Mathematics Preparatory Course - MSc in EEBL

Lecture Notes

Filippo Maurici

Disclaimer: These notes are for exclusive use of the students of the Mathematics Preparatory Course, MSc in European Economy and Business Law, University of Rome Tor Vergata. Their aim is purely instructional and they are not for circulation.

General Information

- **Instructor**: Filippo Maurici
- **Email**: filippo.maurici@uniroma2.it
- **References**: The main reference is "Mathematics for Economists" by C. Simon and L. Blume, Norton & Company
- **Objective:** The Mathematics preparatory course aims to review the basic concepts of calculus and linear algebra and to provide students with the necessary tools to understand the notions of Economics where a quantitative approach is needed.
- **Course content**: The course consists in five lectures which will cover the key concepts of the following subjects:
  - One variable calculus
  - Linear Algebra
  - Function of several variable
  - Optimization

1 Preliminary notions

1.1 Set theory

A **set** is a well-defined collection of distinct objects, considered as an object in its own right. In order to characterize a set we can directly specify the elements that compose it

Example 1.1. $A = \{a, b, c\}$

or it can be identified listing the features of the elements in the set

Example 1.2. $A = \{n | (\text{n is odd}) \wedge (\text{n is positive})\}$

Once we have the defined what a set is, and we have seen how to characterize it, we can define the notion of **subset**

Definition 1.1. Let A and B be two distinct sets. If every element in A is also in B we can state that A is a subset of B, denoted by $A \subseteq B$. More formally:

$$\forall a \in A : a \in A \Rightarrow a \in B \Leftrightarrow A \subseteq B \tag{1}$$

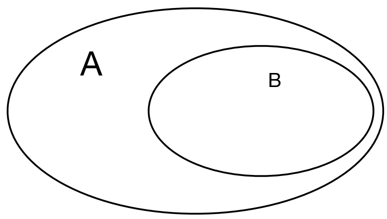


Figure 1: Example of $B \subseteq A$ via Euler diagram

If $A \subseteq B$ and $B \subseteq A$, then A and B coincide $\Rightarrow A = B$.

It is called **empty set** the unique set having no elements, $\varnothing$, and, conventionally, given every set A, the empty set is subset of A

Definition 1.2.

$$\varnothing \subseteq A \;\; \forall A \tag{2}$$

1.2 Set operations

Let X be a set, and A and B subsets of X.

Definition 1.3. **Union**, denoted by $A \cup B$, is the set containing the elements of A *or* the elements of B

$$A \cup B = \{x \in X | (x \in A) \vee (x \in B)\} \quad (3)$$

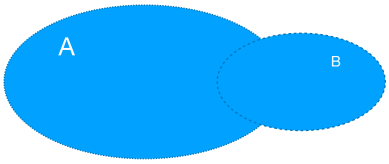


Figure 2: Example of $A \cup B$ via Euler diagram

Example 1.3. Let $A = \{a, b, g\}$ and $B = \{a, c, g\} \Rightarrow A \cup B = \{a, b, c, g\}$

Definition 1.4. **Intersection**, denoted by $A \cap B$, is the set containing the elements either in A and in B

$$A \cap B = \{x \in X | (x \in A) \wedge (x \in B)\} \quad (4)$$

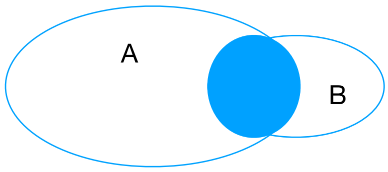


Figure 3: Example of $A \cap B$ via Euler diagram

Example 1.4. Let $A = \{a, b, g\}$ and $B = \{a, c, g\} \Rightarrow A \cap B = \{a, g\}$

Moreover,

Definition 1.5. Two sets are **disjoint** if

$$A \cap B = \varnothing \quad (5)$$

Definition 1.6. **Difference**, denoted by $A \backslash B$, it is given by the set composed by the elements of A minus the elements which are also in $B \Rightarrow A \backslash B = A - (A \cap B)$

Example 1.5. Let $A = \{a, b, g\}$ and $B = \{a, c, g\} \Rightarrow A \backslash B = \{b\}$

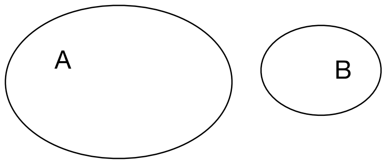


Figure 4: Example of two disjoint sets via Euler diagram

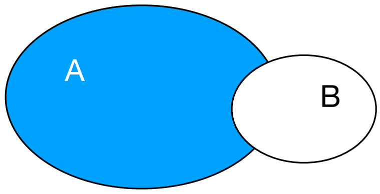


Figure 5: Example of $A \backslash B$ via Euler diagram

1.3 Sets of numbers

Throughout this course we will use the following sets of numbers:

- Set of **Natural numbers**: $\mathbb{N} = \{0, 1, 2, 4, ...\}$
- Set of **Integers**: $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 4, ...\}$
- Set of **Rational numbers**: $\mathbb{Q} = \{a/b \mid a, b \in \mathbf{Z}, b \neq 0\}$
- Set of **Real number** $\mathbb{R}$. This set includes all rational numbers, together with all irrational numbers.

1.4 Cartesian product

The **Cartesian product** of two sets A and B, denoted $A \times B$, is the set of all the ordered pairs (a, b) where $a \in A$ and $b \in B$.

Example 1.6. Let $A = \{0, 1\}$ and $B = \{0, 2\}$, we have that

$$A \times B = \{(0, 0), (0, 2), (1, 0), (1, 2)\} \tag{6}$$

while

$$B \times A = \{(0, 0), (0, 1), (2, 0), (2, 1)\} \tag{7}$$

2 One Variable Calculus

2.1 Introduction

One of the main goals of Economics is to understand mechanisms, interactions and relationships between different variables. Mathematically these relations are described by **functions**.

Definition 2.1. Let X and Y be two subsets of $\mathbb{R}$. A function f defined on X with values in Y is a correspondence associating to each element $x \in X$ **at most** one element $y \in Y$.

$$f : x \to f(x) \tag{8}$$

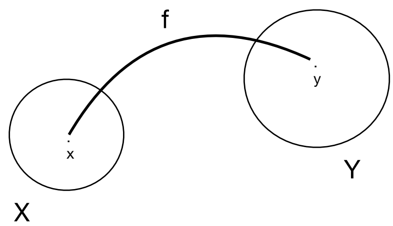


Figure 6: Graphic representation of (8)

the variable $x \in X$ is called **independent**, while the variable $y \in Y$ is called **dependent**.

Roughly speaking a function will take an input (x) and through some rules will give you back an output (y).

Example 2.1. $f(x) = x^2 + 3 \Rightarrow f(5) = 5^2 + 3 = 28$

The set of elements $x \in X$ to which f assigns an element in Y is called the **domain**, while the elements $y \in Y$ associated to x are called **images**. The set of all the images is called **range**.

Useful rules:

Function	*Domain*
$\ln g(x)$	$g(x) > 0$
$\sqrt[2n]{g(x)}$	$g(x) \geq 0$
$\sqrt[2n+1]{g(x)}$	$\forall x \in \mathbb{R}$
$\dfrac{k(x)}{g(x)}$	$g(x) \neq 0$

Example 2.2. Let $f(x) = \frac{1}{x-2} \Rightarrow$ the domain of f is given by all the elements in $\mathbb{R}$ except $+2$

Some of the most important information of a function are contained in the **graph**.

The graph of f is the subset of $\Gamma(f)$ of the Cartesian product $X \times Y$ made of pairs $(x, f(x))$ when x varies in the domain of f, i.e.

$$\Gamma(f) = \{(x, f(x)) \in X \times Y : x \in \operatorname{dom} f\} \tag{9}$$

A function may be even or odd.

Let be the function

$$f : x \to Y \tag{10}$$

Suppose that if $x \in X$ it is also true that $-x \in X$ (It means that the function is symmetric with respect to the origin) then

Definition 2.2. The function f is **even** if

$$f(-x) = f(x),\ \forall x \in X \tag{11}$$

Example 2.3. $f(x) = x^2 - 3$, since $(-x)^2 = x^2$, then $f(x) = f(-x)$

Definition 2.3. The function f is **odd** if

$$f(-x) = -f(x),\ \forall x \in X \tag{12}$$

Example 2.4. $f(x) = 3x^3$, since $(-x)^3 = -(x^3)$ then $f(-x) = -f(x)$

The basic geometric properties of a function are whether it is increasing or decreasing and the location of its local/global maxima/minima.

Definition 2.4. The function f is increasing if

$$\forall x_1, x_2 \in dom(f) \quad x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2) \tag{13}$$

Conversely, it is decreasing if

$$\forall x_1, x_2 \in dom(f) \quad x_1 > x_2 \Rightarrow f(x_1) \geq f(x_2) \tag{14}$$

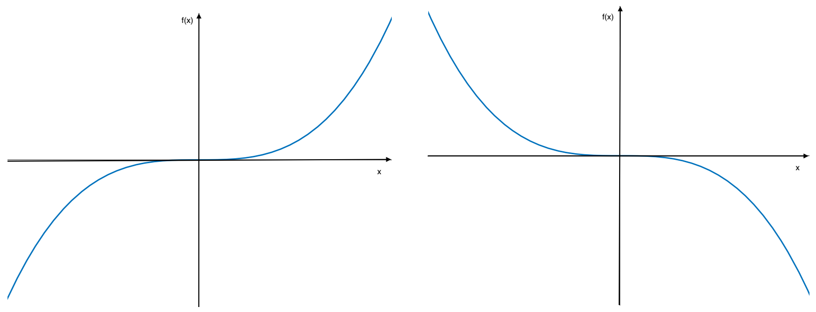


Figure 7: Example of increasing function - Example of decreasing function

The point where the function turns from decreasing to increasing is a minimum for the function and the point where the function turns from decreasing to increasing is a maximum for the function. If there is no greater (smaller) value of the function in its range from that maximum (minimum) then the maximum (minimum) is called global maximum (minimum).

2.2 Function types

- A **Polynomial** is a map of the form $P(x) = \underbrace{a_n x^n}_{monomial} + ... + a_1 x + a_0$ where n is the degree of polynomial.
- A rational function is of the kind $R(x) = \frac{P(x)}{Q(x)}$ where P and Q are polynomials.

- An **exponential function** is defined $f(x) = a^x$, the domain it is all $\mathbb{R}$ and it satisfies $y(0) = a^0 = 1$.

- The inverse $y = \log_a x$ is called logarithm and it is defined in $(0, +\infty)$.

2.3 Linear Function

Polynomials of degree 1 are interesting functions, they are also called **Linear function**

$$f(x) = ax + b \tag{15}$$

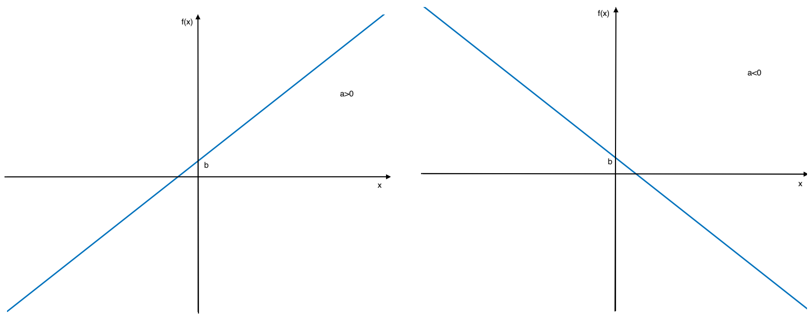


Figure 8: Example of linear functions

the graph of a linear function is a straight line. In order to draw this function knowing two points in the Cartesian plan is enough.

One of the main features which distinguishes two different lines is the **slope** (steepness) that is given by a. This function is increasing as $a > 0$ and decreasing if $a < 0$; if $a = 0$ the function degenerates to the constant function $f(x) = b$. The slope is given by the ratio of the growth in y $(y_2 - y_1)$ and the growth of x $(x_2 - x_1)$, that is $a = \frac{y_2 - y_1}{x_2 - x_1}$.

2.4 Quadratic Function

A quadratic function is a polynomial of degree 2. The graph of a univariate quadratic function is a parabola whose axis of symmetry is parallel to the y-axis

$$f(x) = ax^2 + bx + c \tag{16}$$

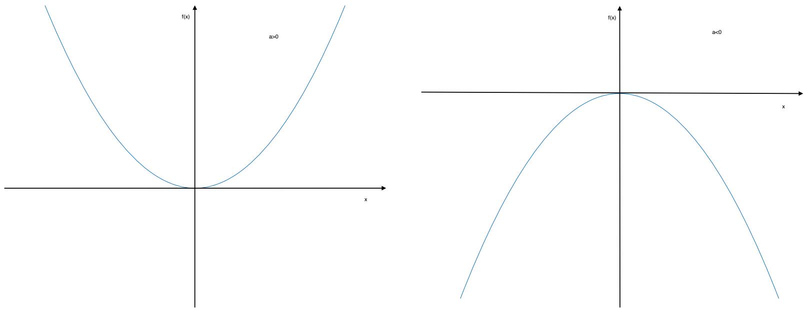


Figure 9: Example of quadratic functions

2.5 Exponential Function

An **exponential function** is a function of the form

$$f(x) = a^x,\ a > 0 \tag{17}$$

2.6 Logarithmic Function

The logarithm is the inverse of the exponential function.

$$f(x) = \log_a x \tag{18}$$

The logarithm of a given number x is the exponent to which another fixed number, the base a, must be raised, to produce that number x. A particular case of the logarithm is the **natural logarithm** which has the number e (that is $e \approx 2.718$) as the base.

3 Limit and continuity

Limit is the value that a function (or sequence) approaches as the input (or index) approaches some value.

Let $X, Y \subseteq \mathbb{R}$ and $f : X \to Y$ be a function of real variables. If

$$\lim_{x \to x_0} f(x) = l \tag{19}$$

means that $f(x)$ can be made to be as close to l as desired, by making x sufficiently close to x_0. Or equivalently, $f(x)$ goes to l as x approaches x_0

In the most simple cases, if $f(x)$ is a real valued function which goes to l as x goes to x_0, we can find l (the limit) substituting x_0 in the function.

Example 3.1.

$$\lim_{x\to 2}(3x - x^2) = 3 \cdot 2 - 2^2 = 2 \tag{20}$$

Example 3.2.

$$\lim_{x\to 2} \ln(x-1) = \ln(2-1) = \ln(1) = 0 \tag{21}$$

The limit of the function can be also $+\infty/-\infty$. In order to find these limits, it is important to know the graph of the function.

Example 3.3.

$$\lim_{x\to +\infty} e^x = +\infty \tag{22}$$

In calculus, a one-sided limit is either of the two limits of a function $f(x)$ of a real variable x as x approaches a specified point either from the left or from the right

Definition 3.1. Right limit

$$\lim_{x\to x_0^+} f(x) = l \tag{23}$$

Definition 3.2. Left limit

$$\lim_{x\to x_0^-} f(x) = l \tag{24}$$

The $\lim_{x\to x_0} f(x) = l$ exists only if the right limit and the left limit exist and they are equal.

Example 3.4. Suppose

$$f(x) = \frac{1}{x-1} \tag{25}$$

We want to study

$$\lim_{x\to 1} \frac{1}{x-1} \tag{26}$$

When $x \to 1$ the function will assume infinite values; more precisely if $x \to 1$ with $x > 1$ the function will tend to $+\infty$, but if $x \to 1$ with $x < 1$ the function will tend to $-\infty$. In this case,

$$\lim_{x\to 1} \frac{1}{x-1} \tag{27}$$

does not exist since

$$\lim_{x\to 1^+} f(x) = +\infty \quad \text{and} \quad \lim_{x\to 1^-} f(x) = -\infty \tag{28}$$

3.1 Mathematical operations

$$\lim_{x\to x_0}[f(x)\pm g(x)] = \lim_{x\to x_0} f(x) \pm \lim_{x\to x_0} g(x) \tag{29}$$

$$\lim_{x\to x_0} f(x)g(x) = \lim_{x\to x_0} f(x) \lim_{x\to x_0} g(x) \tag{30}$$

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{\lim_{x\to x_0 f(x)}}{\lim_{x\to x_0} g(x)} \tag{31}$$

We may end up in a **indeterminate form**

$$\frac{0}{0}, \frac{\infty}{\infty}, 0\times\infty, \infty-\infty \tag{32}$$

Example 3.5.

$$\lim_{x\to 0} \frac{2x^2-3x^5}{x+2x^2} = \lim_{x\to 0} \frac{2x^2(1-3x^3/2)}{x(1+2x)} = \lim_{x\to 0} \frac{2x(1-3x^3/2)}{1+2x} = 0 \tag{33}$$

Example 3.6.

$$\lim_{x\to+\infty} \frac{2x^3+5x^2-x+7}{4x^3-x^2+x-3} = \frac{2x^3(1+\frac{5}{2x}-\frac{1}{2x^2}+\frac{7}{2x^3})}{4x^3(1-\frac{1}{4x}+\frac{1}{4x^2}-\frac{3}{4x^3})} = \frac{1}{2} \tag{34}$$

Example 3.7.

$$\lim_{x\to+\infty} e^x - \sqrt{x} = +\infty \tag{35}$$

3.2 Continuity

We could say if a function is continuous just looking at the graph. A function is continuous if its graph has no breaks (no jumps for the same point) in its domain. However, more rigorously, let $f: X \to Y$ and $x_0 \in X$. $f(x)$ is continuous in x_0 if

$$\lim_{x\to x_0} f(x) = f(x_0) \tag{36}$$

3.3 Differential

Steepness is a key concept in Economics. We are often interested in evaluating what is the effect an increase of the independent variable to the dependent variable. As we have already seen, regarding linear function the steepness is easily given by the slope coefficient. However, most of the economic functions are non-linear. How do we measure the marginal effects of these nonlinear functions?

Let $f : X \to Y$ and $x_0 \in X$. If x increases by Δx, e.g. from x_0 to $x_0 + \Delta x$, also the function $f(x)$ will increase as follows

$$\Delta f(x) = f(x_0 + \Delta x) - f(x_0) \tag{37}$$

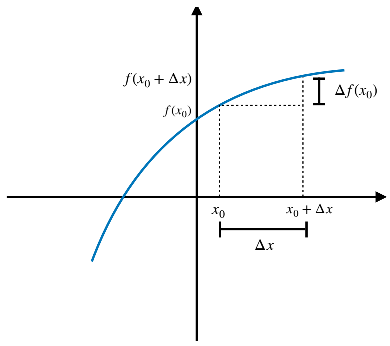


Figure 10: Graphic representation of (37)

However, mostly, it is more interesting studying Difference quotient

$$\frac{\Delta f(x)}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \tag{38}$$

particularly, when Δx is very small

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \tag{39}$$

the limit of the difference quotient as h approaches zero, if it exists, should represent the slope of the tangent line to $(x_0, f(x_0))$. This limit is defined to be the derivative of the function f at x_0:

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0) \tag{40}$$

if the limit of the difference quotient does not exist $f(x)$ is said to be not not differential. If a function is differentiable then it continous the opposite it is not true.

Some derivatives

Function	Derivative	Function	Derivative
$f(x) = k$	$f'(x) = 0$	$f(x) = \sin(x)$	$f'(x) = \cos(x)$
$f(x) = x$	$f'(x) = 1$	$f(x) = \cos(x)$	$f'(x) = -\sin(x)$
$f(x) = x^{\alpha}$, $\alpha \in \mathbb{R}$	$f'(x) = \alpha x^{\alpha-1}$	$f(x) = \cot(x)$	$f'(x) = -\frac{1}{sin^2(x)}$
$f(x) = \alpha^x$	$f'(x) = \alpha^x \ln(\alpha)$	$f(x) = \tan(x)$	$f'(x) = \frac{1}{cos^2(x)}$
$f(x) = e^x$	$f'(x) = e^x$	$f(x) = \arcsin(x)$	$f'(x) = \frac{1}{\sqrt{1-x^2}}$
$f(x) = \log_{\alpha}(x)$	$f'(x) = \frac{1}{x\ln(\alpha)}$	$f(x) = \arccos(x)$	$f'(x) = -\frac{1}{\sqrt{1-x^2}}$
$f(x) = \ln(x)$	$f'(x) = \frac{1}{x}$	$f(x) = \arctan(x)$	$f'(x) = \frac{1}{1+x^2}$
$f(x) = \|x\|$	$f'(x) = \frac{\|x\|}{x}$	$f(x) = arccot(x)$	$f'(x) = -\frac{1}{1+x^2}$

3.4 Derivative of elementary function

- $f(x) = k$, then $f'(x) = 0$, proof;

$$\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{k-k}{\Delta x} = 0$$

- $f(x) = x$, then $f'(x) = 1$, proof;

$$\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{x+\Delta x - x}{\Delta x} = 1$$

- $D\left[\alpha f(x)\right] = \alpha D f(x),\ \alpha \in \mathbb{R}$,
- $D\left[f(x) + g(x)\right] = Df(x) + Dg(x)$
- $D\left[f(x)g(x)\right] = \left[Df(x)\right]g(x) + f(x\left[Dg(x)\right])$
- if $g(x) \neq 0$, $D\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$

For a function to be differentiable, it must be continuous but not vice versa.

Example 3.8. $f(x) = |x|$ is continuous but not differentiable $\forall x \in \mathbf{D}$

$$f'(0) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{h \to 0} \frac{|h|}{h} \tag{41}$$

Therefore, $f(x) = |x|$ is not differentiable in 0.

3.5 Higher order derivative

If $f(x)$ be a differentiable function, and $f'(x)$ its derivative. If $f'(x)$ is again differentiable, $D[f'(x)] = f''(x)$ is its second derivative. If we can do the same with the second derivative we will have the third derivatie, and so on and so forth. Generally, we can talk about **higher order derivative**, defined as $f^{(n)}(x)$.

3.6 Max and min

Let $f : X \to \mathbb{R}$ be a function, and x_0 a point in the domain in X of $f(x)$. We can state that

- x_0 is a minimum if $\exists I_{x_0}$ s.t.

$$f(x) > f(x_0) \;\; \forall x \in I_{x_0} \backslash \{x_0\} \tag{42}$$

- x_0 is a maximum if $\exists I_{x_0}$ s.t.

$$f(x) < f(x_0) \;\; \forall x \in I_{x_0} \backslash \{x_0\} \tag{43}$$

- If x_0 is a max/min, then $f'(x_0) = 0$.
- If $f'(x) > 0 \;\; \forall x \in (a, b)$, then $f(x)$ is increasing in (a, b)
- If $f'(x) < 0 \;\; \forall x \in (a, b)$, then $f(x)$ is decreasing in (a, b)

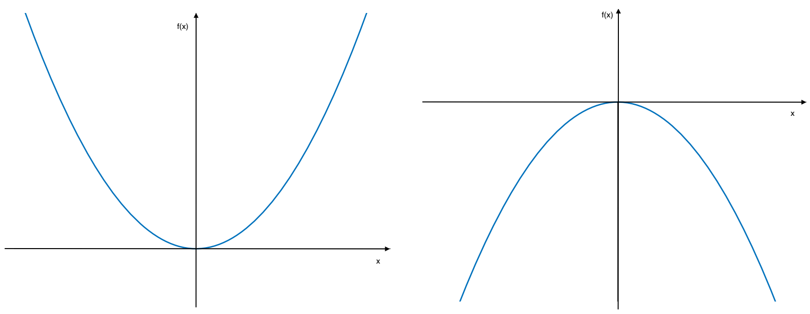


Figure 11: Example of Minimimum - Example of Maximum

4 Integral

The **indefinite integral** represents a class of functions (antiderivative) whose derivative is the integrand.

Definition 4.1. Let $f(x)$ be a function in (a, b). If a function $F(x)$ exists and it is continous in $[a, b]$ and it is also differentiable in (a, b) such that

$$F'(x) = f(x) \ \ \forall x \in (a, b) \tag{44}$$

$F(x)$ is said to be the antiderivative of $f(x)$.

If $F(x)$ is an antiderivative function of $f(x)$, also $G(x) = F(x)+c$, $c \in \mathbb{R}$ is an antiderivative of $f(x)$, given that

$$G'(x) = D[F(x) + c] = F'(x) = f(x) \tag{45}$$

Definition 4.2. Let $f(x)$ be a function in (a, b) and let it have antiderivatives. All the antiderivatives of $f(x)$ are defined as

$$\int f(x)dx \tag{46}$$

$f(x)$ is said to be the integrand.

As the derivative, the integral is a **linear operator**, therefore

$$\begin{aligned} \int[\alpha f(x) + \beta g(x)]dx \ &= \alpha \int f(x)dx + \beta \int g(x)dx \\ &= \alpha[F(x) + c_1] + \beta[G(x) + c_2] \\ &= \alpha F(x) + \beta G(x) + c, \quad c \in \mathbb{R} \end{aligned} \tag{47}$$

The **integration by part** rule is presented here:

$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx + c, \ c \in \mathbb{R} \tag{48}$$

Sometimes, in order to evaluate

$$\int f(x)dx \tag{49}$$

$\int x^{\alpha}dx$	$\frac{x^{\alpha+1}}{\alpha+1},\ c \in \mathbb{R}\ if\ \alpha \neq -1$
$\int \frac{1}{\alpha}$	$\ln\lvert x\rvert + c,\ c \in \mathbb{R}$
$\int [f(x)]^{\alpha} f'(x)dx$	$\frac{[f(x)]^{\alpha+1}}{\alpha+1} + c,\ c \in \mathbb{R}\ if\ \alpha \neq -1$
$\int \frac{f'(x)}{f(x)}dx$	$\ln\lvert f(x)\rvert + c,\ c \in \mathbb{R}$
$\int e^{f(x)} f'(x)dx$	$e^{f(x)} + c,\ c \in \mathbb{R}$
$\int e^{x}dx$	$e^{x} + c,\ c \in \mathbb{R}$
$\int \cos[f(x)]f'(x)dx$	$\sin[f(x)] + c, c \in \mathbb{R}$
$\int \sin[f(x)]f'(x)dx$	$-\cos[f(x)] + c, c \in \mathbb{R}$

it is useful to make a change of variables. For instance, let $g(t)$ be a derivable function, and assume

$$x = g(t) \tag{50}$$

then, dx is changed with

$$dx = g'(t)dt \tag{51}$$

Therefore, we can evaluate Equation (48) as

$$\int f(g(t))g'(t)dt = G(t) + c,\ c \in \mathbb{R} \tag{52}$$

and then solving

$$t = g^{-1}(x) \tag{53}$$

Sometimes, we are interested in evaluating the definite integral; it can be interpreted informally as the signed area of the region in the xy-plane that is bounded by the graph of $f(x)$, the x-axis and the vertical lines $x = a$ and $x = b$. If $f(x)$ is a continuous real-valued function defined on a closed interval $[a, b]$, then, once an antiderivative $F(x)$ of $f(x)$ is known, the definite integral of $f(x)$ over that interval is given by

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a) \tag{54}$$

5 Matrix algebra

The analysis of many economic models reduces to the study of systems of equations. We will, look at the simplest possible system if equations - linear system.

Definition 5.1. Generally, an equation is said to be linear if it has the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \tag{55}$$

where $a_1, ..., a_n$ are **parameters** and $x_1, ..., x_n$ are **variables**.

We can solve linear system of equations using **matrices**.

6 Matrix algebra

We can generate $\mathbb{R}^n$ by n-ary Cartesian product of $\mathbb{R}$. Therefore, we can represent $\mathbb{R}^2$ and $\mathbb{R}^3$ though arrows which connect the origin of the cartesian plan to the coordinates given by the pair (x_1, x_2) or (x_1, x_2, x_3)

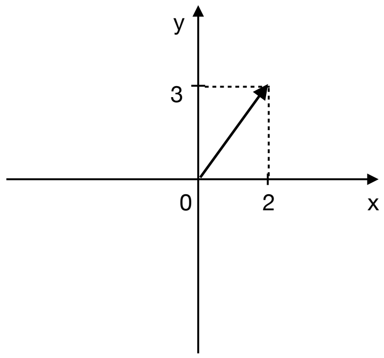


Figure 12: Graphic representation of the vector in $\mathbb{R}^2$ $\underline{v} = (2, 3)$

Definition 6.1. Let $\underline{v}_1 = (x_1, ..., x_n)$ and $\underline{v}_2 = (y_1, ..., y_n)$ be two vectors, their are said to be equal if $x_i = y_i$, $\forall i = 1, 2, .., N$.

Definition 6.2. Let $\underline{v}_1 = (x_1, ..., x_n)$ and $\underline{v}_2 = (y_1, ..., y_n)$ be two vectors, $\underline{v}_3 = (x_1 + y_1, ..., x_n + y_n)$ is the sum of former vectors

We can multiply a vector by a scalar, and we will obtain $c\underline{v}_1 = (cx_1, ..., cx_n)$
Given m vectors $\underline{v}_1, \underline{v}_2, ..., \underline{v}_m$ of $\mathbb{R}^n$ and m scalars $c_1, c_2, ..., c_m$ in $\mathbb{R}$, a **linear combination** of the vectors $\underline{v}_1, \underline{v}_2, ..., \underline{v}_m$ with coefficients $c_1, c_2, ..., c_m$ the vector given by

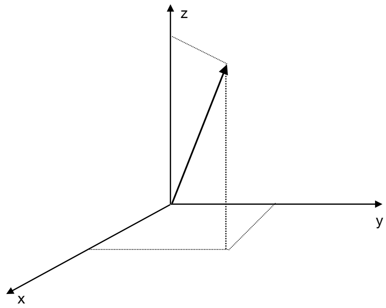


Figure 13: Graphic representation of the vector in $\mathbb{R}^3$

$$\underline{x} = \sum_{k=1}^{m} c_k \underline{v_k} \tag{56}$$

Definition 6.3. A sequence of vectors $(\underline{v_1}, \underline{v_2}, \ldots, \underline{v_k})$ from a vector space V is said to be linearly dependent, if there exist scalars $a_1, a_2, \ldots, a_k$, not all zero, such that

$$a_1\underline{v_1} + a_2\underline{v_2} + \cdots + a_k\underline{v_k} = 0$$

On the other hand, two vectors are said to be linearly independent if there is no linear combination which gives as a result the null vector expect for the one with null coefficients.

6.1 Matrix algebra

A **matrix** is a rectangular array of numbers. The size of a matrix is indicated by the number of its rows and number of its columns. A matrix with k rows and n columns is called a $k \times n$ matrix. The element in row i and column j is called the (i, j)th entry, and it is often written as a_{ij}. A matrix with the number of columns equal to the number of rows is called **square matrix**.

$$A_{k,n} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k,1} & a_{k,2} & \cdots & a_{k,n} \end{pmatrix} \tag{57}$$

6.2 Operations

Definition 6.4. Let A and B two matrices kxn, their sum is the matrix C whose elements are $c_{ij} = a_{ij} + b_{ij}\ \forall i, j$

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{k1} & \cdots & a_{kn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & b_{ij} & \vdots \\ b_{k1} & \cdots & b_{kn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & a_{ij}+b_{ij} & \vdots \\ a_{k1}+b_{k1} & \cdots & a_{kn}+b_{kn} \end{pmatrix} \tag{58}$$

Matrices may be multiplied by scalars. This operation is called **scalar multiplication**. More generally:

Definition 6.5. The product of the matrix A and the number α, denoted αA, is the matrix whose elements are $\alpha a_{ij}\ \forall i, j$

$$\alpha \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{k1} & \cdots & a_{kn} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & \alpha a_{ij} & \vdots \\ \alpha a_{k1} & \cdots & \alpha a_{kn} \end{pmatrix} \tag{59}$$

We can define the **matrix product** AB iff:

$$\text{number of columns of } A = \text{number of rows of } B \tag{60}$$

To obtain the (i, j)th entry of AB, multiply the ith row of A and the jth column of B as follows

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{pmatrix} \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{ni} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj} \tag{61}$$

In other words, the (i, j)th entry of the product AB is defined to be

$$\sum_{n-1}^{m} a_{in}b_{hj}$$

If A is a $k \times m$ and B is $m \times n$, then the product $C = AB$ will be $k \times n$.

Usually, $AB \neq BA$.

The $n \times n$ matrix

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{62}$$

with $a_{ii} = 1$, $\forall i$ and $a_{ij} = 0$, $\forall i \neq j$, has the following property

$$AI = A \tag{63}$$

for any mxn matrix A. I is called **identity matrix**.

Definition 6.6. The **transpose** of a kxn matrix A is the nxk matrix obtained by interchanging the rows and the columns.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix} \tag{64}$$

Definition 6.7. A **Triangular matrix** is a square matrix containing element different from zero only above/below the main diagonal.

Example 6.1. Upper Triangular

$$A = \begin{pmatrix} 1 & -3 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{pmatrix} \tag{65}$$

Example 6.2. Lower Triangular

$$A = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & 5 & -2 \end{pmatrix} \tag{66}$$

Definition 6.8. To any square matrix is associated a **determinant**, $|A|$. From a geometric point of view, it represents the area (volume) of the parallelogram generated by the vectors of the matrix.

Let A be a $2x2$ matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{67}$$

the determinant is $|A| = a_{11}a_{22} - a_{12}a_{21}$

In case of a $n \times n$ matrix with $n > 3$, in order to find the determinant, we can use the **minor** of the matrix.

Definition 6.9. A minor of a matrix A is the determinant of some smaller square matrix, cut down from A by removing one or more of its rows and columns. Minors obtained by removing just one row and one column from square matrices (first minors) are required for calculating matrix cofactors, which in turn are useful for computing both the determinant and inverse of square matrices.

The cofactor is:

$$A_{ik} = (-1)^{i+k} M_{ik} \tag{68}$$

where M_{ik} is a minor of the matrix.

Therefore, using the **Laplace theorem** we can obtain the determinant of a matrix $n \times n$ as the sum of the product, of any row or column, by their cofactor.

The **rank** of A is the largest order of any non-zero minor in A.

Definition 6.10. Let A be a $n \times n$ matrix. The $n \times n$ matrix A^{-1} is an **inverse** for A if $AA^{-1} = A^{-1}A = I_n$

A matrix can have at most one inverse

$\Rightarrow$ Nor every matrix are invertible. In order to be invertible, $|A| \neq 0$

Definition 6.11. Any symmetric A is:

- **Positive semidefinite** if $x'Ax > 0$, $\forall x \neq 0$
- **Positive definite** if $x'Ax \geq 0$, $\forall x \neq 0$
- **Negative definite** if $x'Ax < 0$, $\forall x \neq 0$
- **Negative semidefinite** if $x'Ax \leq 0$, $\forall x \neq 0$

6.3 Linear system of equations

As we have stated above, we can use matrix to solve linear system of equations.

Recall: Generally, an equation is said to be linear if it has the form

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b \tag{69}$$

where $a_1, ..., a_n$ are **parameters** and $x_1, ..., x_n$ are **variables**.

The solution of the linear equation is given by $(x_1, ..., x_n)$ which substituted into the equation solving it.

If there are several linear equations which have to be true all toghether we talk about **Linear System of equations**

Example 6.3.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + ...a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ...a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ...a_{mn}x_n = b_m \end{cases} \tag{70}$$

The solution of the system is given by $(x_1, x_2, ..., x_n)$ which solves all the equations contemporaneously.

The system above, can be expressed much more compactly using matrix notations. Let A denote the coefficient matrix of the system:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \tag{71}$$

Also, let

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{72}$$

Then, the system of equations can be written as

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \tag{73}$$

or simply as

$$A\mathbf{x} = \mathbf{b} \tag{74}$$

Then, if A is *nonsingular* ($|A| \neq 0$), we can solve the system as $\mathbf{x} = A^{-1}\mathbf{b}$. To solve a linear system of simultaneous equations we can use also the **Cramer's rule**. If the matrix A is nonsingular, the linear system of system of n linear equations and n unknowns. Then the theorem states that in this case the system has a unique solution, whose individual values for the unknowns are given by:

$$x_i = \frac{det(A_i)}{det(A)} \qquad i = 1, ..., n \tag{75}$$

where A_i is the matrix formed by replacing the ith column of A by the column vector b.

Example 6.4. Solve the following system

$$\begin{cases} x - 2y + z = 1 \\ 3x + y - 7z = 0 \\ x - z = 1 \end{cases} \tag{76}$$

Since $|A| \neq 0$, we can use Cramer's rule

$$x = \frac{det \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -7 \\ 1 & 0 & -1 \end{pmatrix}}{6} = 2$$
$$y = \frac{det \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & -7 \\ 1 & 1 & -1 \end{pmatrix}}{6} = 1 \quad (77)$$
$$z = \frac{det \begin{pmatrix} 1 & -2 & 1 \\ 3 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}}{6} = 1$$

In Economics, we may be interested in system of the following form

$$Ax = \lambda x \quad (78)$$

where A is a square matrix.

It is equal to write:

$$(A - \lambda I)x = 0 \quad (79)$$

Definition 6.12. The values λ that solve $det(A - \lambda I) = 0$ are called **eigenvalues**.

Definition 6.13. While the non-trivial vectors, $\mathbf{x}$, obtained as the solution of $(A - \lambda I)x = 0$ is called **eigenvector**.

From a geometric point of view, the eigenvector (x) is the vector which is only scaled by a value λ when we apply to it a transformation A.

7 Function of several variables

As we have seen before, in Economics is interested looking what is the effect of a change in one variable with respect to another one. However, in most real cases, variables depend on several variables. We may be interested in $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

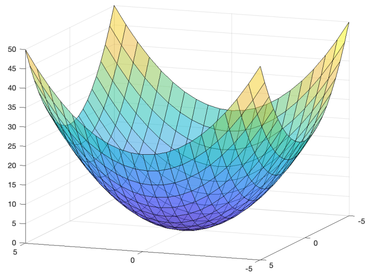


Figure 14: The graph of $f(x, y) = x^2 + y^2$

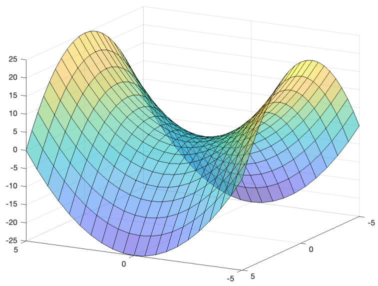


Figure 15: The graph of $f(x, y) = x^2 - y^2$

We can graph function up to $\mathbb{R}^3$.

There is another way to visualize function from $\mathbb{R}^2$ to $\mathbb{R}^1$ which requires only two dimensional sketching: the **level curves**. They are given by

$$L_c = (x, y) : f(x, y) = c \tag{80}$$

It is like to slice the function in many pieces.

7.1 Calculus of several variable

When we deal with function of several variable, we are often interested in the partial variation - the variation brought about by the change in only one variable. We will talk, therefore, about **partial derivative**.

Definition 7.1. Let $f : \mathbb{R}^n \to \mathbb{R}$. Then for each variable x_i at each point x^0 in the domain of f

$$\frac{\partial f}{\partial x_i}\left(x_1^0, \ldots, x_n^0\right) = \lim_{h \to 0} \frac{f\left(x_1^{(0)}, \ldots, x_i^0 + h, \ldots, x_n^0\right) - f\left(x_1^0, \ldots, x_i^0, \ldots, x_n^0\right)}{h} \tag{81}$$

From a practical point of view, we can apply the same rules we apply to one variable function; we take the derivative with respect one variable treating all the other variables as constant.

If we are interested in the behavior of a function $F(x, y)$ of two variables in the neighborhood of a given point $(x^\star, y^\star)$, we can, therefore, look at the effect of the change from $x^\star$ $(y^\star)$ to $x^\star+\Delta x$ $(y^\star + \Delta y)$ respectively:

$$F(x^* + \Delta x, y^*) - F(x^*, y^*) \approx \frac{\partial F}{\partial x}(x^*, y^*)\,\Delta x \tag{82}$$

$$F(x^*, y^* + \Delta y) - F(x^*, y^*) \approx \frac{\partial F}{\partial y}(x^*, y^*)\,\Delta y \tag{83}$$

or we can look at the effect of a change of x and y simultaneously as follows

$$F(x^* + \Delta x, y^* + \Delta y) - F(x^*, y^*) \approx \frac{\partial F}{\partial x}(x^*, y^*)\,\Delta x + \frac{\partial F}{\partial y}(x^*, y^*)\,\Delta y \tag{84}$$

7.2 Higher-order derivatives

The partial derivative $\partial f/\partial x_i$ is itself a function - as we have seen in the case of the second derivative. When we take the first derivative with respect to a variable and the second one with respect to another variable we talk about **mixed partial derivatives** and it is usually written as

$$\frac{\partial^2 f}{\partial x_j \partial x_i} \tag{85}$$

with $i \neq j$.

If a function has n variables, then, it will have n^2 second order partial derivatives. It is common to arrange these n^2 partial derivatives into an $n \times n$ matrix whose (i, j)th entry is the $(\partial^2 f/\partial x_j \partial x_i)(x^*)$. This matrix is called **Hessian matrix**:

$$D^2 f_{\mathbf{x}} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \tag{86}$$

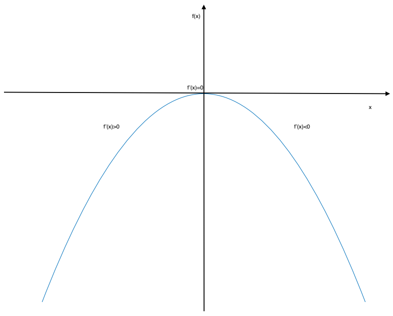
Since $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$, the Hessian matrix is often symmetric.

8 Optimization

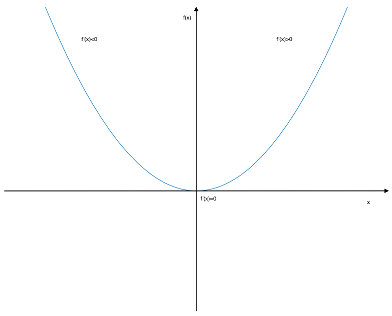
We want to find the point (x, y) s.t. $f'(x) = 0$, this point is called **stationary point** (**critical point**).

A critical point may correspond to:

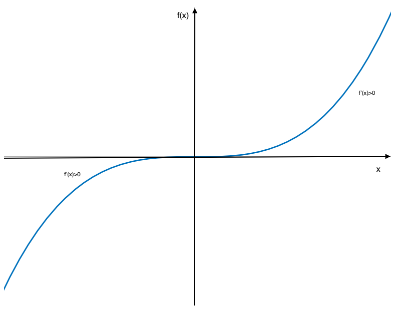
- **Maximum**



- **Minimum**



- **Saddle point**



Once we have found the stationary point with $f'(x) = 0$ (first order condition), we can check what kind of stationary point is looking at the second derivative (second order conditions)

- if $f''(x) < 0 \Rightarrow$ max

- if $f''(x) > 0 \Rightarrow$ min
- if $f''(x) = 0 \Rightarrow$ No clue. It can be max, min or saddle point

in $x^\star$ we have a local max if $f(x^\star) \geq f(x)$, $\forall x$ in the of $x^\star$

8.1 Optimization with several variables

The **gradient vector** is the vector whose components are the partial derivatives of f.

$$\nabla f(p) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(p) \\ \vdots \\ \frac{\partial f}{\partial x_n}(p) \end{bmatrix} \tag{87}$$

The **Hessian Matrix** is the $n \times n$ matrix whose (i,j)th entry is the $\left(\partial^2 f / \partial x_j \partial x_i\right)(\mathbf{x}^*)$.

$$D^2 f_{\mathbf{x}} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} \tag{88}$$

In order to have a stationay point $\nabla f(p) = 0$

In order to attain the second order conditions:

- If the Hessian is positive definite (equivalently, has all eigenvalues positive) at a, then f attains a local minimum at a
- If the Hessian is negative definite (equivalently, has all eigenvalues negative) at a, then f attains a local maximum at a
- If the Hessian has both positive and negative eigenvalues then a is a saddle point for f (and in fact this is true even if a is degenerate).

In those cases not listed above, the test is inconclusive.

8.2 Constrained optimization

Let f and h be $\mathcal{C}^1$ functions of two variables. Suppose that $\mathbf{x}^\star = (x_1^\star, x_2^\star)$ is a solution of the problem

$$\begin{aligned} &\text{maximize} && f(x_1, x_2) \\ &\text{subject to} && h(x_1, x_2) = c \end{aligned} \tag{89}$$

Suppose that $(x_1^\star, x_2^\star)$ is not a critical point of h. Then, there is a real number $\mu^\star$ s.t. $(x_1^\star, x_2^\star, \mu^\star)$ is a critical point of the Lagrangian function

$$L(x_1, x_2, \mu) \equiv f(x_1, x_2) - \mu [h(x_1, x_2) - c] \tag{90}$$

In other words at (x_1^*, x_2^*, μ^*)

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \text{and} \quad \frac{\partial L}{\partial \mu} = 0 \tag{91}$$

Reference

Simon, Carl P., and Lawrence Blume. Mathematics for economists. Vol. 7. New York: Norton, 1994