C'(100)$ . What is the difference between  $C'(q)$  and  $C(q + 1) - C(q)$ ?

### 3.3. Second Derivative

Nothing prevents us from applying the same differentiation process to a first derivative of a function. Suppose that we are given a function  $f$ , for which we can calculate the first derivative  $f'$ . We can then think of  $f'$  as a function, say,  $g = f'$ , and calculate the first derivative of  $g$ . This gives us the derivative  $g'$  of  $g$ , which can be associated back to the original function  $f$  by  $g' = (f')'$ .

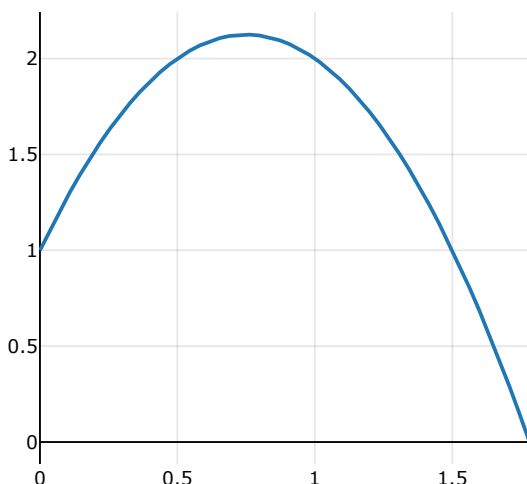
Why would we be interested in something like this? The derivative of the derivative of  $f$  reveals information about the curvature of  $f$ . The curvature of a function can be geometrically thought of as the function's bending. In economics and finance, the curvatures are pivotal in calculating the risk preferences of individuals (see the *relative* and *absolute risk aversion* measures).

The **second derivative** of a function  $f$ , if exists, is the limit of second-order finite differences of  $f$  as the domain change increments go to zero, i.e.,

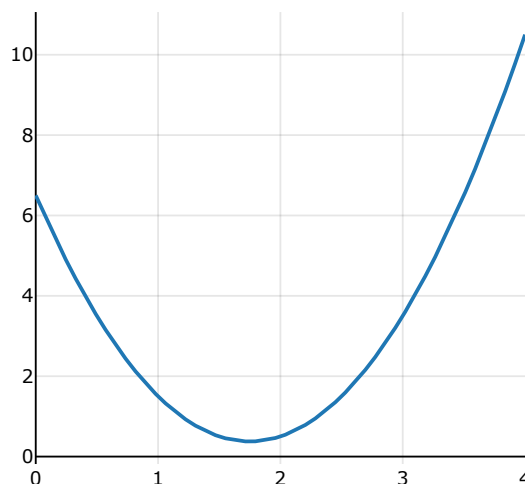
$$\begin{aligned}
 \frac{d^2 f(x)}{dx^2} &= \lim_{\Delta x \rightarrow 0} \frac{f(x_1 + \Delta x) - 2f(x_1) + f(x_1 - \Delta x))}{(\Delta x)^2} \\
 &= \lim_{\Delta x \rightarrow 0} \frac{\frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} - \frac{f(x_1) - f(x_1 - \Delta x)}{\Delta x}}{\Delta x} \\
 &= \lim_{\Delta x \rightarrow 0} \frac{f'(x_1 + \Delta x) - f'(x_1)}{\Delta x}.
 \end{aligned}$$

If a function's second derivative is negative, the function is **concave** (curvature opening down). If its second derivative is positive, it is **convex** (curvature opening up).

A Concave Function



A Convex Function



### Exercise:

Let  $g(x) = 3x^3 - \frac{1}{5}x^5$ .

1. Find  $g'$  and  $g''$ .
2. Check where  $g$  is increasing and where it is concave.

## 3.4. Derivatives of Higher Order

We can recursively continue defining derivatives of higher order. The interpretations of higher-order derivatives become more obscure. A notable exception is the third derivative, which is used to measure the *prudence* of preferences (attitude towards savings for rainy days) in household finance and macroeconomics. Higher-order derivatives are, nevertheless, very useful in economics because they can be used to approximate many functions with arbitrary precision. See the discussion of the [Taylor Expansion Section](#) for an example.

For any  $k \geq 2$ , the  $k$ -th order **derivative** of a function  $f$ , if exists, is the limit of the secant of the  $(k - 1)$ -th derivative of  $f$  as the domain change increments go to zero, i.e.,

$$\frac{d^k f(x)}{dx^k} = \lim_{\Delta x \rightarrow 0} \frac{\frac{d^{k-1} f(x_1 + \Delta x)}{dx^{k-1}} - \frac{d^{k-1} f(x_1)}{dx^{k-1}}}{\Delta x}.$$

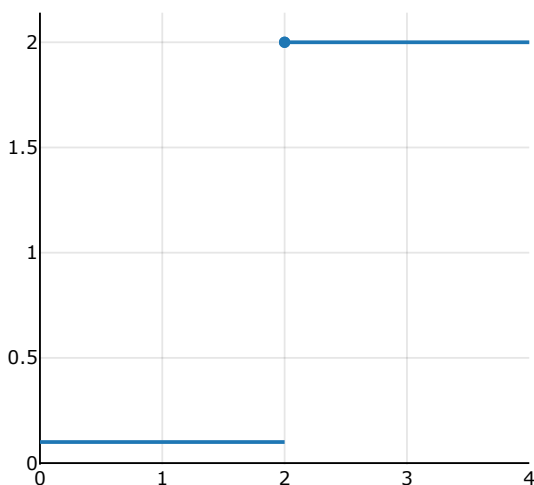
The ***k*-th** order derivative of a function ***f*** is more compactly denoted as ***f*<sup>(*k*)</sup>**. With this notation we can rewrite the definition for the ***k*-th** order derivative as ***f*<sup>(*k*)</sup>(*x*) = (*f*<sup>(*k*−1)</sup>)'(*x*)**.

### 3.5. Continuous and Smooth Functions

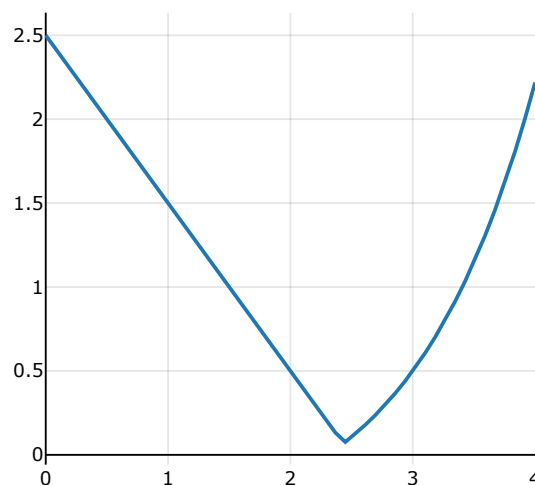
A function is **continuous** if arbitrarily small changes in the domain result in arbitrarily small changes in the range of the function. Giving an exact definition or a good intuition of continuity requires introducing concepts that go way beyond the scope of the material. In business and economic studies, continuity is mostly treated as a technicality that is always present in the used functions. The good news is that whenever a function is differentiable, it is also continuous. Therefore, familiarity with the usual calculus toolbox can serve as a guide for continuity.

A stronger concept (that means more restrictive, i.e., fewer functions are satisfying it) is that of smoothness. We can define smoothness based on the ideas that we have already introduced. A function ***f*** is said to be **smooth** if its derivative ***f*<sup>(*k*)</sup>** exists for any integer ***k* ∈ ℕ**. The graphs of smooth functions do not exhibit any kinks or corners, which is from where these functions are named after.

A Non Continuous Function



A Non Smooth Function



### 3.6. Taylor Expansion

Smooth functions are very useful in economics and finance because many of these functions can be approximated by expressions based on their derivatives. An example of such an approximation is the *Campbell-Shiller decomposition*, which provides a simple way to describe asset returns as functions of prices and dividends in finance.

For a smooth function ***f*** the **Taylor series expansion** of ***f*** at ***x*<sub>0</sub>** is the function given by

$$\begin{aligned}
 T_f(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\
 &= f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2} (x - x_0)^2 + \dots
 \end{aligned}$$

For many well-behaved functions, the Taylor series is convergent and it approximates  $f$ . For such cases, we simply write

$$f(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{f^{(2)}(x_0)}{2} (x - x_0)^2 + \dots$$

### Exercise:

Perform a first-order Taylor approximation to the function  $R(x) = \log(1 + e^x)$  around  $x = d - p$ . This is the approximation used in the Campbell-Shiller decomposition.

## 3.7. Product and Quotient Rule

The **product rule** is used to calculate the derivatives of products of functions. The derivative of the product of two functions  $f$  and  $g$  is given by

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

The **quotient rule** is used to calculate the derivatives of ratios of functions. The derivative of the ratio of  $f$  to  $g$  is given by

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

## 3.8. Chain Rule

A **composite** function is a function that combines the transformations of two functions. Suppose that we are given two functions  $f: X \rightarrow Y$  and  $g: Y \rightarrow Z$ . We can define a composite function  $h: X \rightarrow Z$  by  $h(x) = g(f(x))$ . The composition of  $g$  and  $f$  is sometimes denoted as  $g \circ f$ .

The derivatives of composite functions are calculated according to the **chain rule**. The derivative of the composition of two functions  $f$  to  $g$  is given by

$$(g \circ f)'(x) = g'(f(x))f'(x).$$

### Exercise:

Compute the following derivatives:

1.  $\frac{dZ}{dt}$  when  $Z = (u^2 - 1)^3$  and  $u = t^3$
2.  $\frac{dK}{dt}$  when  $K = \sqrt{L}$  and  $L = 1 + \frac{1}{t}$

## 3.9. Partial Derivatives

The idea of differentiation is not restricted to the functions of one variable. We can calculate the rate of change for functions with more variables by letting one variable vary while keeping all other variables fixed. This concept has many applications in economics, business, and finance because many commonly used functions have more than one variable.

Suppose that we are given a function  $f: X_1 \times X_2 \rightarrow Y$ . The **partial derivative** of  $f$  with respect to the first variable is defined as

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \lim_{\Delta x_1 \rightarrow 0} \frac{f(x_1 + \Delta x_1, x_2) - f(x_1, x_2)}{\Delta x_1}.$$

The partial derivative of  $f$  with respect to the second variable is defined as

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = \lim_{\Delta x_2 \rightarrow 0} \frac{f(x_1, x_2 + \Delta x_2) - f(x_1, x_2)}{\Delta x_2}.$$

Albeit a bit tedious, it is straightforward to generalize the concept for functions of more than two variables. The partial derivative of a function  $f$  of  $k$  variables with respect to the  $j$ -th variable is given by

$$\begin{aligned} \frac{\partial f(x)}{\partial x_j} &= \frac{\partial f(x_1, \dots, x_j, \dots, x_k)}{\partial x_j} \\ &= \lim_{\Delta x_j \rightarrow 0} \frac{f(x_1, \dots, x_j + \Delta x_j, \dots, x_k) - f(x)}{\Delta x_j}. \end{aligned}$$

It is common to denote the partial derivatives using a shorthand notation based on the differentiation variable. In this notation, the partial derivative with respect to the first variable is written as  $f_{x_1}$ , and the partial derivative with respect to the second variable as  $f_{x_2}$ .

### Exercise:

Calculate the partial derivatives of the Cobb-Douglas function

$$u(x_1, x_2) = Ax_1^\alpha x_2^\beta,$$

where  $A$ ,  $\alpha$ , and  $\beta$  are positive constants. Can you, in addition, calculate the second-order partial derivatives?

## 4. Optimization

Decision problems in economics are predominantly described as optimization problems. For example, suppose that one has to decide among a finite number of alternatives  $\alpha_1, \alpha_2, \dots, \alpha_n$ . Alternatives are evaluated based on a payoff function  $u$ . The payoff she receives by choosing alternative  $\alpha_j$  is given by  $u(\alpha_j)$ . One way to model the agent's decision is to assume that she chooses the alternative that maximizes the payoff she gains.

We say that  $\alpha_j$  **maximizes**  $u$  or  $\alpha_j$  is a **maximizer** of  $u$  if  $u(\alpha_j) \geq u(\alpha_i)$  for all  $i = 1, \dots, n$ . Then, we also say that  $u(\alpha_j)$  is the **maximum** of  $u$  or that  $u$  attains its maximum at  $\alpha_j$ . Analogously we can define the minimizer and minimum of  $u$  based on the condition  $u(\alpha_j) \leq u(\alpha_i)$  for all  $i = 1, \dots, n$ .

In our small example, optimization reduces to comparing (sorting) all the values  $u(\alpha_i)$  for  $i = 1, \dots, n$  and picking the greatest or smaller values. Nevertheless, in many business applications, the set of alternatives from which the agent is choosing is either infinite or so large that it is practically impossible to perform all required comparisons for detecting maxima and minima. Luckily, calculus comes to the rescue in such cases.

### 4.1. Unconstrained Optimization

Suppose that we are given a decision problem described by a function  $f(c, p)$ , where  $c$  is a decision variable under the agent's control and  $p$  is a parameter variable that the agent does not affect. The goal of the decision problem is to maximize  $f$ . The **unconstrained optimization problem** corresponding to this decision setting is

$$\max_c f(c, p).$$

The function  $f$  is known as the **objective** and the optimization variable  $c$  is the **control variable** of the optimization problem.

Using Fermat's theorem, we know that if  $f$  attains its maximum at  $c_0$  given  $p$ , then its first (partial) derivative should necessarily be equal to zero, namely  $f'(c_0, p) = 0$ . We can use this idea to locate any local maximizers. By imposing the condition

$$f'(c, p) \stackrel{!}{=} 0,$$

and solving it for  $c$ , we can locate potential local maximizers of the optimization problem. This condition is so commonly used that it is named. It is called the **first-order** or **necessary** condition. The candidate maximizers are called **critical points**.

Thus, we have a way of obtaining candidate maximizers, but how do we know which candidates are actually maximizers? Some of the candidates might as well be minimizers, as those too satisfy the first-order condition. Whether a candidate is a local maximizer or minimizer depends on the (local) curvature of the function. If the function is concave in a small neighborhood around the candidate (curvature opening down), then the candidate is a local maximizer. Instead, if the function is convex in a small neighborhood around the candidate (curvature opening up), then the candidate is a local minimizer.

Therefore, the problem's maximizers are all the candidates that satisfy the condition

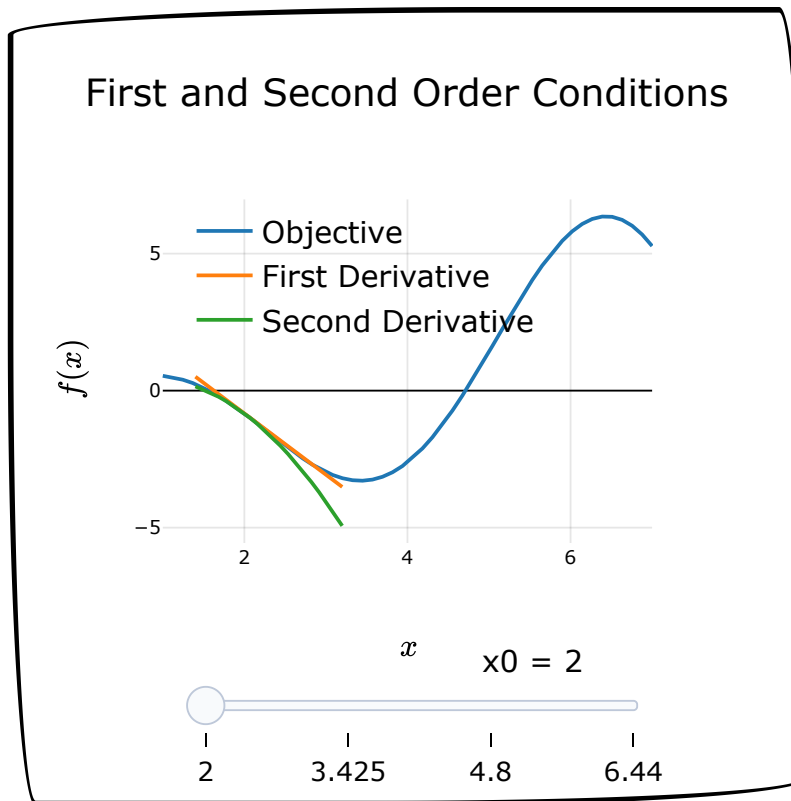
$$f''(c, p) \stackrel{!}{<} 0.$$

This condition is called the **second-order** or **sufficient** condition.

We have kept the parameter value  $p$  fixed in the preceding discussion. Thus, any maximizer that we have obtained depends on the fixed value of  $p$ . We can change the value of  $p$  to  $p'$  and repeat the process. The new maximizer will be based on the new parameter value  $p'$  in this case. In this manner, we can define a function such that for each parameter value  $p$ , we get an optimizer  $c(p)$ . This mapping is called the **optimal control function**. Substituting the optimal control in the objective function we get

$$v(p) = f(c(p), p) = \max_c f(c, p).$$

In this way, we have eliminated the dependence of the objective on the control variable and expressed the maximized values exclusively as a function of the parameter  $p$ . The function  $v$  gives the maximum value of  $f$  for each parameter  $p$ , and it is called the **value function**.

**Exercise:**

Consider the function  $f(c, p) = p - (c - p)^2$  for  $c, p \in \mathbb{R}$ .

1. Calculate the maximization candidates using the first-order condition for a fixed value of  $p$ .
2. Show that these candidates are indeed maximizers using the second-order condition.
3. What is the optimal control function of this problem?
4. What is the value function of this problem?

**Exercise:**

Let  $f$  be defined by  $f(x) = 2 - \frac{4x}{x^2+3}$

1. Find  $f'$  and  $f''$ .
2. Find all the critical points and classify them into maxima and minima.

**4.2. Constrained Optimization**

In some decision problems, agents do not have access to the full universe of alternatives. For instance, firms are subject to technological constraints, and they can only choose to produce technologically feasible quantities. Similarly, households can only buy commodities and services that are affordable for their budgets. How can we describe such decision problems?

Suppose that we are given a decision problem described again by a function  $f(c, p)$ , where  $c$  is the control variable and  $p$  is a parameter. The goal of the decision problem is to maximize  $f$ . However, the agent cannot freely choose  $c$ . The agent's choices have to satisfy the **constraint**  $g(c, I) = 0$ , where  $I$  is another parameter that the agent does not control. The **constrained optimization problem** corresponding to this decision setting is

$$\begin{aligned} & \max_c f(c, p) \\ \text{s.t. } & g(c, I) = 0. \end{aligned}$$

The general approach with which one solves constraint optimization problems is by forming the **Lagrangian** function

$$\mathcal{L}(c, \lambda, p, I) = f(c, p) - \lambda g(c, I).$$

The Lagrangian introduces an additional variable  $\lambda$ , which is called the **Lagrange multiplier**. The Lagrange multiplier takes many useful interpretations in economics and finance decision problems. Specifically, the Lagrange multiplier gives us the *shadow price* of the constraint, i.e., the change in the value function when the constraint is infinitesimally relaxed.

Under some conditions, the maximizers or minimizers of the constrained maximization problem are obtained by solving the unconstrained optimization problem with  $\mathcal{L}(c, \lambda, p, I)$  as the objective and  $c, \lambda$  as control variables. Therefore, we can use the Lagrangian to reduce a constrained optimization problem to the relatively simpler unconstrained problem

$$\max_{c, \lambda} \mathcal{L}(c, \lambda, p, I),$$

We can get the maximizers or minimizers of the original constrained problem using the unconstrained problem's first and second-order conditions.

### Exercise:

Maximize the area of a rectangle with length  $\alpha$  and width  $\beta$  such that its perimeter is equal to **16cm**.

## 5. A Hitchhiker's Guide to Mathematics for Economic Courses

We can only cover so much in a short preparatory course. This topic proposes a guide to help you navigate through your future economic studies.

One introductory book for the mathematics you will deal with in the economics course is ([Sydsæter, Hammond, and Strøm 2012](#)). I am using the 4th edition of the book.

The following points do not constitute homework you *have* to do. Keep in mind that it will take time to cover them, and you cannot do them in a week. Keep these points more as a companion for later on when you read the corresponding sections.

- Chapters 1, 2: Avoid until you need something in particular from there, or you do it for fun (very low priority)
- From chapter 3, we cover section 3.6. An additional section that you could read is 3.4. This section does not directly apply to what we do, but it will help you understand many arguments you will find in other textbooks.
- Chapter 4, we mostly cover. See sections
  - 4.5: Examples 3, 4 and exercises 2, 3, 4, 6 (relevant to the [Economics of the Market](#) course)
  - 4.6: Examples 1, 2 (relevant to the [Economics of the Market](#) course) and exercises 3 (use the first order conditions and conclude how easier it is from what the exercise suggests), 5, 6
  - 4.7 - 4.10: Use only as a reference if you want to refresh your knowledge about particular functions.
- Chapter 5: Skip for now. Read only for fun or if you need to review our discussion about the inverse function.
- Chapter 6: We cover most of the ideas. Use for review. See sections
  - 6.2: Exercises 1, 2, 3, 4, 7 (only for practice)
  - 6.4: Exercises 1, 2, 3, 6, 7 (nice that these exercises have economic context)
  - 6.6-6.8: Use as a reference if you forget something.
- Chapter 7: A bit more advanced. Skip for now.
  - 7.7 Elasticities: We do not cover this, but it is quite important for all the economics courses. Prioritize its reading if you can.
- Chapter 8: More mathematical details than our approach in the class. Read with caution.
  - 8.3: Examples 1, 2, 4 and exercises 2, 4, 5 (these are very relevant to the [Economics of the Market](#))

- Chapter 13: Quite more mathematical than our approach. Avoid it for now because you might get lost. Potential exceptions are
  - 13.1: Examples 3,4, and exercise 3. These are helpful for Macro courses.
- Chapter 14: We cover some elements together.
  - Read 14.1: Useful for microeconomics 1

## 6. References

Sydsæter, Knut, Peter J. Hammond, and Arne Strøm. 2012. *Essential Mathematics for Economic Analysis*. 4th ed. Harlow ; Singapore: Pearson.

---

Date: Summer Term 2024

Author: Pantelis Karapanagiotis

Email: [karapanagiotis\[at\]ebs\[dot\]edu](mailto:karapanagiotis[at]ebs[dot]edu)

[Validate](#)