NOTES ON MATHEMATICS

Gonzalo Rodríguez

Departament de Matemàtica Econòmica,

Financera i Actuarial

GEI (Grau d'Empresa Internacional)

PROLOGUE

The aim of this manual is to provide GEI students with a thorough introduction to the contents of *Mathematics* course and, to this end, it will be used and further expanded upon in class. Its contents are essential for ensuring that students are in handling many of the formal and quantitative matters that underpin other subjects in this GEI.

The manual comprises three main sections. The first, *Linear Algebra*, deals with the basic properties of vectors – i.e., directed line segments - which unlike scalar magnitudes are determined by more than one numerical value. A good grounding in this subject is fundamental for understanding the second section, devoted to *Multivariable Optimization*, in which we search for the maxima and minima (extreme points, optima) of functions of several variables without any constraints ("Classical Optimization"). Finally, the third section, *Dynamic Analysis*, deals with integrals of continuous functions of one real variable. This is an indispensable tool for calculating planar areas determined by functions as well as for solving differential equations, which in turn play a vital role in Mathematical Economics. Precisely this manual ends up studying the simplest cases of these differential equations.

Note that each of the subjects tackled here contains various relevant applications in Economics as well as a few interesting exercises intended to solve in the classroom. The manual concludes with some bibliographic references and a glossary of terms and page references to help quickly find the most important concepts quoted here. Finally point out that this manual has been filed in the UB Digital Repository in the URL adress:

http://hdl.handle.net/2445/45663.

Please note that all contents as well as any errors that might be found here are the sole and exclusive responsibility of the author.

Gonzalo Rodríguez Pérez

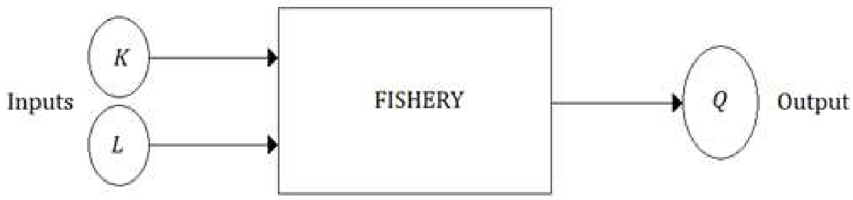
INDEX

1. LINEAR ALGEBRA

Numerical magnitudes that scientists usually deal with are basically of two types: *scalar* and *vectorial.* Scalar magnitudes are those determined by a sole numerical value; consider, for instance, the weight of a body, the price of a good or the rate of interest. In the other hand, vectorial magnitudes need, unlike scalars, more than one value to be explicited. In Economics, vectorial magnitudes or vectors appear frequently as independent variables of functions. Let us consider for example a lobster fishery where its "output" $Q \geq 0$ stands for the estimated catch of lobster depending on two "inputs", the stock of lobster $K \geq 0$ and the harvesting effort $L \geq 0$ in this way: [1]

$$Q = 2.26 \cdot K^{0.44} \cdot L^{0.48}.$$

Schematically:



In this scenario, the lobster fishery is "modeled" by one function of two independent variables: [2] the first is the stock of lobster, $K \geq 0$, and the second is the harvesting effort $L \geq 0$. It is very important to emphasize that these variables are non-related each other in any way. On the contrary, the third variable, the output $Q \geq 0$, depends on both at the same time. Therefore, the domain of functions like this are not a set of points of the real line $\mathbb{R}$ but a set of *vectors* of the real plane $\mathbb{R}^2$ defined by:

$$\mathbb{R}^2 = \{\vec{x} = (x_1, x_2) : x_1, x_2 \in \mathbb{R}\}.$$

Hence, if we want to understand how functions of several variables (i.e., functions depending on more than one variable) work, we need prior to tackle with vectors like these. This section discusses some of the seminal properties of *vector spaces* (i.e., sets of vectors with a sum and product operations associated) to further apply them to study some metric concepts like norms, angles and distances between vectors. Finally, we introduce quadratic forms that we will need in the next section.

---

[1] See Sydsaeter. K & Hammond, P. J. (1995), page 490. This function is a particular case of a Cobb-Douglas's function of production.

[2] We deal with this type of functions in the second section.

1.1. Vector Spaces

1.1.1. $\mathbb{R}^n$-*Vector Space, Vectors and Scalars*

The concept of a vector space incorporates the basic properties of the sum and the external product (the product by a number) of vectors.

<u>Definition</u>: The **vector space** $\mathbb{R}^n$ ($\mathbb{R}^n$ for short) is a mathematical structure given by:

1. A set of **vectors**:

$$\mathbb{R}^n = \{\vec{x} = (x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\}$$

2. A vector **addition**:

$$\vec{x} + \vec{y} = (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

3. A **multiplication** by scalars:

$$\lambda \cdot \vec{x} = \lambda \cdot (x_1, \dots, x_n) = (\lambda \cdot x_1, \dots, \lambda \cdot x_n), \text{ for any } \lambda \in \mathbb{R}.$$

In this context, the elements of the set of real numbers $\mathbb{R}$ are called **scalars**.

From now onwards we will denote the **zero-vector** of $\mathbb{R}^n$ by:

$$\vec{0} = \left(\overbrace{0, \dots 0}^{n}\right) \in \mathbb{R}^n.$$ [3]

On the other hand, the **opposite** vector of $\vec{x} \in \mathbb{R}^n$ will be the vector:

$$-\vec{x} = (-1) \cdot \vec{x}.$$

Any vector space satisfies the following basic properties:

<u>Properties</u>: In a vector space $\mathbb{R}^n$ we have:

1. $\lambda \cdot \vec{x} = \vec{0}$ implies either $\lambda = 0$ or $\vec{x} = \vec{0}$.
2. $-(\lambda \cdot \vec{x}) = (-\lambda) \cdot \vec{x}$.
3. $\vec{x} = \vec{y}$ is equivalent to $\vec{x} - \vec{y} = \vec{0}$.

---

[3] This vector $\vec{0}$ should not be confused with the scalar 0.

1.1.2. *Linear Combination of Vectors*

We begin by introducing the most basic concept related to vectors, namely, the linear combination. For instance, the vector $(3,-1,7)$ is a linear combination of the vectors $(1,0,4)$ and $(-1,1,1)$ since it can be written as:

$$(3,-1,7) = 2 \cdot (1,0,4) + (-1) \cdot (-1,1,1).$$

Thus:

<u>Definition</u>: The vector $\vec{x} \in \mathbb{R}^n$ is a **linear combination** of k vectors $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$ provided that k scalars $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ exist with the condition:

$$\vec{x} = \lambda_1 \cdot \vec{x}_1 + \dots + \lambda_k \cdot \vec{x}_k = \sum_{i=1}^{k} \lambda_i \cdot \vec{x}_i \text{.}$$ [4]

Consider this example:

<u>Example</u>: Prove that the vector $(-1,9,4)$ is a linear combination of the two vectors $(1,2,0)$ and $(-4,3,4)$. Reason that this is not the case for the vector $(7,3,4)$.

SOLUTION: In the first case we must solve the vector-equation:

$$(-1,9,4) = \lambda_1 \cdot (1,2,0) + \lambda_2 \cdot (-4,3,4).$$

Every vector-equation always leads to a linear system of equations. In this case:

$$\left.\begin{array}{l} -1 = \lambda_1 - 4\lambda_2 \\ 9 = 2\lambda_1 + 3\lambda_2 \\ 4 = 4\lambda_2 \end{array}\right\} \text{with the scalars } \lambda_1 \text{ and } \lambda_2 \text{ as variables.}$$

Since this system is consistent, we deduce that λ_1 and λ_2 exist and hence the linear combination.[5] However, this is not the case for the vector $(7,3,4)$ because the linear system associated:

$$(7,3,4) = \lambda_1 \cdot (1,2,0) + \lambda_2 \cdot (-4,3,4) \text{ equivalent to: } \begin{cases} 7 = \lambda_1 - 4\lambda_2 \\ 3 = 2\lambda_1 + 3\lambda_2 \\ 4 = 4\lambda_2 \end{cases}$$

is inconsistent■

[4] Note that the number k of vectors does not have to match with the number n of the vector components.

[5] The values of the two scalars are $\lambda_1 = 3$ and $\lambda_2 = 1$. It must be pointed out that it is not necessary to figure out these values in order to prove the existence of the linear combination.

1.1.3. *Linear Dependence and Independence of Vectors. Basis of a Vector Space*

As we have just seen, the vector $(-1,9,4)$ can be expressed as a linear combination of the two vectors $(1,2,0)$ and $(-4,3,4)$. In this case, we say that the three vectors:

$$(-1,9,4),\ (1,2,0) \text{ and } (-4,3,4) \text{ are linearly dependent.}$$

On the other hand, we know that this is not true for the vector $(7,3,4)$. In this case, we say that:

$$(7,3,4),\ (1,2,0) \text{ and } (-4,3,4) \text{ are linearly independent.}$$

Thus, by definition:

Definition: We say that k vectors $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$ are:

1. **Linearly dependent** if at least one of them, $\vec{x}_i$ for instance, can be expressed as a linear combination of the others. Formally:

$$\vec{x}_i = \sum_{j=1, j\neq i}^{k} \lambda_j \cdot \vec{x}_j.$$

2. **Linearly independent** otherwise, i.e., when none of the k vectors can be expressed as a linear combination of the rest.[6]

The next theorem characterizes the linear independence through the so-called zero-linear combinations of vectors. Indeed:

Theorem: k vectors $\vec{x}_1, \dots, \vec{x}_k \in \mathbb{R}^n$ are linearly independent if and only if every zero-linear combination has the corresponding scalars equal to zero. In other words:

$$\sum_{i=1}^{k} \lambda_i \cdot \vec{x}_i = \lambda_1 \cdot \vec{x}_1 + \cdots + \lambda_k \cdot \vec{x}_k = \vec{0} \text{ implies: } \lambda_1 = \cdots = \lambda_k = 0.$$

Obviously, any set of vectors will be linearly dependent in the case there exists a zero-linear combination with at least one non-zero scalar.

[6] In the case of $k = n$ a set of linearly independent vectors is called a **basis** of the vector space $\mathbb{R}^n$ being n its **dimension.**

1.1.3.1. Study of the Linear Independence by means of Zero-Linear Combinations

The above theorem is very suitable in those cases in which the components of the vectors under study are unknown. Consider the following example:

Example: Prove that:

1. Every set of vectors of any vector space containing the zero-vector is a set of linearly dependent vectors.
2. If three vectors $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ are linearly independent vectors so they are:

$$\begin{cases} \vec{u} = \vec{x} \\ \vec{v} = \vec{x} + \vec{y} \\ \vec{w} = \vec{x} + \vec{y} + \vec{z} \end{cases}$$

SOLUTION: (1) Consider k vectors of $\mathbb{R}^n$ with the associated zero-vector $\vec{0} \in \mathbb{R}^n$:

$$\vec{x}_1, \dots, \vec{x}_k, \vec{0}.$$

Following the guidelines of the last theorem we just have to to build a zero-linear combination among them with not all the scalars equal to zero. Indeed, the zero-linear combination:

$$0 \cdot \vec{x}_1 + \cdots + 0 \cdot \vec{x}_k + 1 \cdot \vec{0} = \vec{0}$$

shows us that this is the case. Consequently, $\vec{x}_1, \dots, \vec{x}_k, \vec{0} \in \mathbb{R}^n$ are linearly dependent■

(2) Consider any zero-linear combination of the three vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$:

$$\vec{0} = \lambda_1 \cdot \vec{u} + \lambda_2 \cdot \vec{v} + \lambda_3 \cdot \vec{w}.$$

In order to show they are linearly independent we must only prove that:

$$\lambda_1 = \lambda_2 = \lambda_3 = 0$$

taking into account that, by hypothesis, the vectors $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ are so. First note that the above zero-linear combination can be written as follows:

$$\vec{0} = \lambda_1 \cdot \vec{u} + \lambda_2 \cdot \vec{v} + \lambda_3 \cdot \vec{w} = \lambda_1 \cdot \vec{x} + \lambda_2 \cdot (\vec{x} + \vec{y}) + \lambda_3 \cdot (\vec{x} + \vec{y} + \vec{z}) =$$
$$= (\lambda_1 + \lambda_2 + \lambda_3) \cdot \vec{x} + (\lambda_2 + \lambda_3) \cdot \vec{y} + \lambda_3 \cdot \vec{z}.$$

Since $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ are linearly independent we conclude that these scalars must be zero:

$$\lambda_1 + \lambda_2 + \lambda_3 = \lambda_2 + \lambda_3 = \lambda_3 = 0.$$

These homogeneus lineal equalities match only under the assumption that:

$$\lambda_1 = \lambda_2 = \lambda_3 = 0$$

as we wanted to prove. Thus, the vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ are also linearly independent■

1.1.3.1. Relationship between Linear Independence and Rank of a Matrix

If we recall that the rank of a matrix is equal to the order of the largest associated non-zero minor, this relationship states that:

Theorem: A set of k vectors of $\mathbb{R}^n$ are linearly independent if and only if the matrix they form (either by rows or columns) has rank equal to k.[7]

Consider the following example:

Example: Prove that the vectors:

1. $(1,-1,3,7)$, $(5,2,4,-2)$ and $(0,-6,0,1)$ are linearly independent.
2. $(2,-2,8)$, $(5,1,3)$ and $(4,2,-1)$ are linearly depedent.

SOLUTION:

(1) The matrix formed by these three row-vectors is:

$$\begin{pmatrix} 1 & -1 & 3 & 7 \\ 5 & 2 & 4 & -2 \\ 0 & -6 & 0 & 1 \end{pmatrix}.$$

Since we have a non-zero minor of order 3 formed by the firsts three columns:

$$\begin{vmatrix} 1 & -1 & 3 \\ 5 & 2 & 4 \\ 0 & -6 & 0 \end{vmatrix} = -66 \neq 0$$

we deduce, thanks to the above theorem, that these vectors are linearly independent since the rank of the above matrix matches the number of vectors■

(2) In this case, since the determinat of the square matrix that these three vectors form is equal to zero:

$$|A| = \begin{vmatrix} 2 & -2 & 8 \\ 5 & 1 & 3 \\ 4 & 2 & -1 \end{vmatrix} = 0$$

we have three linearly dependent vectors because the rank now is less than 3■

[7] Consequently, the rank of a matrix is the maximum number of linearly independent row (or column) vectors. In particular, n vectors of $\mathbb{R}^n$ form a basis if and only if the associated matrix is regular, i.e., a square matrix with a non-zero determinant.

1.1.3.2. Application of Linear Independence in Economics

Let us consider this application of linear independence of vectors:[8]

Example: [9] An oil company can convert one barrel of crude oil into three different kinds of fuel according to the vector equation: [10]

$$(x, y, z) = (1-\theta) \cdot (2,2,4) + \theta \cdot (5,0,3)$$

where $x, y, z \geq 0$ stand for the amounts of the three types of fuel obtained, and $0 \leq \theta \leq 1$ is the percentage of lead additives added. Check if the following output vectors:

$$\vec{u} = (3.5,1,3.5), \vec{v} = \left(4, {}^{1}/_{3}, {}^{10}/_{3}\right) \text{ and } \vec{w} = (1,6,9)$$

are possible and, if so, determine the percentage of lead additives included.

SOLUTION:

First, we must see whether the three set of the vectors:

$$\{\vec{u}, (2,2,4), (5,0,3)\}, \{\vec{v}, (2,2,4), (5,0,3)\} \text{ and } \{\vec{w}, (2,2,4), (5,0,3)\}$$

are linearly dependent. Since:

$$\begin{vmatrix} 3.5 & 1 & 3.5 \\ 2 & 2 & 4 \\ 5 & 0 & 3 \end{vmatrix} = 0, \begin{vmatrix} 4 & {}^{1}/_{3} & {}^{10}/_{3} \\ 2 & 2 & 4 \\ 5 & 0 & 3 \end{vmatrix} = -{}^{14}/_{3} \text{ and: } \begin{vmatrix} 1 & 6 & 9 \\ 2 & 2 & 4 \\ 5 & 0 & 3 \end{vmatrix} = 0$$

we conclude that $\vec{u}$ and $\vec{w}$ are two possible output vectors. Taking into account that:

$$\vec{u} = (3.5,1,3.5) = \frac{1}{2} \cdot (2,2,4) + \frac{1}{2} \cdot (5,0,3)$$

and:

$$\vec{w} = (1,6,9) = 3 \cdot (2,2,4) + (-1) \cdot (5,0,3)$$

the vector $\vec{u}$ is the only case where the percentage $0 \leq \theta \leq 1$ exists since:

$$\theta = \frac{1}{2}$$

while the vector $\vec{w}$ fails because there is no value $0 \leq \theta \leq 1$ such that:

$$1 - \theta = 3 \text{ and: } \theta = -1.$$

Thus:

$$\vec{u} = (3.5,1,3.5)$$

is an output vector with $\theta = 50\%$ of lead additives■

[8] Roughly speaking, this is an application of linear dependence.
[9] Sydsaeter, K. & Hammond, P. J. (1995), page 384.
[10] In this scenario, the vector (x, y, z) is called "output vector".

1.2. Scalar Product on Vector Spaces

1.2.1. *Scalar Product on* $\mathbb{R}^n$

We are going to introduce the algebraic notion of scalar product with the aim to deal with some metric concepts in $\mathbb{R}^n$ such as the norm of a vector (i.e., its length), the angle formed by two vectors and the distance between two vectors. Among others, these concepts are geometrically related with the length of a linear segment, the angle between two straight lines and the distance between points.[11]

Definition:[12] The **scalar** (or **inner**) **product** of two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is equal to the real number (scalar) obtained through the expression:

$$\vec{x} \cdot \vec{y} = (x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = x_1 \cdot y_1 + \cdots + x_n \cdot y_n = \sum_{i=1}^{n} x_i \cdot y_i \in \mathbb{R}.$$

Example: Calculate:

1. The scalar product of the vectors $(1,0,-5)$ and $(5,-2,1)$.
2. The value of $k \in \mathbb{R}$ so that the scalar product of $(5,-2,1)$ and $(1,-1,k)$ is equal to 10.

SOLUTION:

(1) In this case we trivially have:

$$(1,0,-5) \cdot (5,-2,1) = 1 \cdot 5 + 0 \cdot (-2) + (-5) \cdot 1 = 5 - 5 = 0 \blacksquare$$

(2) The value of the parameter $k \in \mathbb{R}$ will be:

$$10 = (5,-2,1) \cdot (1,-1,k) = 5 \cdot 1 + (-2) \cdot (-1) + 1 \cdot k = 7 + k \text{ implies: } k = 3 \blacksquare$$

It must be emphasized that the scalar product of two vectors is not a vector but a number, i.e., a scalar and hence the name. Likewise, as we have just seen in this example, the scalar product of two non-zero vectors can nevertheless be equal to 0; recall that this is not the case when we have to cope with real numbers.

[11] Roughly speaking every vector of a vector space with origin at the zero-vector $\vec{0}$ can be viewed as the **position-vector** of its ending point and vice versa. So, the same set of coordinates can be used to express vectors and points related in this way.

[12] In general, a vector space with a scalar product is called a **Euclidian vector space**.

1.2.1.1. Application of Scalar Product in Economics[13]

Example: If the n positive numbers $p_1, \dots, p_n > 0$ denote the unit sale prices of n products:

1. Prove that the total income obtained selling $q_1, \dots, q_n > 0$ units of them matches the scalar product of the two vectors:

$$\vec{p} = (p_1, \dots, p_n) \text{ and: } \vec{q} = (q_1, \dots, q_n).$$

2. If €5, €2 and €3 are respectively the unit sale prices of three economic goods, determine how many we need to sell if we want to obtain an income of €105 where the amount sold of the second good is half that of the first, and the amount sold of the third equals the sum of the other two.

SOLUTION: (1) It is clear that the total income I obtained selling these n products equals the sum of the partial incomes associated with each product. Thus, as the sale of $q_i \geq 0$ units of the i-product provides a partial income of €$p_i \cdot q_i$ then the total income will be:

$$I = \sum_{i=1}^{n} p_i \cdot q_i = p_1 \cdot q_1 + \dots + p_n \cdot q_n = \vec{p} \cdot \vec{q} \blacksquare$$

(2) In the specific case of:

$$\vec{p} = (p_1, p_2, p_3) = (5,2,3) \text{ and: } I = €105$$

we would have the linear equality:

$$105 = I = 5q_1 + 2q_2 + 3q_3$$

associated, by hypothesis, with the two conditions:

$$q_2 = {}^{q_1}/_2 \text{ and: } q_3 = q_1 + q_2.$$

Since:

$$q_2 = {}^{q_1}/_2 \text{ implies: } q_3 = q_1 + q_2 = {}^{3q_1}/_2$$

we can write the total income of €105 as:

$$105 = I = 5q_1 + 2q_2 + 3q_3 = 5q_1 + 2\left({}^{q_1}/_2\right) + 3\left({}^{3q_1}/_2\right) = {}^{21q_1}/_2$$

from which it follows that:

$$q_1 = 10 \text{ and: } q_2 = {}^{q_1}/_2 = 5, q_3 = {}^{3q_1}/_2 = 15.$$

So, we must sell 10 units of the first good, 5 units of the second and 5 units of the third good if we want to obtain an income of €105 ■

[13] Adillon, R.; Alvarez, M.; Gil, D.; Jorba, L. (2015), pages 125-126.

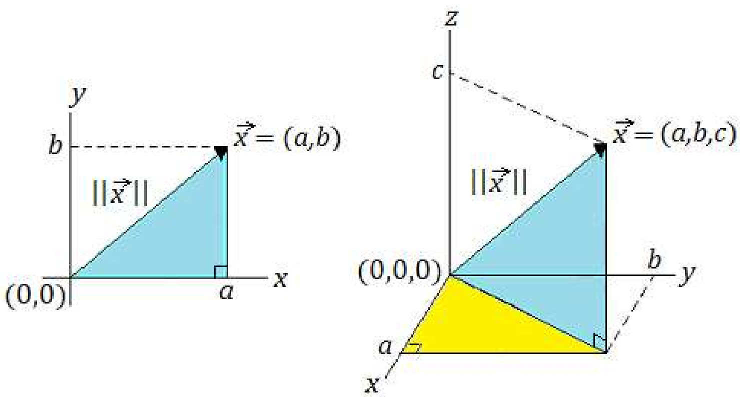
1.2.2. *Norm of a Vector and Properties*

We are now ready to deal with some basic metric concepts in $\mathbb{R}^n$ through the associated scalar product. First, we are going to introduce the notion of the norm (i.e., the length) of a vector. As the scalar product of any vector by itself is always positive, we can consider:

Definition: The **norm** of the vector $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ is the positive square root:

$$\|\vec{x}\| = +\sqrt{\vec{x} \cdot \vec{x}} = +\sqrt{x_1^2 + \cdots + x_n^2} = +\sqrt{\sum_{i=1}^{n} x_i^2}.$$

In fact, the norm of any vector in $\mathbb{R}^n$ reproduces the theorem of Pythagoras. Indeed:



We will have to deal with some of the following properties of the norm of a vector:

Properties:

1. $\|\vec{x}\| \geq 0$. Moreover $\|\vec{x}\| = 0$ is equivalent to: $\vec{x} = \vec{0}$.
2. "Triangular inequality": $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$.[14]
3. "Schwarz inequality": $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \cdot \|\vec{y}\|$.[15]
4. If $\vec{x} \neq \vec{0}$, the vector:

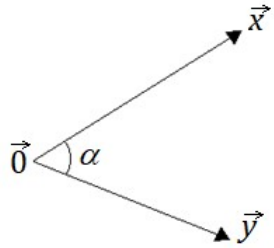
$$\frac{1}{\|\vec{x}\|} \cdot \vec{x}$$

is an **unit vector**, i.e., a vector whose norm is 1.

[14] This property is related to the mathematical statement that claims that the length of any of the three sides of a triangle is always less or equal to the sum of the other two.
[15] The Schwarz inequality enable us to define the notion of angle between vectors in any Euclidian vector space.

1.2.3. *Angle between Vectors and Orthogonal Vectors*

One of the main concepts in metric geometry is the angle formed by two straight lines or, in our case, by two vectors. Graphically:



The Schwarz inequality enables us to define the cosine of α since as we have pointed out:

$$|\vec{x}\cdot\vec{y}| \leq ||\vec{x}||\cdot||\vec{y}|| \text{ implies: } -1 \leq \frac{\vec{x}\cdot\vec{y}}{||\vec{x}||\cdot||\vec{y}||} \leq 1.$$

<u>Definition</u>: Assuming that $\vec{x},\vec{y} \in \mathbb{R}^n$ are two non-zero vectors:

1. The **angle** α between $\vec{x},\vec{y} \in \mathbb{R}^n$ is such that $\cos\alpha = \frac{\vec{x}\cdot\vec{y}}{||\vec{x}||\cdot||\vec{y}||}$.
2. The two vectors $\vec{x},\vec{y} \in \mathbb{R}^n$are said to be **orthogonal** (or **perpendicular**) if the angle that they form is a right angle. In other words:

$$\vec{x}\cdot\vec{y} = 0.$$[16]

Note that the cosine allows us to define the scalar product using the norm. Indeed:

$$\vec{x}\cdot\vec{y} = ||\vec{x}||\cdot||\vec{y}||\cdot\cos\alpha.$$[17]

<u>Example</u>: Find the angle formed by $\vec{x} = (1,1,0)$ and $\vec{y} = (2,9,6)$.

SOLUTION: As we have:

$$\cos\alpha = \frac{\vec{x}\cdot\vec{y}}{||\vec{x}||\cdot||\vec{y}||} = \frac{11}{\sqrt{2}\cdot\sqrt{121}} = \frac{1}{2}$$

then:

$$\alpha = \cos^{-1}\left(1/2\right) = 45^\circ \blacksquare$$[18]

[16] This is a consequence to the fact that the cosine of a right angle is 0. Note that the two first vectors that appear in the exemple on page 11 are orthogonal.
[17] Thus, the scalar product shows the "tendency" of two vectors to point in the same direction.
[18] Or $\alpha = \pi/_{a}$ radians. The function $\cos^{-1}(x)$ gives us the angle α with cosine equal to x.

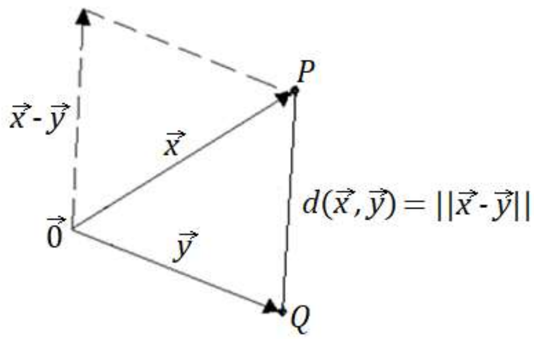
1.2.4. *Distance between Vectors*

Another basic metric concept is the distance between two vectors. This concept will enable us to rigorously define the distance between points related to vector spaces as graphically illustrated bellow.[19]

Definition: The **distance** between two vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$ is defined by the norm of their difference:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = +\sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Graphically:



In this case we can also talk about the distance between the ending points P and Q of the associated position-vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Example: Calculate the distance between the two vectors:

$$\vec{x} = (3, -2, 0, 1) \text{ and } \vec{y} = (1, -4, 0, 2)$$

SOLUTION: In this case we have:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \|(2, 2, 0, -1)\| = +\sqrt{2^2 + 2^2 + 0^2 + (-1)^2} = +\sqrt{9} = 3 \blacksquare$$ [20]

[19] See the first footnote of paragraph 1.2.1.

[20] Consequently the distance between the points $P = (3, -2, 0, 1)$ and $Q = (1, -4, 0, 2)$ is equal to 3.

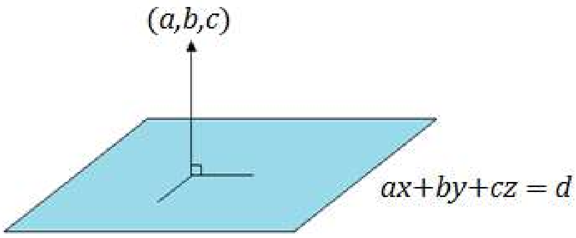
1.2.4.1. Application in Geometry: Hyperplans

An interesting mathematical object with important applications in Economics is the concept of hyperplane.[21] A **hyperplane** in $\mathbb{R}^n$ is the set of all points $(x_1, \dots, x_n) \in \mathbb{R}^n$ satisfying a linear equation like this:

$$a_1 \cdot x_1 + \cdots + a_n \cdot x_n = b$$

with the scalars $a_1, \dots, a_n \in \mathbb{R}$ as a coefficients and $b \in \mathbb{R}$ as a constant term. [22]

An interesting property of hyperplanes is that the vector of coefficients $(a_1, \dots, a_n) \in \mathbb{R}^n$ is always orthogonal to the hyperplane. Graphically:



Example: Find the equation of the hyperplane (plane) perpendicular to the vector $(5, -2, 1)$ passing through the point $(2,0,1)$.

SOLUTION:

Using the above-mentioned property, the equation of the plane orthogonal to the vector $(5, -2, 1)$ is a linear equation of the form:

$$5x - 2y + z = b.$$

We must find the value of the constant term. Bearing in mind that point $(x, y, z) = (2,0,1)$ lies on the plane we have that it must satisfy the above equation. So, the constant term is:

$$5 \cdot 2 - 2 \cdot 0 + 1 = b \text{ implies: } b = 11.$$

Hence the equation of the plane is:

$$5x - 2y + z = 11 \blacksquare$$

[21] Linear Programing for instance.

[22] Straight lines in $\mathbb{R}^2$ and planes in $\mathbb{R}^3$ are the corresponding hyperplanes.

1.3. Quadratic Forms on Vector Spaces

Basically, we are interested in quadratic forms due to the sufficient condition of the existence of extreme points of functions of several variables.[23] For example, the mapping defined on $\mathbb{R}^3$ by the quadratic expression:

$$Q(x,y,z) = 3x^2 + 4y^2 + 3z^2 + 2xy - 2yz$$

is an example of quadratic form in $\mathbb{R}^3$.[24] Notice that if we consider the symmetric matrix defined by:[25]

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix}$$

where the entries of the main diagonal a_{ii} are the coefficients of the square variables x^2, y^2, z^2, and a_{ij} are half of the coefficients of the crossing products xy, xz, yz, then the following matricial product delivers us the initial quadratic form. Indeed:

$$(x,y,z)\cdot A \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x,y,z)\cdot \begin{pmatrix} 3 & 1 & 0 \\ 1 & 4 & -1 \\ 0 & -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x,y,z)\cdot \begin{pmatrix} 3x+y \\ x+4y-z \\ -y+3z \end{pmatrix} =$$

$$= x\cdot(3x+y) + y\cdot(x+4y-z) + z\cdot(-y+3z) =$$

$$= 3x^2 + 4y^2 + 3z^2 + 2xy - 2yz = Q(x,y,z).$$

Thus, in this case, we can say that the quadratic form has the matrix A as its “associated” symmetric matrix. When we put together this we have:

Definition: A mapping $Q(x_1, \dots, x_n)$ defined on $\mathbb{R}^n$ is a **quadratic form** provided that a symmetric matrix $A = (a_{ij})$ of order n exists such that:

$$Q(x_1, \dots, x_n) = (x_1, \dots, x_n)\cdot A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i,j=1}^{n} a_{ij}\cdot x_i \cdot x_j\,.$$ [26]

---

[23] We return to this question in the next section.

[24] Note that the sum of the exponents of the variables in each term is 2.

[25] Recall that a matrix is symmetric when it is equal to its associated transpose matrix.

[26] Observe that the two coefficients a_{ij} and a_{ji} must be added in the expresión of the quadratic form $Q(x_1, \dots, x_n)$ since $x_i \cdot x_j = x_j \cdot x_i$. And reverse, this is the reason why the coefficients of the cross products of $Q(x_1, \dots, x_n)$ must be divided by two prior to put them in the matrix A.

1.3.1. *Sign of a Quadratic Form*

What is of interest to us about this issue is basically the sign of quadratic forms and in this regard, we need to bear in mind the following definition:

Definition: The quadratic form $Q(x_1, \dots, x_n)$ is:

1. **Positive definite** if $Q(x_1, \dots, x_n) > 0$ for any non-zero vector $(x_1, \dots, x_n) \in \mathbb{R}^n$.[27]
2. **Negative definite** if $Q(x_1, \dots, x_n) < 0$, for any non-zero vector $(x_1, \dots, x_n) \in \mathbb{R}^n$.
3. **Positive semidefinite** if $Q(x_1, \dots, x_n) \geq 0$, for any vector $(x_1, \dots, x_n) \in \mathbb{R}^n$. Besides one non-zero vector $(y_1, \dots, y_n) \in \mathbb{R}^n$ must exist such that $Q(y_1, \dots, y_n) = 0$.
4. **Negative semidefinite** if $Q(x_1, \dots, x_n) \leq 0$, for any vector $(x_1, \dots, x_n) \in \mathbb{R}^n$. Besides one non-zero vector $(y_1, \dots, y_n) \in \mathbb{R}^n$ must exist such that $Q(y_1, \dots, y_n) = 0$.
5. **Indefinite** if it does not match any of the definitions listed above, i.e., when two vectors $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$ exist such that $Q(x_1, \dots, x_n) < 0 < Q(y_1, \dots, y_n)$.

Example: Prove that the quadratic form:

$$Q(x, y, z) = x^2 + y^2 + z^2 - 2xy$$

is positive semidefinite.

SOLUTION:

This is clear because this quadratic form is always positive (greater than zero) if we rewrite its expression in this way:

$$Q(x, y, z) = x^2 + y^2 + z^2 - 2xy = (x - y)^2 + z^2 \geq 0$$

taking into account, for instance, that the image of the non-zero vector (1,1,0) is 0:

$$Q(1,1,0) = 1^2 + 1^2 + 0^2 - 2 \cdot 1 \cdot 1 = 0 \blacksquare$$

Determining the sign of a quadratic form using the above definition is not an easy task. Thus, we are going to introduce a relatively simple method for doing so in the easiest cases.

[27] Note that the corresponding image of the zero-vector $\vec{0}$ by any quadratic form is always equal to 0. So the zero-vector must be dropped out of those definitions.

1.3.2. *Classification of Quadratic Forms by means of Principal Minors*

First, we need to recall the definition of principal minor of a matrix. Indeed, the **principal minors** of a square matrix of the type:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

are, precisely, the minors having a main diagonal formed by some of the n coefficients:

$$a_{11}, \dots, a_{nn}$$

of the main diagonal of matrix A.

Let us consider the following example:

Example: Calculate the principal minors of the symmetric matrix:

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

SOLUTION:

First, we have the three first-order principal minors, Mp_1, that following the above definition match the three entries of the main diagonal of A. Indeed:

$$Mp_1 = \{-5; -2; -3\}■$$

Note that the three second-order principal minors, Mp_2, are:

$$Mp_2 = \left\{\begin{vmatrix} -5 & 3 \\ 3 & -2 \end{vmatrix}; \begin{vmatrix} -2 & 0 \\ 0 & -3 \end{vmatrix}; \begin{vmatrix} -5 & 1 \\ 1 & -3 \end{vmatrix}\right\} = \{1; 6; 14\}■$$

Finally as the matrix A is a third order matrix, the sole third-order principal minor Mp_3 is precisally its determinant:

$$Mp_3 = \{A\} = \left\{\begin{vmatrix} -5 & 3 & 1 \\ 3 & -2 & 0 \\ 1 & 0 & -3 \end{vmatrix}\right\} = \{-1\}■$$

1.3.2.1. Relationship between the Sign of Quadratic Forms and Principal Minors

The theorem we enunciate below characterizes the sign of a quadratic form through the sign of the principal minors associated with the related matrix. Indeed:

Theorem: A quadratic form $Q(x_1, \dots, x_n)$ with associated symmetric matrix A is:

1. Positive definite if and only if all the principal minors of A are strictly positive.
2. Negative definite if and only if all the principal minors of even order of A are strictly positive and those of odd order are strictly negative.
3. Positive semidefinite if and only if all the principal minors of A are positive being always $|A| = 0$.
4. Negative semidefinite if and only if all the principal minors of even order of A are positive and those of odd order are negative or zero being always $|A| = 0$.[28]
5. Indefinite otherwise.

Example: Calculate both the analytic expression $Q(x, y, z)$ and the sign of the quadratic form with associated symmetric matrix:

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

SOLUTION:

According to the definition of quadratic form, the analytic expression demanded will be:

$$Q(x, y, z) = -5x^2 - 2y^2 - 3z^2 + 6xy + 2xz \blacksquare$$

Since the principal minors are:

$$Mp_1 = \{-5; -2; -3\}, Mp_2 = \{1; 6; 14\} \text{ and: } Mp_3 = \{-1\}$$

we can conclude according to the previous theorem that this quadratic form is negative definite$\blacksquare$

[28] Note that a necessary condition for a quadratic form to be semidefinite (either positive or negative) is that the determinant of the associated matrix A is equal to 0.

1.3.2.2. Application of Quadratic Forms in Economics

Let us consider this situation in terms of Economics:

Example: A company dedicated to develop three types of wine has annual revenues given by the function:

$$B(x,y,z) = x^2 + y^2 + 2z^2 - 2xy - 2\sqrt{2}xz$$

where $x, y, z \geq 0$ denote respectively the amount of hectoliters of each type of wine elaborated. Under these assumptions prove that: (1) The company may have losses. (2) This is not the case if the number of hectoliters of the first type is half of the third.

SOLUTION: (1) First of all, obseve that the function of revenues is a quadratic form. Since the associated symmetric matrix is:

$$A = \begin{pmatrix} 1 & -1 & -\sqrt{2} \\ -1 & 1 & 0 \\ -\sqrt{2} & 0 & 2 \end{pmatrix}$$

and the principal minors are:

$$Mp_1 = \{1; 1; 2\}, Mp_2 = \{0; 2; 0\} \text{ and: } Mp_3 = \{-2\}$$

the function of revenues is indefinite according to the previous results. Following the definition of indefinite quadratic forms, we might conclude that this company may have losses. In fact, this is the case when the production of the three types of wine is the same. Indeed:

$$B(x,x,x) = x^2 + x^2 + 2x^2 - 2x \cdot x - 2\sqrt{2}x \cdot x = \left(2 - 2\sqrt{2}\right) \cdot x^2 < 0 \text{ for all } x \neq 0 \blacksquare$$

(2) In this case being:

$$x = {}^{z}/_{2} \text{ implies: } z = 2x$$

the function of revenues would be the "restricted" quadratic form on two variables:

$$B(x,y,2x) = x^2 + y^2 + 2(2x)^2 - 2xy - 2\sqrt{2}x(2x) = \left(9 - 4\sqrt{2}\right)x^2 + y^2 - 2xy.$$

Since the associated matrix is:

$$A = \begin{pmatrix} 9 - 4\sqrt{2} & -1 \\ -1 & 1 \end{pmatrix}$$

and the principal minors are strictly positive:

$$Mp_1 = \left\{9 - 4\sqrt{2}; 1\right\} \text{ and: } Mp_2 = \{|A|\} = \left\{8 - 4\sqrt{2}\right\}$$

the restricted quadratic form is definite positive, which means that now the company does not have losses$\blacksquare$

1.4. Exercises

1. Find the values of the parameter $a \in \mathbb{R}$ so that the vector $(-2,-1,5,0)$ is a linear combination of $(2,4,7,6)$ and $(a, 2, -1, a)$.
2. Find the values of $a \in \mathbb{R}$ so that $(a, 0, -3)$, $(2, -a, 5)$ and $(0,1, a)$ form a basis of $\mathbb{R}^3$.[29]
3. Prove that the vectors $(-1,0,4,3)$, $(6,5,0,3)$ and $(0,-2,1,0)$ are linearly independent and find another vector which when taken altogether form a basis of $\mathbb{R}^4$.
4. Given the vectors $\vec{u} = (-m^2, -m^2, 0)$, $\vec{v} = (1,1,-1)$ and $\vec{w} = (m, 1, -m)$ find the values of $m \in \mathbb{R}$ so that: (a) $\vec{u}$ can be expressed as a linear combination of $\vec{v}$ and $\vec{w}$. (b) These vectors are linearly dependent.
5. Given the vectors $\vec{u} = (2,1,1)$ and $\vec{v} = (3,-2,2)$ determine:
    a. Their scalar product.
    b. The associated norms.
    c. The associated unit vectors.
    d. The angle that they form.
    e. The distance between them.
6. Given the vectors $\vec{x} = (k, -k, 0)$ and $\vec{y} = (1,0,2k)$:
    a. Prove that despite the values of $k \in \mathbb{R}$ these vectors cannot be orthogonal.
    b. Find the values of $k \in \mathbb{R}$ such that these vectors form an angle of $\pi/3$ radians.
    c. Idem so that the distance between them is equal to 3.
7. Prove that two non-zero orthogonal vectors are always linearly independent.
8. Find the sign of the following quadratic forms:
    a. $Q(x, y, z) = 2x^2 + 3y^2 + 2z^2 + 2xz$.
    b. $Q(x, y, z) = x^2 + 2y^2 + z^2 - 2xy + 2yz$.
    c. $Q(x, y, z) = -3x^2 - 4y^2 - 3z^2 + 4\sqrt{2}xy + 4xz$.
    d. $Q(x, y, z) = -x^2 - 14y^2 - 2z^2 + 4xz$ restricted to the hyperplane: $x + y + z = 0$.[30]

[29] For a definition of basis see the footnote of paragraph 1.1.3.
[30] In this case we must obtain one "restricted" quadratic form depending on two variables.

SOLUTIONS:

1. $a = 2$.
2. $a \neq -1$.
3. For example, (0,0,0,1).
4.
   a. $m = 0$.
   b. Either $m = 0$ or $m = 1$.
5.
   a. $\vec{u} \cdot \vec{v} = 6$.
   b. $\|\vec{u}\| = +\sqrt{6}$ and: $\|\vec{v}\| = +\sqrt{17}$.
   c. $\vec{u}_0 = \left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$ and: $\vec{v}_0 = \left(\frac{3}{\sqrt{17}}, -\frac{2}{\sqrt{17}}, \frac{2}{\sqrt{17}}\right)$.
   d. $\alpha \cong 0{,}935$ radians.
   e. $d(\vec{u}, \vec{v}) = +\sqrt{11}$.
6.
   b. Either $k = -0.5$ or $k = 0.5$.
   c. Either $k = -1$ or $k = \frac{4}{3}$.
8.
   a. Positive definite.
   b. Positive semidefinite.
   c. Indefinite.
   d. Negative definite.

2. MULTIVARIABLE OPTIMIZATION

This section is mainly devoted to learning how to obtain maxima and minima (optimal points) of functions of several variables.[31] In the sequel we are going to take for granted that the lector knows how to manage functions of one-real variable on this issue. Obviously, we must start introducing the notion of function of several variables and viewing some of the basic properties that we have bear in mind in order to accomplish the task mentioned. Consider the following situation in terms of Economics. Let:

$$C = C(q_1, \dots, q_n)$$

be the total cost of a firm that produces n articles, $A_1, \dots, A_n$, in quantities $q_1, \dots, q_n \geq 0$. According to this, suppose that these articles are sold in a market generating as a total income of:

$$I(q_1, \dots, q_n) = p_1 \cdot q_1 + \dots + p_n \cdot q_n = \sum_{i=1}^{n} p_i \cdot q_i$$

where $p_1, \dots, p_n > 0$ stand for the sale price of $A_1, \dots, A_n$. An interesting problem in Economics appears when we want to know the production of the n articles that maximizes benefits, i.e., total income minus total cost:

$$B(q_1, \dots, q_n) = I(q_1, \dots, q_n) - C(q_1, \dots, q_n) = \left(\sum_{i=1}^{n} p_i \cdot q_i\right) - C(q_1, \dots, q_n).$$

For example, suppose that the total cost of the production of two articles A and B were:

$$C(q_1, q_2) = 3q_1 \cdot q_2 + 1.5q_1^2 + 2q_2^2$$

being the two sale prices constant and equal to:[32]

$$p_1 = €42 \text{ and: } p_2 = €51.$$

It easy to see that benefits (or profit) adopt the form:

$$B(q_1, q_2) = 42q_1 + 51q_2 - 3q_1q_2 - 1.5q_1^2 - 2q_2^2.$$

It can be proved that $q_1 = 5$ units of A and $q_2 = 9$ units of B maximize benefits with a total amount of $B(5,9) = €334.5$.[33] As mentioned, this section will start analyzing some of the most important concepts related to functions of several variables such as partial derivatives and Hessian matrices. These mathematical objects are the main tools in the search for optimal points of this class of functions.

---

[31] If they exist.

[32] In this scenario we are dealing with a monopolistic market.

[33] See page 45.

2.1. Function of Several Variables

2.1.1. *Function of Several Variables and Domain*

The basic structure in which we are going to move is the vector space $\mathbb{R}^n$ associated with the scalar product defined above, i.e., the Ecuclidian vector space $\mathbb{R}^n$. Formally:

<u>Definition</u>: A **function of several variables** (**function** for short) is a mapping f that assigns, obviously in the case it makes sense, a real number $z \in \mathbb{R}$ to each point or vector $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$:

$$z = f(\vec{x}) = f(x_1, \dots, x_n) \in \mathbb{R}.$$

So, by definition, we must consider the set $Domf \subset \mathbb{R}^n$ formed by the points of $\mathbb{R}^n$ supporting image by f. This set is called the **domain** of the function f.

Note that if $n = 1$ a function of several variables is simply a function of one real variable:

$$y = f(x) \in \mathbb{R}.$$

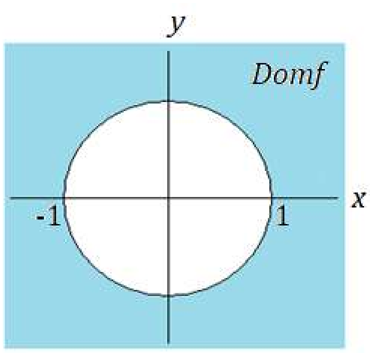
<u>Example</u>: Determine graphically the domain of the function of two variables:

$$z = f(x, y) = +\sqrt{x^2 + y^2 - 1}.$$

SOLUTION: It is easy to see that the domain of this function is the set of points:

$$Domf = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \geq 1\}.$$

Observe that the domain is the set of the plane formed by the points that lie either on the circumference centerd at (0,0) and radius $r = 1$ or outside of it■ [34] Graphically:



[34] Recall that any equation of the form: $(x - a)^2 + (y - b)^2 = r^2$ represents the circumference of radius $r > 0$ centered at point (a, b).

2.1.1.1. Example of Graphical Representation of Domains

From a practical point of view, we only will graph domains of functions of two variables. This is the first task that we must learn.[35]

Let us now see another example:

Example: Determine graphically the domain of the function:

$$z = f(x, y) = \frac{x^2 y}{x^2 - y^2}.$$

SOLUTION:

Taking into account that both the numerator and the denominator regardlees the values of x and y always exist, the only condition that we have to impose here is:

$$x^2 - y^2 \neq 0.$$[36]

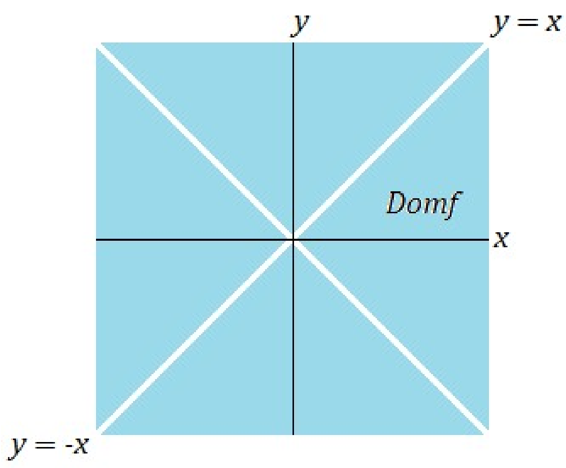
Since:

$$x^2 - y^2 = (x - y)(x + y) \neq 0 \text{ is equivalent to: } y \neq x \text{ and: } y \neq -x$$

we conclude that the domain is formed by the set of points of the plane that lie outside the two bisectors of the reference system:

$$Domf = \{(x, y) \in \mathbb{R}^2 : y \neq x \text{ and: } y \neq -x\}$$■

Graphically:



[35] We recomend to review the graphic representation of lines, parabolas and hyperbolas.
[36] Recall that dividing by zero is forbbiden.

2.1.2. *Level Curves of a Function of Two Variables*

We know that any function of a single variable can be represented by a "curve" in the plane. However, in the case of two variables, the "surfaces" that functions of this type generate in the space are very difficult to draw. In practice what happens is that this problem is simplified by considering only the curves that join those points of the domain that have the same value by the function. It should be stressed that these curves provide a lot of information when drawing functions is not a straightforward task. Formally:

Definition: The **level curves** of $z = f(x, y)$ are the plane-curves defined by the equation:

$$f(x, y) = k \in \mathbb{R}$$

and are denoted by c_k.[37]

Example: Represent graphically the level curves of the function:

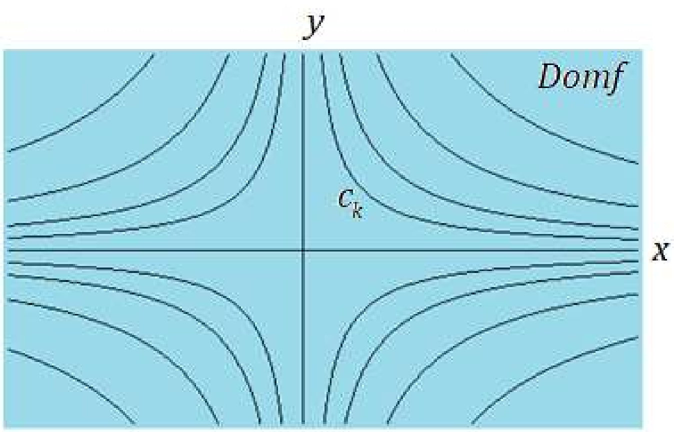
$$z = f(x, y) = x \cdot y.$$

SOLUTION:

The level curves c_k are the plane curves provided by the equation:

$$x \cdot y = k.$$

Note that these curves are in fact hyperboles having the two reference axes as asymptotes. For any $k > 0$ the hyperbola appears in the first and third quadrants, whereas for $k < 0$ it does in the second and fourth. Graphically:



■

[37] In Economics, the level curves associated with production functions are called *isoquants*, while those associated with utility functions are called *indifference curves*.

2.1.2.1. Application of Level Curves in Economics: Maximal Income

Let us consider the next example in Economics on which it is possible to calculate the maximum income from the sole knowledge of the associate level curves.

Example: Let:

$$q_1 = 60 - 5p_1 \text{ and } q_2 = 20 - p_2$$

the demand functions of a product P in two markets, being $q_1 \geq 0$ the quantity demanded of P in the first market at a unit price of $p_1 \geq 0$ and $q_2 \geq 0$ the corresponding quatity of P demanded in the second one at a unit price of $p_1 \geq 0$. Determine the total maximum income of P as well as the relating prices and quantities.

SOLUTION:

The total income obtained when P is sold in both markets is the function depending on the two prices $p_1 \geq 0$ and $q_2 \geq 0$ as variables:

$$I(p_1, p_2) = p_1 \cdot q_1 + p_2 \cdot q_2 =$$

$$= p_1 \cdot (60 - 5p_1) + p_2 \cdot (20 - p_2) = 60p_1 + 20p_2 - 5p_1^2 - p_2^2$$

and the k-level c_k curve associated is:

$$60p_1 + 20p_2 - 5p_1^2 - p_2^2 = k.$$

The main question here is to figure out the maximum value that parameter $k \geq 0$ can reach. In order to solve this, we need to rewrite the income function in a proper maner. It can be proved that:

$$60p_1 + 20p_2 - 5p_1^2 - p_2^2 = k$$

is equivalent to the equation:

$$5(p_1 - 6)^2 + (p_2 - 10)^2 = 280 - k.$$

Since the left-side of this equation is always positive, it is straigtforward to be aware that the maximum value of $k \geq 0$ is that of 280. Hence:

$$k = 280 \text{ implies: } 5(p_1 - 6)^2 + (p_2 - 10)^2 = 0 \text{ implies: } p_1 = 6 \text{ and } p_2 = 10.$$

and:

$$p_1 = 6 \text{ and } p_2 = 10 \text{ implies: } q_1 = 60 - 5 \cdot 6 = 30 \text{ and: } q_2 = 20 - 10 = 10.$$

So, in order to maximize the total income of 280 it is necessary to sold 30 units of P at a unit price of 6 in the first market, and 10 units of P at a unit price of 10 in the second■

2.2. Partial Derivatives of Functions

2.2.1. *Partial Derivatives of a Function of Several Variables*

Similarly to the way in which we define the derivative of a function of a single variable:

$$\lim_{t\to 0}\frac{f(a+t)-f(a)}{t}=f'(a)\in\mathbb{R}$$

the same procedure can be adopted with functions of several variables. This gives raise to the concept of the partial derivative. Formally:

<u>Definition</u>: The function $z=f(x_1,\dots,x_n)$ has **partial derivative with respect to the variable $\boldsymbol{x_i}$** at a point $\vec{a}=(a_1,\dots,a_n)\in Domf$ of its domain in the case of the existence of the limit: [38]

$$\lim_{t\to 0}\frac{f(a_1,\dots,a_i+t,\dots,a_n)-f(a_1,\dots,a_i,\dots,a_n)}{t}=\frac{\partial f(\vec{a})}{\partial x_i}\in\mathbb{R}.$$

Thus, given a function of n variables and a point from its domain, we have in principle n partial derivatives at this point. Obviously, their existence will depend on the existence of the corresponding limits. Let us consider the following example:[39]

<u>Example</u>: Calculate the partial derivatives at point $(1{,}1)$ of the function:

$$z=f(x,y)=\frac{y}{x}.$$

SOLUTION: In this case $\vec{a}=(1{,}1)\in Domf$ and, hence, we have:

$$\frac{\partial f(1{,}1)}{\partial x}=\lim_{t\to 0}\frac{f(1+t,1)-f(1{,}1)}{t}=\lim_{t\to 0}\frac{\left(\frac{1}{1+t}\right)-\frac{1}{1}}{t}=\lim_{t\to 0}\left(\frac{-t}{t(1+t)}\right)=\lim_{t\to 0}\left(\frac{-1}{1+t}\right)=-1\ ■$$

and:

$$\frac{\partial f(1{,}1)}{\partial y}=\lim_{t\to 0}\frac{f(1{,}1+t)-f(1{,}1)}{t}=\lim_{t\to 0}\frac{\left(\frac{1+t}{1}\right)-\frac{1}{1}}{t}=\lim_{t\to 0}\left(\frac{t}{t}\right)=\lim_{t\to 0}1=1\ ■$$ [40]

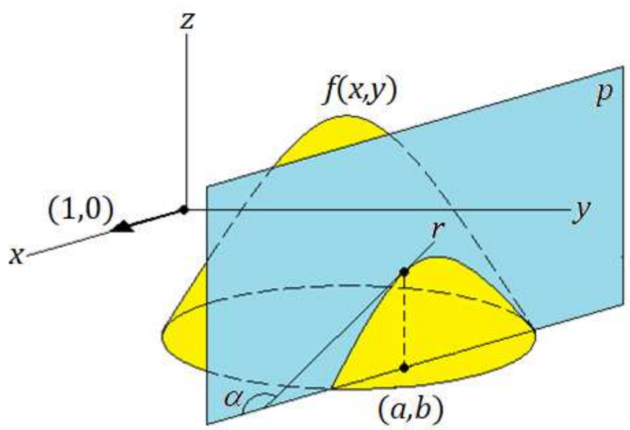
[38] Take care of this symbolism of partial derivatives.

[39] However we will calculate partial derivatives through the one-variable derivative rules that we already know. See the example on page 31.

[40] Note that the partial derivatives of a given function evaluated at a point of its domain do not necessarly have to match each other.

2.2.1.1. Geometrical Interpretation of Partial Derivatives

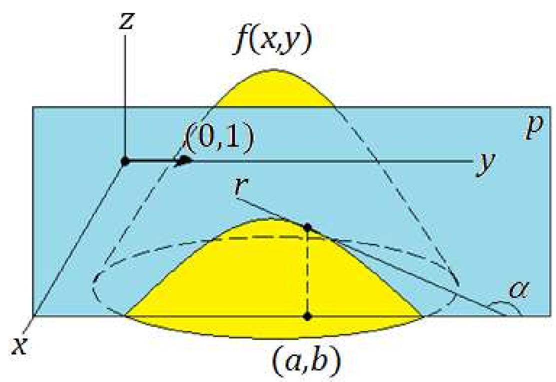
Geometrically speaking partial derivatives also measure slopes of tangent lines in a specific manner as we are going to see. Indeed, the partial derivative with respect to x of the function $z = f(x, y)$ at point $(a, b) \in Domf$ is the slope of the tangent r on the plane p determined as it is shown below:



In other words:

$$\frac{\partial f(a,b)}{\partial x} = \tan \alpha.$$

Similarly, the partial derivative with respect to y of the function $z = f(x, y)$ at point $(a, b) \in Domf$ is the slope of the tangent r on the plane p:



In other words:

$$\frac{\partial f(a,b)}{\partial y} = \tan \alpha.$$

2.2.1.2. Calculation of Partial Derivatives

From the definition of partial derivatives, it is easy to see that the calculus of partial derivatives can be reduced to that of the derivatives in a single variable. This is due to the fact that the partial derivative with respect to any specific variable is obtained by "deriving" the function (as we already know) respect to this variable being constant the remaining variables. Let us look at an example:

Example: Calculate the partial derivatives of the following functions:

1. $z = f(x, y) = 4x^2 - 7y^3 + 2x^2y - 75y + 8$.
2. $z = f(x, y) = \frac{1}{2} \cdot \ln\left(\frac{x-y}{x+y}\right)$.
3. $z = f(x, y) = x^y$.

SOLUTION: (1) In this case the partial derivative with respect to x is:

$$\frac{\partial z}{\partial x} = \{y \text{ is constant}\} = 8x - 0 + 4x \cdot y - 0 + 0 = 8x + 4xy \blacksquare$$

and that of corresponding to y will be:

$$\frac{\partial z}{\partial y} = \{x \text{ is constant}\} = 0 - 21y^2 + 2x^2 \cdot 1 - 75 \cdot 1 + 0 = -21y^2 + 2x^2 - 75 \blacksquare$$

(2) Now the two partial derivatives will be applying the chain rule:

$$\frac{\partial z}{\partial x} = \{y \text{ is constant}\} = \frac{1}{2} \cdot \left(\frac{1}{\left(\frac{x-y}{x+y}\right)}\right) \cdot \left(\frac{1 \cdot (x+y) - (x-y) \cdot 1}{(x+y)^2}\right) = \frac{y}{x^2 - y^2} \blacksquare$$

and:

$$\frac{\partial z}{\partial y} = \{x \text{ is constant}\} = \frac{1}{2} \cdot \left(\frac{1}{\left(\frac{x-y}{x+y}\right)}\right) \cdot \left(\frac{(-1) \cdot (x+y) - (x-y) \cdot 1}{(x+y)^2}\right) = \frac{-x}{x^2 - y^2} \blacksquare$$

(3) In this case we have:

$$\frac{\partial z}{\partial x} = \{y \text{ is constant}\} = y \cdot x^{y-1} \blacksquare$$

and:

$$\frac{\partial z}{\partial y} = \{x \text{ is constant}\} = x^y \cdot \ln x \blacksquare$$

2.2.2. *Applications of Partial Derivatives*

2.2.2.1. Application in Economics I: Marginalism

Let $z = f(x_1, \dots, x_n)$ be for example an economic function[41] having partial derivative with respect to x_i at point $\vec{a} = (a_1, \dots, a_n) \in Domf$ of its domain, i.e:

$$\lim_{t \to 0} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n)}{t} = \frac{\partial f(\vec{a})}{\partial x_i} \in \mathbb{R}.$$

If we consider $t = 1$ in this equality, we would obtain the approximation:

$$f(a_1, \dots, a_i + 1, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n) \cong 1 \cdot \frac{\partial f(\vec{a})}{\partial x_i} = \frac{\partial f(\vec{a})}{\partial x_i}.$$ [42]

Thus, this partial derivative approximately measures the change in the function value $f(a_1, \dots, a_i, \dots, a_n)$ caused by an increase of 1 unit in the component $a_i \in \mathbb{R}$.

As an example consider this application in terms of Economics:

Example: Let:

$$B(x, y) = 1100x + 1300y - 2x^2 - 2.5y^2 - 70{,}000$$

be a function of the benefit in euros associated with the production and sale of two articles A and B. If currently $x = 250$ and $y = 220$ units are produced, estimate the change in the benefit if the production of B is increased by one unit.

SOLUTION: In order to estimate the change in the benefit we need to consider the partial derivative of the function with respect to the variable y at the point $(x, y) = (250{,}220)$. Since:

$$\frac{\partial B}{\partial y} = 1300 - 5y \text{ implies: } \frac{\partial B(250{,}220)}{\partial y} = 1300 - 5 \cdot 220 = 200$$

we deduce that the benefit increases by approximately €200.[43] Indeed:

$$\Delta B = B(250{,}221) - B(250{,}220) \cong 1 \cdot \frac{\partial B(250{,}220)}{\partial y} = 200 \blacksquare$$

[41] For instance, a function of benefits, of costs, of utility, etc.

[42] In principle we can consider others values than $t = 1$ but the greater in absolute value they be the poorer is the approximation.

[43] This approximation works quite well since the real value of this increase is equal to €197.5 as can be easily figure it out.

2.2.2.2. Application of in Economics II: Partial Elasticity

If $z = f(x_1, \dots, x_n)$ is an economic function having partial derivative with respect to x_i at point $\vec{a} = (a_1, \dots, a_n) \in Domf$ such that $f(\vec{a}) \neq 0$, we can always define the **partial elasticity with respect to the variable** x_i as the number:

$$E_{x_i} f(\vec{a}) = \frac{a_i}{f(\vec{a})} \cdot \frac{\partial f(\vec{a})}{\partial x_i} \in \mathbb{R}.$$

It can be proved that the value of $E_{x_i} f(\vec{a})$, taken as a percentage, is approximately equal to the percentage change in the function value $f(\vec{a}) = f(a_1, \dots, a_n)$ caused by a 1% increase in the component $a_i \in \mathbb{R}$. Moreover, if this component increases (or decreases) at a r% rate, the value $f(\vec{a}) = f(a_1, \dots, a_n)$ varies approximately by:

$$\left(r \cdot E_{x_i} f(\vec{a})\right)\%.$$[44]

Example: Let:

$$B(x, y) = 1100x + 1300y - 2x^2 - 2.5y^2 - 70{,}000$$

be the precedent function of the benefit in euros associated with the production and sale of two articles A and B. If currently $x = 250$ and $y = 220$ units are produced, estimate the percentage of increasing (or decrasing) in benefits if the production of A increases by 2% from the present level of 250 units.

SOLUTION: Note that the partial elasticity of the benefit function respect to x is:

$$E_x B(x, y) = \frac{x}{B(x, y)} \cdot \frac{\partial B}{\partial x} = \frac{x}{B(x, y)} . (1100 - 4x) = \frac{1100x - 4x^2}{B(x, y)}.$$

Since:

$$B(250{,}220) = 245{,}000$$

the partial elasticity of the benefit function respect to x at point $(x, y) = (250{,}220)$ is:

$$E_x B(250{,}220) = \frac{1100 \cdot 250 - 4 \cdot 250^2}{245{,}000} = 0.1020408.$$

So, the percentage of change in benefits when the production of A increases by $r = 2\%$ is approximately equal to:

$$\left(2 \cdot E_x B(250{,}220)\right)\% = (2 \cdot 0.1020408)\% = 0.20408\% \blacksquare$$[45]

[44] The increase or decrease of the function depends on the sign of the corresponding partial elasticity.

[45] In fact the real percentatge of change is equal to 0.183%.

2.2.3. *Gradient of a Function*

We are ready to introduce the seminal concept of the vector gradient of a function of several variables at a point of its domain. Formally:

Definition: The **gradient vector** (**gradient** for short) of the function $z = f(x_1, \dots, x_n)$ at a point $\vec{a} = (a_1, \dots, a_n) \in Domf$ of its domain is the vector formed by all the partial derivatives of the function at this point. Namely:

$$\nabla f(\vec{a}) = \left(\frac{\partial f(\vec{a})}{\partial x_1}, \dots, \frac{\partial f(\vec{a})}{\partial x_n}\right) \in \mathbb{R}^n.$$

We must stress that the gradient associated with a function of n variables is a vector with n components. Hence, if any of the partial derivatives appearing in the definition do not exist, neither does the gradient. Let us see an example:

Example: Calculate, in the case that it does exist, the gradient at point $(1,0)$ of the following functions:

1. $z = f(x,y) = \frac{y}{x}$.
2. $z = f(x,y) = +\sqrt{xy}$.

SOLUTION:

(1) In this case the gradient does exist and it is the vector:

$$\left.\begin{array}{l} \dfrac{\partial f}{\partial x} = -\dfrac{y}{x^2} \\ \dfrac{\partial f}{\partial y} = \dfrac{1}{x} \end{array}\right\} \text{implies: } \nabla f(1,0) = \left(\frac{\partial f(1,0)}{\partial x}, \frac{\partial f(1,0)}{\partial y}\right) = \left(-\frac{0}{1^2}, \frac{1}{1}\right) = (0,1) \blacksquare$$

(2) Here, the gradient does not exist since there no is partial derivative respect to y at this point. Indeed:

$$\frac{\partial f}{\partial y} = \frac{1}{2\sqrt{xy}} \cdot x = \frac{1}{2} \cdot \sqrt{\frac{x}{y}} \text{ implies: } \frac{\partial f(1,0)}{\partial y} = \frac{1}{2} \cdot \sqrt{\frac{1}{0}} \notin \mathbb{R} \blacksquare$$ [46]

[46] The partial derivative respect to x exists and it is equal to 0.

2.2.3.1. Application of the Gradient: Maximal Increase of a Function

Among the main properties of the gradient of a function we should mention that, as a vector, it always points in the direction of maximal increase of the function.[47] Consequently if we want to maximize functions as fast as possible, we are forced to follow the direction of gradient vectors; conversely, if we wish to minimize, we have to reverse their direction.[48]

Example: Given the function:

$$z = f(x,y) = x^2y - y^3 + x^2 + 1$$

1. Calculate the gradient at points $(1,0)$ and $(3,-1)$.
2. If we want to maximize this function as fast as possible from point $(1,0)$, what would be the direction to follow? What can we say about the point $(3,-1)$?

SOLUTION: (1) Since the gradient of the function is the vector:

$$\nabla f(x,y) = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}\right) = (2xy + 2x, x^2 - 3y^2)$$

the associated gradients at points $(1,0)$ and $(3,-1)$ are:

$$\nabla f(1,0) = (2,1) \text{ and } \nabla f(3,-1) = (0,0) \blacksquare$$

(2) Thus, if we want to maximize the function value as fast as posible from point $(1,0)$, we have to follow the direction of this gradient, i.e., the direction of the vector:

$$\vec{u} = (2,1) \blacksquare$$ [49]

Now, since:

$$\nabla f(3,-1) = (0,0)$$

we cannot say anything about the increase (or decrease) of the function at point $(0,0)$ ■

Consequently, as long as the gradient vector of a function is not the zero-vector the function can be either increased or decreased.

---

[47] This property deals with *differentiable* functions. In fact, all the functions we study here are differentiable. See Sydsaeter, K.; Hammond, P.J. (1995) for a definition of this concept.

[48] In the case the gradient is equal to the zero-vector we cannot say anything in principle. We will discuss this issue in the next topic.

[49] If minimizing is the case, we have to follow the opposite vector.

2.2.4. *Second-Order Partial Derivatives of a Function*

A fundamental concept in the field we are studying is that of the Hessian matrix of a function of several variables. This matrix is formed by the second-order partial derivatives of that function. Hence, we must start by introducing the concept of the second-order partial derivative:

Definition: The **second-order partial derivative** of a function $z = f(x_1, \dots, x_n)$ at a point $\vec{a} = (a_1, \dots, a_n) \in Domf$ of its domain with respect to the variables (x_i, x_j) is the partial derivative with respect to the second variable x_j at this point of the partial derivative function with respect to the first variable x_j. In other words:

$$\frac{\partial^2 f(\vec{a})}{\partial x_i \partial x_j} = \frac{\partial\left(\frac{\partial f}{\partial x_i}\right)(\vec{a})}{\partial x_j}.$$

When $x_i = x_j$ we put:

$$\frac{\partial^2 f(\vec{a})}{\partial x_i^2}.$$

Example: Evaluate at point (1,1) the second-order partial derivatives of $f(x,y) = x \cdot \ln y$.
SOLUTION: Since the gradient is:

$$\nabla f(x,y) = \left(\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y}\right) = \left(\ln y, {}^{x}\!/_{y}\right)$$

we deduce that the four second partial derivatives of the function at point (1,1) are:[50]

$$\left.\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial(\ln y)}{\partial x} = 0 \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial(\ln y)}{\partial y} = {}^{1}\!/_{y} \\ \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial({}^{x}\!/_{y})}{\partial x} = {}^{1}\!/_{y} \\ \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial({}^{x}\!/_{y})}{\partial y} = -{}^{x}\!/_{y^2} \end{aligned}\right\} \text{implica: } \begin{cases} \frac{\partial^2 f(1,1)}{\partial x^2} = 0 \blacksquare \\ \frac{\partial^2 f(1,1)}{\partial x \partial y} = {}^{1}\!/_{1} = 1 \blacksquare \\ \frac{\partial^2 f(1,1)}{\partial y \partial x} = {}^{1}\!/_{1} = 1 \blacksquare \\ \frac{\partial^2 f(1,1)}{\partial y^2} = -{}^{1}\!/_{1^2} = -1 \blacksquare \end{cases}$$

[50] Observe that the value of the two "crossing" second-order partial derivatives $\frac{\partial^2 f(1,1)}{\partial x \partial y}$ and $\frac{\partial^2 f(1,1)}{\partial y \partial x}$ is the same. This crossing second-order partial derivatives coincidence is true within the whole functions that we are studying here.

2.2.4.1. Hessian Matrix of a Function

As we shall see, the concept of the Hessian matrix is of great significance in the context of the optimization of functions of several variables.[51]

Definition: The **Hessian matrix** of $z = f(x_1, \dots, x_n)$ at a point $\vec{a} = (a_1, \dots, a_n) \in Domf$ of its domain is the square matrix defined by:

$$Hf(\vec{a}) = \begin{pmatrix} \frac{\partial^2 f(\vec{a})}{\partial x_1^2} & \frac{\partial^2 f(\vec{a})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\vec{a})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\vec{a})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\vec{a})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\vec{a})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_n^2} \end{pmatrix}.$$

Let us see the following example that completes the one on the previous page:

Example: Evaluate at point $(1,1)$ the Hessian matrix of the function:

$$z = f(x,y) = x \cdot \ln y.$$

SOLUTION:

Applying the results just obtained, wee see that the Hessian matrix of this function at point $(1,1)$ is:

$$Hf(1,1) = \begin{pmatrix} \frac{\partial^2 f(1,1)}{\partial x^2} & \frac{\partial^2 f(1,1)}{\partial x \partial y} \\ \frac{\partial^2 f(1,1)}{\partial y \partial x} & \frac{\partial^2 f(1,1)}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \blacksquare$$ [52]

With all this in mind, we are in a good position to address the main topic of this section, which is none other than the study of the maxima and minima (optima) of functions of several variables. Let us start.

[51] In fact, this concept plays a similar role to that of the second derivative of a function of a single variable.

[52] However if the point changes, so does the Hessian matrix.

2.3. Classical Optimization

2.3.1. *Global and Local Extreme Points of a Function. Extreme Value of a Function*

As we shall see below, the definition of extreme points (maxima and minima)[53] of a function of several variables is similar to that corresponding to a single variable. In other words:

Definition: The function $z = f(\vec{x}) = f(x_1, \dots, x_n)$ at a point $\vec{a} = (a_1, \dots, a_n) \in Domf$ has:

1. A **global maximum** if $f(\vec{x}) \leq f(\vec{a})$ for any point $\vec{x} = (x_1, \dots, x_n) \in Domf$.
2. A **global minimum** if $f(\vec{a}) \leq f(\vec{x})$ for any point $\vec{x} = (x_1, \dots, x_n) \in Domf$.
3. A **local maximum** if $f(\vec{x}) \leq f(\vec{a})$ takes place in an open ball centered at $\vec{a} \in Domf$ and included in the domain $Domf$.[54]
4. A **local minimum** if $f(\vec{a}) \leq f(\vec{x})$ takes place in an open ball centered at $\vec{a} \in Domf$ and included in the domain $Domf$.

The real value that an extreme point (either global or local) reaches by the function:

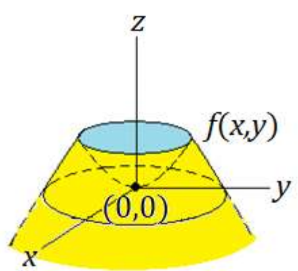
$$z_0 = f(\vec{a}) = f(a_1, \dots, a_n) \in \mathbb{R}$$

is, by definition, an **extreme value** of the function.[55]

In fact, not every global optimum is local. For example, the picewise function defined by:

$$f(x,y) = \begin{cases} x^2 + y^2, & 0 \leq x^2 + y^2 \leq 1 \\ 2 - x^2 - y^2, & \text{otherwise} \end{cases}$$

has a local minimum at point (0,0) that is not a global minimum. Graphically:



[53] Also called **optima.**

[54] An "open ball" in $\mathbb{R}^n$ centered at a point is the set of alls points placed at a distance of it smaller than a radius done. For instance, open balls in the plane are disks without the corresponding circumference.

[55] Also called **optimal value.**

2.3.2. *Necessary First-Order Condition of Existence of Extreme Points* [56]

In what follows we assume that the functions have 1st and 2nd order continuous partial derivatives.[57] As we have seen few pages back (paragraph 2.2.3.1) provided that the gradient of a function is different from the zero-vector the function can increase or decrease. So, the necessary condition for a function to have a extreme point is that the gradient at this point, in the case that it does exist, is precisely the zero-vector. So that:

<u>Theorem</u>: A necessary condition for a point $\vec{a} = (a_1, \dots, a_n)$ to be a local extreme point of a function $z = f(x_1, \dots, x_n)$ is that:

$$\nabla f(\vec{a}) = \left(\frac{\partial f(\vec{a})}{\partial x_1}, \dots, \frac{\partial f(\vec{a})}{\partial x_n}\right) = \left(\overbrace{0, \dots, 0}^{n}\right).$$

Consider this example:

<u>Example</u>: Check that the function:

$$z = f(x, y) = 1 - 3x^2 - y^2 + 2y$$

satisfies the necessary first-order condition at point (0,1).

SOLUTION: Since the gradient of this function is:

$$\nabla f(x, y) = \left(\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y}\right) = (-6x, -2y + 2)$$

it is straightforward to see that it satisfies the necessary first-order condidition. Indeed:

$$\nabla f(0,1) = \left(\frac{\partial f(0,1)}{\partial x}, \frac{\partial f(0,1)}{\partial y}\right) = (-6 \cdot 0, -2 \cdot 1 + 2) = (0,0) \blacksquare$$ [58]

It should be emphasized that this first-order condition is satisfied only when the extreme points have associated gradient.

---

[56] And we say "first-order" condition because the gradient is involved on it.

[57] This condition is essential and all the functions studied here match it.

[58] Consequently, if this function has an extreme point, this must be (0,1). In fact, it can be proved that this function has a global maximum at this point.

2.3.2.1. Stationary and Saddle Points of a Function

Because of the previous theorem, all those points with associated zero-gradient play a key role in the process of searching extreme points. So, consider the following definition:

Definition: A point $\vec{a} = (a_1, \dots, a_n)$ is a **stationary point** of $z = f(x_1, \dots, x_n)$ provided that it has an associated zero-gradient vector.

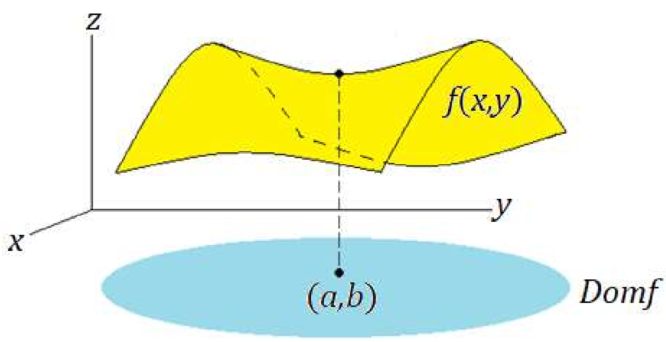
Example: Calculate the stationary points of $f(x, y, z) = 2x^2 + y^2 - z^2 + x + 5$.
SOLUTION:
The sole stationary point of this function is $\left(-\frac{1}{4}, 0,0\right)$ since:

$$(0,0,0) = \nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (4x + 1, 2y, -2z) \text{ implies: } (x, y, z) = \left(-\frac{1}{4}, 0,0\right) \blacksquare$$

Yet a stationary point does not always have to be an extreme point. A stationary point that is not an extreme point is known as a saddle point:

Definition: A point $\vec{a} = (a_1, \dots, a_n)$ is a **saddle point** of $z = f(x_1, \dots, x_n)$ provided that it is a stationary point but not an extreme point of the function. [59]

Graphically, the point (a, b) is a saddle point of $z = f(x, y)$:



[59] The stationary point of the above example, as we shall see below, is a saddle point.

2.3.3. *Necessary Second-Order Condition of Existence of Extreme Points*

This condition involves the Hessian matrices of the function under study. We must stress here that every Hessian matrix can be viewed as the symmetric matrix associated to a certain quadratic form.

<u>Theorem</u>: [60] Let $\vec{a} = (a_1, \dots, a_n)$ be a stationary point of the function $z = f(x_1, \dots, x_n)$. If this function has:

1. A local minimum at $\vec{a} = (a_1, \dots, a_n)$, then the Hessian matrix $Hf(\vec{a})$ is associated to an either positive definite or semidefinite quadratic form.
2. A local maximum at $\vec{a} = (a_1, \dots, a_n)$, then the Hessian matrix $Hf(\vec{a})$ is associated to an either either negative definite or semidefinite quadratic form.

<u>Example</u>: Assuming that the function:

$$z = f(x, y) = x^2 + y^2$$

has a global minimum at point (0,0), check the necessary second-order condition.

SOLUTION:

It is straightforward since the Hessian matrix of this function:

$$Hf(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

is positive definite■

This theorem has an interesting conclusion: every stationary point with an associated Hessian matrix neither definite nor semidefinite cannot be an extreme point in any way.[61] This issue is developed in the next paragraph.

---

[60] This theorem and the sequel justify the study of the sign of quadratic forms.

[61] From now onwards, we will "assign" to any Hessian matrix the sign of the associated quadratic form.

2.3.3.1. The Necessary Second-Order Condition and Saddle Points

Remember that the necessary second-order condition introduced below affirms that the Hessian matrix associated with any extreme point cannot be indefinite. Hence, we have:

Theorem: Let $\vec{a} = (a_1, \dots, a_n)$ be a stationary point of the function $z = f(x_1, \dots, x_n)$. If the Hessian matrix $Hf(\vec{a})$ is indefinite, then this point is a saddle point.

Let us look at an example:

Example: Prove that the function:

$$f(x, y, z) = 2x^2 + y^2 - z^2 + x + 5$$

has a saddle point at $\left(-\frac{1}{4}, 0,0\right)$.

SOLUTION:

We already know that this point is a stationary point of the function.[62] Therefore since the Hessian matrix is constant and equal to:

$$\nabla f(x, y) = (4x + 1, 2y, -2z) \text{ implies: } Hf(x, y, z) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

we deduce that the Hessian matrix at this point:

$$Hf\left(-\frac{1}{4}, 0,0\right) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

is indefinite and, thanks to the previous theorem, the point $\left(-\frac{1}{4}, 0,0\right)$ is a saddle point of this function ■

[62] See page 40.

2.3.4. *Sufficient Second-Order Condition of Existence of Extreme Points*

This condition is also related to the Hessian matrices associated with the function under study[63] and shows us when stationary points are local extreme points.[64] The next theorem is very important:

Theorem: Let $\vec{a} = (a_1, \dots, a_n)$ be a stationary point of the function $z = f(x_1, \dots, x_n)$. Then:

1. If the Hessian matrix $Hf(\vec{a})$ is positive definite, $\vec{a} = (a_1, \dots, a_n)$ is a local minimum.
2. If the Hessian matrix $Hf(\vec{a})$ is negative definite, $\vec{a} = (a_1, \dots, a_n)$ is a local maximum.[65]

Let us take a look at this following example:

Example: Prove that the function:

$$z = f(x, y) = 1 - 3x^2 - y^2 + 2y$$

has a local maximum at point (0,1).

SOLUTION:

We already know that (0,1) is a stationary point of the function.[66] Since the associated Hessian matrix:

$$Hf(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -6 & 0 \\ 0 & -2 \end{pmatrix} \text{ implies: } Hf(0,1) = \begin{pmatrix} -6 & 0 \\ 0 & -2 \end{pmatrix}$$

is negative definite, we can conclude that the stationary point (0,1) is a local maximum of the function ■[67]

Remember that we cannot apply the above analytical procedure in order to obtain extreme points in those situations in wich there is neither gradient nor Hessian matrix.

[63] Hereby we refer to it as a “second-order” condition.
[64] Local not global. The precedent results can only enable us to get local extreme points.
[65] In the case the Hessian matrix is semidefinite we cannot say anything in principle.
[66] See page 39.
[67] Moreover this local maximum is a global maximum as we have already mentioned.

2.3.4.1. Example of Application

Example: Find the extreme points and the associated extreme values of the function:

$$z = f(x,y) = 8xy + \frac{1}{x} + \frac{1}{y}.$$

SOLUTION:

Note that the domain of this function is:

$$Domf = \{(x,y) \in \mathbb{R}^2 : x \neq 0, y \neq 0\}.$$

Finding their extreme points means, first, to study the necessary first-order condition, i.e., to solve the following system of non-linear equations:

$$\left.\begin{matrix} 0 = \dfrac{\partial f}{\partial x} = 8y - \dfrac{1}{x^2} \\ 0 = \dfrac{\partial f}{\partial y} = 8x - \dfrac{1}{y^2} \end{matrix}\right\} \text{ equivalent to: } \begin{cases} 8x^2y = 1 \\ 8xy^2 = 1 \end{cases}$$

Since $x \neq 0$ and $y \neq 0$ we deduce that:

$$0 = 1 - 1 = 8x^2y - 8xy^2 = 8xy(x - y) \text{ implies: } x - y = 0 \text{ implies: } x = y.$$

Hence, taking the first equation for example, we have:

$$x = y \text{ implies: } 1 = 8x^2y = 8x^3 \text{ implies: } x = y = \tfrac{1}{2}.$$

Thus, this function has a stationary point at $\left(\frac{1}{2}, \frac{1}{2}\right)$. Now since the associated Hessian matrix at this point:

$$Hf(x,y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2/_{x^3} & 8 \\ 8 & 2/_{y^3} \end{pmatrix} \text{ implies: } Hf\left(\frac{1}{2}, \frac{1}{2}\right) = \begin{pmatrix} 16 & 8 \\ 8 & 16 \end{pmatrix}$$

is positive definite we conclude that:

$$\left(\frac{1}{2}, \frac{1}{2}\right) \text{ is a local minimum.}$$

Finally the extreme (minimum) value of the function will be:

$$z_0 = f\left(\frac{1}{2}, \frac{1}{2}\right) = 8 \cdot \left(\frac{1}{2}\right) \cdot \left(\frac{1}{2}\right) + \frac{1}{\left(\frac{1}{2}\right)} + \frac{1}{\left(\frac{1}{2}\right)} = 6 \blacksquare$$

2.3.4.2. Application of Extreme Points in Economics

Example: Let $C(q_1, q_2) = 3q_1 \cdot q_2 + 1{,}5q_1^2 + 2q_2^2$ be the function of costs in € of a firm that produces two goods A and B, where $q_1 \geq 0$ and $q_2 \geq 0$ are, respectively, the quantities produced. The firm wants to know how many quantities of A and B maximize benefits in the case that the unit sale prices of A and B are:

1. $p_1 = €42$ and $p_2 = €51$.
2. $p_1 = 66 - 3q_1$ and $p_2 = 72 - q_2$.

SOLUTION: (1) In this market the benefits function is:

$$B(q_1, q_2) = (p_1 \cdot q_1 + p_2 \cdot q_2) - C(q_1, q_2) = -1{,}5q_1^2 - 2q_2^2 - 3q_1q_2 + 42q_1 + 51q_2.$$

Since:

$$\left.\begin{array}{l} 0 = \dfrac{\partial B}{\partial q_1} = -3q_1 - 3q_2 + 42 \\ 0 = \dfrac{\partial B}{\partial q_2} = -4q_2 - 3q_1 + 51 \end{array}\right\} \text{implies: } q_1 = 5 \text{ and } q_2 = 9$$

and:

$$Hf(5{,}9) = \begin{pmatrix} -3 & -3 \\ -3 & -4 \end{pmatrix} \text{ is negative definite}$$

we see that the firm must sell $q_1 = 5$ units of A and $q_2 = 9$ units of B in order to maximize benefits. In this case, the maximum of benefits will be of $B(5{,}9) = €333{,}5$ ■

(2) In this other market the benefit function is:

$$\begin{aligned} B(q_1, q_2) &= (p_1 \cdot q_1 + p_2 \cdot q_2) - C(q_1, q_2) = \\ &= \big((66 - 3q_1) \cdot q_1 + (72 - q_2) \cdot q_2\big) - (3q_1q_2 + 1{,}5q_1^2 + 2q_2^2) = \\ &= -4{,}5q_1^2 - 3q_2^2 - 3q_1q_2 + 66q_1 + 72q_2. \end{aligned}$$

Since:

$$\left.\begin{array}{l} 0 = \dfrac{\partial B}{\partial q_1} = -9q_1 - 3q_2 + 66 \\ 0 = \dfrac{\partial B}{\partial q_2} = -6q_2 - 3q_1 + 72 \end{array}\right\} \text{implies: } q_1 = 4 \text{ and } q_2 = 10$$

and:

$$Hf(4{,}10) = \begin{pmatrix} -9 & -3 \\ -3 & -6 \end{pmatrix} \text{ is negative definite}$$

the firm now must sell $q_1 = 4$ units of A and $q_2 = 10$ units of B to maximize benefits. The value of these benefits will be of $B(4{,}10) = €492$, being the sales prices of A and B:

$$p_1 = 66 - 3 \cdot 4 = €54 \text{ and } p_2 = 72 - 10 = €62 \blacksquare$$

2.3.4.3. Application of Extreme Points in Econometrics

Consider now this example of application of the "linear regression" method commonly used in Econometrics:

Example: It is known that a certain phenomenon under study follows a linear law of type:

$$Y = f(X) = a \cdot X + b$$

where X and Y represent accurate economic indicators being $a, b > 0$ parameters. On the other hand, suppose we have just experimentaly obtained four values of Y corresponding to four values of X according to the table:

X_t	1	2	4	6
Y_t	3.1	3.4	4.2	4.5

Under these conditions find the values of parameters $a, b > 0$ such that the above linear law describes as best as possible the phenomenon.

SOLUTION:

The linear regression method supposes to minimize the "loss" function defined by:

$$L(a, b) = \sum_{i=1}^{4} \left(Y_t - f(X_t)\right)^2 =$$

$$= \left(3.1 - f(1)\right)^2 + \left(3.4 - f(2)\right)^2 + \left(4.2 - f(4)\right)^2 + \left(4.5 - f(6)\right)^2 =$$

$$= \left(3.1 - (a \cdot 1 + b)\right)^2 + \left(3.4 - (a \cdot 2 + b)\right)^2 + \left(4.2 - (a \cdot 4 + b)\right)^2 + \left(4.5 - (a \cdot 6 + b)\right)^2 =$$

$$= 57a^2 + 4b^2 + 26ab - 107.4a - 30.4b + 59.06.$$

Since:

$$\left.\begin{array}{l} 0 = \dfrac{\partial L}{\partial a} = 114a + 26b - 107.4 \\ 0 = \dfrac{\partial L}{\partial b} = 8b + 26a - 30.4 \end{array}\right\} \text{implies: } a = \frac{86}{295} \cong 0.291 \text{ and: } b = \frac{1683}{590} \cong 2.852$$

and the general Hessian matrix:

$$HL\left(\frac{86}{295}, \frac{1683}{590}\right) = \begin{pmatrix} 114 & 26 \\ 26 & 8 \end{pmatrix} \text{ is positive definite}$$

we conclude that the function:

$$Y = f(X) = 0.291 \cdot X + 2.852$$

is the best choice in order to linearly describe the phenomenon■

2.3.5. *Constrained Optimization: Direct-Case Method*

This method of optimization may be applyied when a system of linear equations with the independent variables as unknowns acts as a "constrain" on the function. Basically, what is required here is to introduce the solution of this system into the function and then optimizing the resulting "auxiliary" function. Let us see an example:

Example: Find the extreme points of $u = f(x, y, z) = 2xy - 3xz + yz$ constrained to the equation $3x - 2y + z = 1$.[68]
SOLUTION:
Since the "solution" in terms of the variable z of this linear equation is:

$$z = 1 - 3x + 2y$$

the direct-case method basically means to introduce this dependence into the initial function and optimizing the "auxiliary" function of two variables thus obtained:

$$u = u(x, y) = f(x, y, 1 - 3x + 2y) = 9x^2 + 2y^2 - 7xy - 3x + y.$$

Since:

$$\left.\begin{array}{l} 0 = \dfrac{\partial u}{\partial x} = 18x - 7y - 3 \\ 0 = \dfrac{\partial u}{\partial y} = 4y - 7x + 1 \end{array}\right\} \text{implies: } x = \frac{5}{23} \text{ and: } y = \frac{12}{92}$$

and the general Hessian matrix:

$$Hu\left(\frac{5}{23}, \frac{12}{92}\right) = \begin{pmatrix} 18 & -7 \\ -7 & 4 \end{pmatrix} \text{ is positive definite}$$

the point $\left(\frac{5}{23}, \frac{12}{92}\right)$ is a minimum of $u = u(x, y)$. Due to the fact that:

$$z = 1 - 3x + 2y \text{ and: } (x, y) = \left(\tfrac{5}{23}, \tfrac{12}{92}\right) \text{ implies: } z = \tfrac{14}{23}$$

the direct-case method enables us to affirm that the point:

$$(x, y, z) = \left(\frac{5}{23}, \frac{12}{92}, \frac{14}{23}\right)$$

is a local minimum of $u = f(x, y, z)$ constrained to the equation $3x - 2y + z = 1$■

[68] It is straighforward to see that $u = f(x, y, z)$ has no optima but a saddle point at (0,0,0).

2.3.5.1. Application in Economics: Marshall's Model of Consumption

Let us see now an interesting application of constrained optimization in Economics:

Example: Let:

$$U(x,y) = 0.05 \cdot x \cdot y$$

be the consumption utility associated with the purchase of $x > 0$ units of a commodity A and $y > 0$ units of a commodity B. If €5 and €8 are their corresponding unit sale prices determine the amount of both A and B that maximize the utility of spending €400 on them.

SOLUTION:

Since the "purchase constraint" is the linear equation:[69]

$$€5 \cdot x + €8 \cdot y = €400 \text{ implies: } 5x + 8y = 400$$

we must solve the problem:

$$\left.\begin{array}{ll} \text{maximize} & U(x,y) = 0.05xy \\ \text{subject to:} & 5x + 8y = 400 \end{array}\right\}^{70}$$

applying the direct-case method. Since:

$$5x + 8y = 400 \text{ implies: } y = \frac{400-}{8} = 50 - 0.625x$$

we have first to maximize the auxiliary function of one-variable:

$$y = 50 - 0.625x \text{ implies: } U = U(x, 50 - 0.625x) = 2.5x - 0.03125x^2.$$

Since:

$$0 = \frac{dU}{dx} = 2.5 - 0.0625x \text{ implies: } x = 40$$

and:

$$\frac{d^2U}{dx^2} = -0.0625 < 0$$

we deduce that $x = 40$ maximizes the auxiliary function. Because:

$$y = 50 - 0.625x \text{ and: } x = 40 \text{ implies: } y = 25$$

we can affirm that 40 units of A and 25 units of B maximize the utility of the consumption of €400 being the optimal utility equal to:

$$U(40,25) = 0.05 \cdot 40 \cdot 25 = €50 \blacksquare$$

---

[69] Note that this linear constrain is a hyperplane, the so-called hyperplane of "monetary constraint".

[70] This type of problems are called *programs* in Mathematical Economics.

2.4. Exercises

1. Determine both the domain and the level curves of the function $f(x,y) = +\sqrt{\frac{x^2+y}{x^2-y}}$.
2. Prove that the function $z = \frac{1}{2} \cdot \ln(x^2 + xy - y^2)$ satisfies the equality $x \cdot \frac{\partial z}{\partial x} + y \cdot \frac{\partial z}{\partial y} = 1$.
3. Calculate the Hessian matrix of $f(x,y,z) = x^2y - yz^2$ at point (1,1,1).
4. Study the existence of extreme points of:
   a. $z = f(x,y) = x^2 + y^2 - 6xy - 39x + 18y + 20$.
   b. $z = f(x,y) = 2x + 3y - x^2 - 2y^2 + xy$.
   c. $z = f(x,y) = x^2 + y^2 - 2 \cdot \ln x - 18 \cdot \ln y$.
   d. $z = f(x,y) = xy - x^3 - y^3$.
   e. $u = f(x,y,z) = x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 12yz$.
5. An employer who manufactures two products A and B knows that, at a price of €25, he can sell 175 units of A and that for every euro he lowers this price, the sales increase by 5 units. If the price of B is always €30 and the function:

   $$C(x,y) = \frac{x^2}{5} + \frac{y^2}{3} + 1{,}925$$

   is the cost function, where $x > 0$ and $y > 0$ represent the output of A and B, determine how many units of these products he should manufacture in order to maximize benefits.
6. A company manufactures fabrics of 103 meters long and 1 meter wide that are dyed green, yellow and red. If the cost of production is:

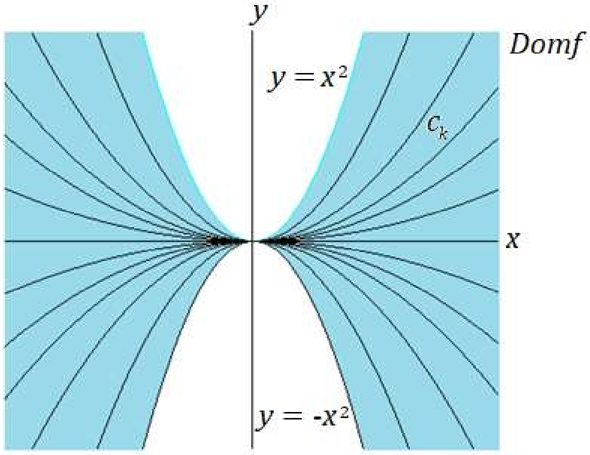
   $$C(x,y,z) = 0.03x^2 + 0.05y^2 + 0.11z^2 + 40.05$$

   where $x > 0$, $y > 0$ and $z > 0$ are the respective number of square meters of each color depicted, determine the area of green, yellow and red of each fabric if this company wants to minimize the cost of production.[71]

[71] This exercise needs to be solved using the direct-case method.

SOLUTIONS:

1.



3.

$$Hf(1,1,1) = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 0 & -2 \\ 0 & -2 & -2 \end{pmatrix}.$$

4.

a. $\left(\frac{15}{16}, -\frac{99}{16}\right)$ is a saddle point.

b. $\left(\frac{11}{7}, \frac{8}{7}\right)$ is a local maximum.

c. $(1,3)$ is a local minimum.

d. $(0,0)$ is a saddle point and $\left(\frac{1}{3}, \frac{1}{3}\right)$ is a local maximum.

e. $(0,0,0)$ is a saddle point.

5. 75 units of A and 45 units of B. The maximum value of benefits is €1,000.

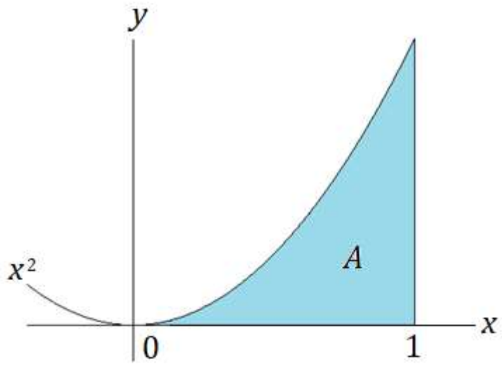
6. 55 m^2 green, 33 m^2 yellow and 42 m^2 red. The minimum cost is €210.

3. DYNAMICAL ANALYSIS

This section is basically devoted to integrals. As an illustrative case of what is coming suppose that we want to evaluate the area A given by the parabola:

$$y = x^2$$

as illustrated bellow:



Note that no classical formula is suitable for this problem to be solved since we do not know up to now how to figure out areas of non-rectilinean planar shapes, as it is the case at hand. Luckily integrals come to give us the correct solution in the following way. Since the parabola viewed as a function of one-real variable has associated a "primitive" function (antiderivative) of the form:

$$F(x) = \frac{x^3}{3}$$

mathematical results[72] enables us to affirm that the area A is equal to the "definite integral":

$$A = \int_0^1 x^2 dx = \{\text{Barrow's rule}\} = F(1) - F(0) = \left(\frac{1^3}{3}\right) - \left(\frac{0^3}{3}\right) = \frac{1}{3}.$$

This procedure can be spread over the mostly part of functions of one-real variable. Thus, this section is devoted to study how to calculate definite integrals, looking at some of their applications, and using them finally to solve some of the most basic differential equations.[73]

[72] Fundamental Theorem of Integral Calculus. See pages 61 and 62.

[73] Differential equations allow us to study a lot of natural phenomena where mathematics has something important to say. It must be emphasize that this type of equations is perhaps one of the most powerful tools in science.

3.1. Indefinite Integrals

3.1.1. *Primitive of a Function of One Real Variable*

From an "informal" point of view we can say that integrating a function of a real variable is the opposite of calculating its derivative.[74] Hence by definition:

Definition: A **primitive** (also called **antiderivative**) of a function of one variable $y = f(x)$ is another function $F(x)$ such that its derivative is precisely $f(x)$. Formally:

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

Example: Prove that:

$$F(x) = \frac{x^3}{3} + 27$$

is a primitive of $f(x) = x^2$.
SOLUTION:
Evidently since:

$$F'(x) = \frac{dF(x)}{dx} = \frac{3x^2}{3} + 0 = x^2 \blacksquare$$

We should emphasize that a primitive of a function is not unique. Indeed:

Theorem: If $F_1(x)$ and $F_2(x)$ are two primitives of a same function, a constant $C \in \mathbb{R}$ always exists such that:

$$F_2(x) = F_1(x) + C.$$

We may note that in the above case all primitives of $f(x) = x^2$ would be the functions like:

$$F(x) = \frac{x^3}{3} + C, \text{ with } C \in \mathbb{R}.$$

[74] That is why the result of the process of calculating the integral of a function is sometimes called "antiderivative".

3.1.2. *Indefinite Integral of a Function and Properties*

The last property allows us to define the concept of the indefinite integral. Formally:

Definition: The **indefinite integral** (**integral** for short) of $y = f(x)$, in the case it exists, is equal to the expression: [75]

$$\int f(x)\, dx = F(x) + C$$

in which $F(x)$ is a primitive of the function $f(x)$ and $C \in \mathbb{R}$ is a constant.

Example: Given the function:

$$y = f(x) = \frac{1}{x}$$

find:

1. The indefinite integral.
2. The primitive passing through the point $(e, 2)$.

SOLUTION:

(1) Since the derivative of $F(x) = \ln x$ is $f(x) = \frac{1}{x}$ we can affirm that the required indefinite integral is:

$$\int \left(\frac{1}{x}\right) dx = \ln x + C \blacksquare$$

(2) Now we have to find a function of the form:

$$F(x) = \ln x + C$$

passing through the point $(e, 2)$. Since:

$$2 = F(e) = \ln e + C = 1 + C \text{ implies: } C = 1$$

then the primitive of $f(x) = \frac{1}{x}$ we have to find is the one-real variable function:

$$F(x) = \ln x + 1 \blacksquare$$

Therefore and from a geometrical point of view, any indefinite integral represents a family of curves in the plane. Each of these curves is precisely a primitive.

[75] This symbolism is related to the concept of the definite integral, which we will study later.

3.1.2.1. Immediate Integrals

In order to calculate indefinite integrals, it is essential to take into account the following list of **immediate integrals**, i.e., integrals that can be directly resolved from the definition. Indeed:

1. $\int a \cdot dx = a \cdot x + C$, for any constant $a \in \mathbb{R}$.
2. $\int x^a dx = \frac{x^{a+1}}{a+1}$, where $a \neq -1$.
3. $\int x^{-1} dx = \int \frac{dx}{x} = \ln x + C$.
4. $\int a^x dx = \frac{a^x}{\ln a}$, where $a > 0$.
5. $\int \sin x \, dx = -\cos x + C$.
6. $\int \cos x \, dx = \sin x + C$.
7. $\int \left(\frac{1}{\cos^2 x}\right) dx = \int (1 + \tan^2 x) dx = \tan x + C$.
8. $\int \left(\frac{1}{1+x^2}\right) dx = \tan^{-1} x + C$.

Unfortunately, the process of integration does not support anything similar to the chain rule of the corresponding process of differentiation. That means that not all functions have an indefinite integral in the above sense. For instance, the following two integrals:

$$\int \left(\frac{\sin x}{x}\right) dx \text{ and } \int e^{-x^2} dx$$

cannot be properly expressed in terms of suitable functions. Consequently, we have to integrate each different type of integral using an appropriate method; in fact, the main purpose that we are going to deal in this paragraph is to study the easiest cases.[76]

[76] We have choosed two specific cases: logarithmic integration and integration by substitution.

3.1.2.2. Propeties of Indefinite Integrals

We have first to consider the following properties along with the immediate integrals noticed in the previous page:

Properties:

1. $\int f'(x)\,dx = f(x) + C$.
2. $\int \big(f(x) + g(x)\big)\,dx = \int f(x)\,dx + \int g(x)\,dx$..
3. $\int \lambda \cdot f(x)\,dx = \lambda \cdot \int f(x)\,dx$, for any constant $\lambda \in \mathbb{R}$.
4. "Logarithm integration": [77]

$$\int \frac{f'(x)}{f(x)}\,dx = \ln f(x) + C.$$

Example: Calculate: (1) $\int (x^4 - 7x^3 + 5)\,dx$. (2) $\int \left(\frac{\sqrt{x}-2}{x^2}\right)dx$. (3) $\int \frac{xdx}{1+x^2}$. (4) $\int \tan x\,dx$.

SOLUTION:

(1) Applying the above properties, we have:

$$\int (x^4 - 7x^3 + 5)\,dx = \int x^4\,dx - 7\int x^3\,dx + 5\int dx = \left(\frac{x^5}{5}\right) - 7\left(\frac{x^4}{4}\right) + 5x + C \blacksquare$$

(2) Now dividing by x^2 and applying the list of immediate integrals:

$$\int \left(\frac{\sqrt{x}-2}{x^2}\right)dx = \int \left(x^{-\frac{3}{2}} - 2x^{-2}\right)dx = \left(\frac{x^{-\frac{3}{2}+1}}{-\frac{3}{2}+1}\right) - 2\left(\frac{x^{-2+1}}{-2+1}\right) + C = -\left(\frac{2}{\sqrt{x}}\right) + \frac{2}{x} + C \blacksquare$$

(3) Thanks to logarithm integration and multiplying and dividing by 2 we deduce that:

$$\int \frac{xdx}{1+x^2} = \frac{1}{2}\cdot\int \left(\frac{2x}{1+x^2}\right)dx = \frac{1}{2}\cdot\ln(1+x^2) + C \blacksquare$$

(4) In the same way as before, and thanks to the definition of the tangent function, we have now:

$$\int \tan x\,dx = \int \left(\frac{\sin x}{\cos x}\right)dx = -\int \left(\frac{-\sin x}{\cos x}\right)dx = -\ln\cos x + C \blacksquare$$

[77] Roughly speaking this property can be seen as a specific integration method.

3.1.2.3. Integration by Substitution

Remember that the derivative of a function $y = f(x)$ is nothing but the expression:

$$\frac{df(x)}{dx} = f'(x) \text{ equivalent to: } df(x) = f'(x) \cdot dx.$$

Bearing in mind this, the method of integration by substitution is based on the so-called "change of variable". In simple words, to introduce a change of variable into an integral means substituting the variable x with a suitable function $\varphi(t)$ depending on a new variable t. Schematically: [78]

$$x = \varphi(t) \text{ implies: } dx = \varphi'(t) \cdot dt.$$ [79]

The by-product of this process is a new integral depending on the new variable t. If this last integral can be solved then, reverting the mentioned process, we can easily get the former integral depending on the primitive variable x. The following example shows us the basic guidelines of how this method works:

Example: Calculate:

$$\int \left(x \cdot \sqrt{1-x}\right) dx$$

using the change of variable $1 - x = t$.

SOLUTION: In this case the function $\varphi(t)$ would be:

$$1 - x = t \text{ implies: } x = \varphi(t) = 1 - t.$$

Consequently:

$$dx = \varphi'(t) \cdot dt = (-1) \cdot dt = -dt.$$

Thus:

$$\int \left(x \cdot \sqrt{1-x}\right) dx = \begin{Bmatrix} x = 1-t \\ dx = -dt \end{Bmatrix} = \int \left((1-t) \cdot \sqrt{t}\right)(-dt) = \int \left((t-1) \cdot t^{\frac{1}{2}}\right) dt =$$

$$= \int \left(t^{\frac{3}{2}} - t^{\frac{1}{2}}\right) dt = \left(\frac{t^{\frac{3}{2}+1}}{3/2+1}\right) - \left(\frac{t^{\frac{1}{2}+1}}{1/2+1}\right) + C = \left(\frac{2t^{\frac{5}{2}}}{5}\right) - \left(\frac{2t^{\frac{3}{2}}}{3}\right) + C =$$

$$= \{t = 1-x\} = \left(\frac{2(1-x)^{\frac{5}{2}}}{5}\right) - \left(\frac{2(1-x)^{\frac{3}{2}}}{3}\right) + C \blacksquare$$

[78] Formally speaking this new function must be bijective and differentiable.
[79] The derivative $\varphi'(t)$ is evaluated respect to the new variable t.

3.1.2.4. Example of Integration by Substitution

Example: Calculate the following integrals:

$$(1) \int (1-2x)^7\, dx.\ (2) \int \frac{dx}{(x-3)^2}.\ (3) \int e^{x^3} x^2\, dx.\ (4) \int \left(\sqrt[3]{1-3x}\right) dx.\ (5) \int \frac{0{,}5dx}{\sqrt{x+5}}.$$

SOLUTION:

(1) In this case we can introduce the change $t = 1 - 2x$. Hence:

$$\int (1-2x)^7\, dx = \left\{ \begin{array}{l} x = \dfrac{1-t}{2} \\ dx = -\dfrac{dt}{2} \end{array} \right\} = \int t^7 \left(-\frac{dt}{2}\right) = -\frac{1}{2}\int t^7 dt = -\frac{1}{2}\cdot\left(\frac{t^8}{8}\right) + C = -\frac{(1-2x)^8}{16} + C \blacksquare$$

(2) Now introducing the change $t = x - 3$ we have:

$$\int \frac{dx}{(x-3)^2} = \left\{ \begin{array}{c} x = t + 3 \\ dx = dt \end{array} \right\} = \int \frac{dx}{t^2} = \int t^{-2}\, dt = \left(\frac{t^{-1}}{-1}\right) + C = -\left(\frac{1}{t}\right) + C = -\left(\frac{1}{x-3}\right) + C \blacksquare$$

(3) Applying the change $x^3 = t$ we have:

$$x^3 = t \text{ implies: } (3x^2)dx = 1 \cdot dt \text{ implies: } dx = \frac{dt}{3x^2}.$$

Therefore:

$$\int e^{x^3} x^2\, dx = \left\{ \begin{array}{c} x^3 = t \\ dx = \dfrac{dt}{3x^2} \end{array} \right\} = \int e^t \cdot x^2 \cdot \left(\frac{dt}{3x^2}\right) = \frac{1}{3}\cdot \int e^t\, dt = \frac{1}{3}\cdot e^t + C = \frac{e^{x^3}}{3} + C \blacksquare$$

(4) With the change $1 - 3x = t$ we have:

$$\int \left(\sqrt[3]{1-3x}\right) dx = \left\{ \begin{array}{l} x = \dfrac{1-t}{3} \\ dx = -\dfrac{dt}{3} \end{array} \right\} = \int \sqrt[3]{t} \cdot \left(-\frac{dt}{3}\right) = -\frac{1}{3}\cdot \int t^{\frac{1}{3}}\, dt = -\frac{1}{3}\cdot \left(\frac{t^{\frac{4}{3}}}{4/3}\right) + C =$$

$$= -\left(\frac{t^{\frac{4}{3}}}{4}\right) + C = -\left(\frac{(1-3x)^{\frac{4}{3}}}{4}\right) + C \blacksquare$$

(5) As in the former case, and considering the change $x + 5 = t$ we deduce that:

$$\int \frac{0{,}5dx}{\sqrt{x+5}} = \left\{ \begin{array}{c} x = t - 5 \\ dx = dt \end{array} \right\} = 0.5 \cdot \int \frac{dt}{\sqrt{t}} = 0.5 \cdot \int t^{-0.5}\, dt = 0.5 \cdot \left(\frac{t^{0.5}}{0.5}\right) + C = t^{0.5} + C =$$

$$= \sqrt{x+5} + C \blacksquare$$

3.2. Definite Integrals

3.2.1. *Definite Integral of a Function of One Real Variable and Properties*

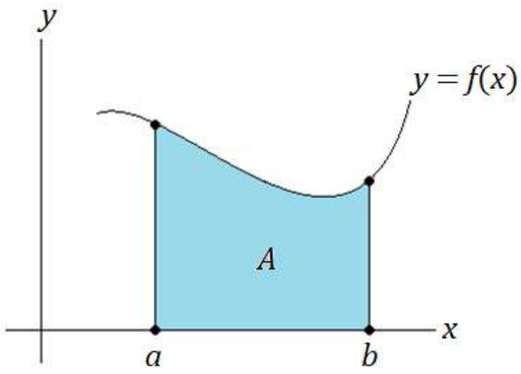
Let $y = f(x)$ be a continuous function with the positive restriction:

$$f(x) \geq 0, \text{ for any } a \leq x \leq b.$$

In this case, the **definite integral** of $y = f(x)$ between these two points a and b, symbolized by the expression:

$$\int_a^b f(x)\, dx$$

measures the area A delimited by the x-axis, the function $y = f(x)$ and the two vertical lines $x = a$ and $x = b$. Graphically:



As we have mentioned we shall put:

$$A = \int_a^b f(x)\, dx.$$ [80]

It is worth noting that the positive restriction is not an obstacle when we are calculating areas like this, as we will have the chance to see it right now.

---

[80] We can find the formal definition of the definite integral in most manuals devoted to this issue. See Sydsaeter. K & Hammond, P. J. (1995), pages 320-323 to become aware about the main idea that underlies this concept. Note that the definite integrals is always a number.

3.2.1.1. Properties of Definite Integrals

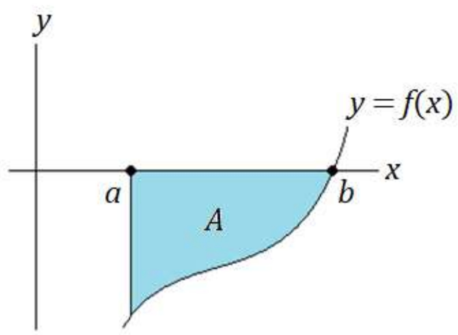
Properties:

1. $\int_a^b \big(f(x)+g(x)\big)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$.
2. $\int_a^b \lambda \cdot f(x)\,dx = \lambda \cdot \int_a^b f(x)\,dx$, for any constant $\lambda \in \mathbb{R}$.
3. $\int_a^a f(x)\,dx = 0$.
4. $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$.
5. $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$, for any $a \le c \le b$.
6. If $f(x) \ge 0$, for all $a \le x \le b$, then $\int_a^b f(x)\,dx \ge 0$.
7. If $f(x) \le 0$, for all $a \le x \le b$, then $\int_a^b f(x)\,dx \le 0$.

From these properties, we can deduce that if a continuous function $y = f(x)$ is negative between points a and b, i.e.:

$$f(x) \le 0, \text{ for any } a \le x \le b$$

as in this case:



the value of the area A is equal to the opposite value of the definite integral. In other words:

$$A = \left|\int_a^b f(x)\,dx\right|.$$ [81]

[81] If we wish to use definite integrals to figure out areas without drawing them, we ought consider their absolute values as we are going to see.

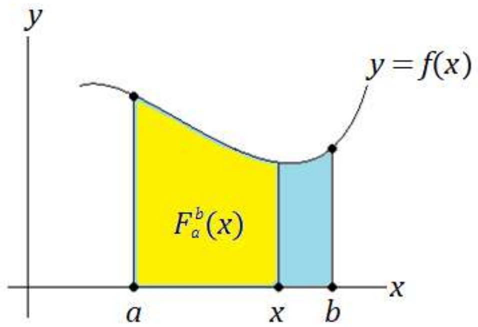
3.2.2. *Fundamental Theorem of Integral Calculus. Integral Function*

The geometrical process followed to define and figure out non rectilinear areas (see Sydsaeter. K & Hammond, P. J. (1995), pages 320-323 as mentioned) is not the main procedure to pursue in most of the cases. Instead, we usually take advantage of a fundamental property that put in relation indefinite and definite integrals. Let us take a look at what the Fundamental Theorem of the Integral Calculus states.

Definition: The **integral function** of a continuous function $y = f(x)$ between points $a \leq b$ is the one-real variable function defined on $a \leq x \leq b$ as:

$$F_a^b(x) = \int_a^x f(t)\, dt.$$

Graphically: [82]



Note that this function is 0 at point $x = a$ and it is the whole definite integral at point $x = b$:

$$F_a^b(a) = \int_a^a f(t)\, dt = 0 \text{ and: } F_a^b(b) = \int_a^b f(t)\, dt.$$

The Fundamental Theorem of the Integral Calculus claims that:[83]

Theorem: The integral function associated to a continuous one-real variable function is one of its primitives. In other words:

$$\frac{dF_a^b(x)}{dx} = f(x).$$

[82] Note that the integral function measures the yellow area

[83] It should be mentioned that this result, proved separately by Newton and Leibniz in the XVII century, was a real breakthrough in mathematics.

3.2.2.1. Calculating Definite Integrals: Barrow's Rule

The Barrow's rule appears as a consequence of the above fundamental theorem and allows us to calculate the definite integral of a continuous function provided it has primitives.[84]

Theorem (Barrow's rule): If $F(x)$ is a primitive of $y = f(x)$ then:

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Example: Calculate the definite integral:

$$\int_{-\pi/2}^{\pi/2} \left(\frac{\cos x}{2 - \sin x}\right) dx.$$

SOLUTION: Since the indefinite integral of this function is:

$$\int \left(\frac{\cos x}{2 - \sin x}\right) dx = -\ln(2 - \sin x) + C$$

we can take as a primitive the following function:

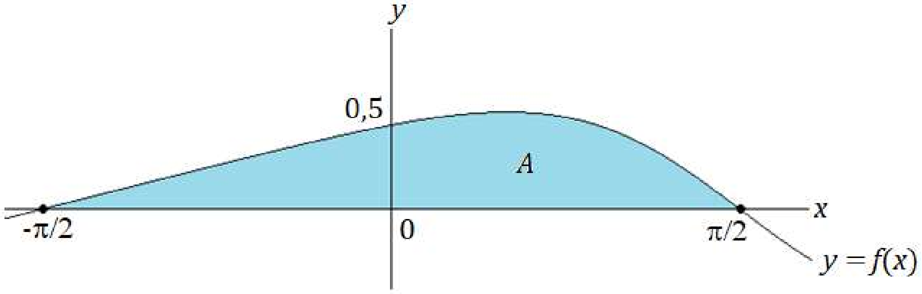
$$F(x) = -\ln(2 - \sin x).$$

Hence, we deduce from the Barrow's rule that:

$$\int_{-\pi/2}^{\pi/2} \left(\frac{\cos x}{2 - \sin x}\right) dx = F\left(\frac{\pi}{2}\right) - F\left(-\frac{\pi}{2}\right) = \left(-\ln\left(2 - \sin\left(\frac{\pi}{2}\right)\right)\right) - \left(-\ln\left(2 - \sin\left(-\frac{\pi}{2}\right)\right)\right) =$$

$$= \ln 3 \blacksquare$$

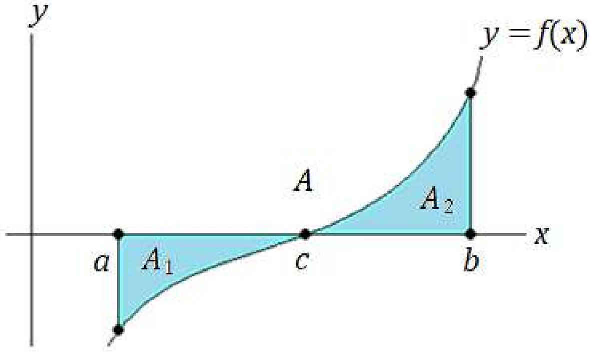
Since this function is positive between $-\pi/2$ and $\pi/2$ this definite integral matches the area determined by the function as illustrated:



[84] Unfortunately, the Barrow's rule cannot be applied on a function without primitives.

3.2.3. *Metric Applications of the Definite Integral* [85]

1. We want to calculate the area A determined as:



In this case:

$$A = A_1 + A_2 = \left|\int_a^c f(x)dx\right| + \left|\int_c^b f(x)dx\right|.$$ [86]

Example: Find the area determined by the function $f(x) = (x-1)^3 + 1,$ the lines $x = -1,$ $x = 1$ and the x–axis.

SOLUTION: Since this function cut the x-axis off at point 0:

$$0 = f(x) = (x-1)^3 + 1 \text{ implies: } x = 0$$

we deduce that the area A demanded is equal to:

$$A = A_1 + A_2 = \left|\int_{-1}^{0} ((x-1)^3 + 1)dx\right| + \left|\int_0^1 ((x-1)^3 + 1)dx\right|.$$

Since:

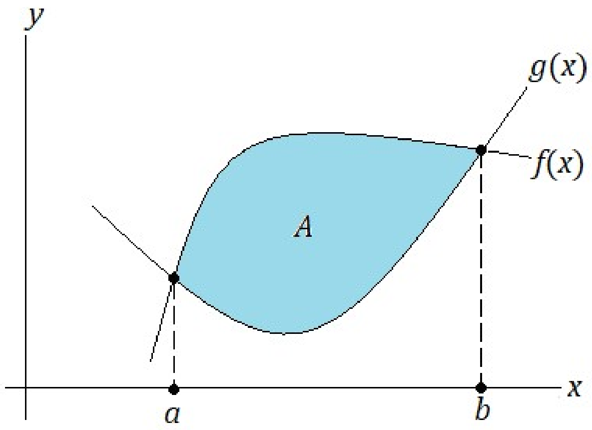
$$\int ((x-1)^3 + 1)\,dx = \frac{1}{4}\cdot(x-1)^4 + x + C \text{ implies: } F(x) = \frac{1}{4}\cdot(x-1)^4 + x$$

we conclude that the area is:

$$A = |F(0) - F(-1)| + |F(1) - F(0)| = \left|-\frac{11}{4}\right| + \left|\frac{3}{4}\right| = \frac{14}{4} = 3.5 \blacksquare$$

[85] It is not neccesary to know the graphics of the following áreas to find the required values.
[86] In the above situation it is not necessary to put absolute value in the second definite integral.

2. We want to calculate now areas enclosed between two functions as it is illustrated:



In this case we have:

$$A = \left| \int_a^b \big(f(x) - g(x)\big) dx \right| .$$[87]

Example: Find the area enclosed between the parabola $y = x^2 + 1$ and the line $x + y = 3$.

SOLUTION:

These two functions cut off at points $(-2,5)$ and $(1,2)$. Indeed:

$$x^2 + 1 = 3 - x \text{ implies: } x^2 + x - 2 = 0 \text{ implies: } x = -2 \text{ or } x = 1.$$

Therefore, the area A required will be equal to:

$$A = \left| \int_{-2}^{1} \big((x^2+1) - (3-x)\big) dx \right| = \left| \int_{-2}^{1} (x^2 + x - 2) dx \right| .$$

Givent that we can take as a primitive the function:

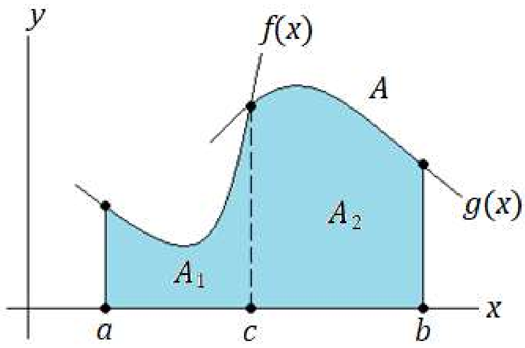
$$\int (x^2 + x - 2) dx = \frac{x^3}{3} + \frac{x^2}{2} - 2x + C \text{ implies: } F(x) = \frac{x^3}{3} + \frac{x^2}{2} - 2x$$

the area A is equal to:

$$A = \left| \int_{-2}^{1} (x^2 + x - 2) dx \right| = |F(1) - F(-2)| = \left| \left(-\frac{7}{6}\right) - \left(\frac{10}{3}\right) \right| = \frac{9}{2} = 4.5 \blacksquare$$

[87] We will always consider the absolute value of the definite integral of the "difference" function. Note that it does not matter which of the two functions is the minuend of the subtraction.

3. We want to calculate the area A determined by two "consecutive" functions:



In this case the area will be:

$$A = A_1 + A_2 = \left| \int_a^c f(x)dx \right| + \left| \int_c^b g(x)dx \right|.$$

<u>Example</u>: Calculate the area determined by the x-axis and the function:

$$f(x) = \begin{cases} \frac{1+x}{1-x}, x < 0 \\ \cos x, x \geq 0 \end{cases} \text{ between the points } x = -1 \text{ and } x = \frac{\pi}{2}.$$

SOLUTION: Since this function changes its definition at point $x = 0$, the area A will be:

$$A = \left| \int_{-1}^{0} \left(\frac{1+x}{1-x} \right) dx \right| + \left| \int_{0}^{\pi/2} \cos x \, dx \right|.$$

Bearing in mind that:

$$\int \left(\frac{1+x}{1-x} \right) dx = -x - 2\ln(1-x) + C \text{ implies: } F_1(x) = -x - 2\ln(1-x)$$

and:

$$\int \cos x \, dx = \sin x + C \text{ implies: } F_2(x) = \sin x$$

we conclude that:

$$A = \left| \int_{-1}^{0} \left(\frac{1+x}{1-x} \right) dx \right| + \left| \int_{0}^{\pi/2} \cos x \, dx \right| = |F_1(0) - F_1(-1)| + \left| F_2\left(\frac{\pi}{2}\right) - F_2(0) \right| =$$

$$= |0 - (1 - 2\ln 2)| + |1 - 0| = 2\ln 2 \cong 1.3863 \blacksquare$$

3.2.4. *Applications of Definite Integrals in Economics*

3.2.4.1. Marginal Economic Fucntions

The first application concerns with *marginal* economic functions such as the marginal cost, the marginal profit, etc. Consider this example:

Example: Let:

$$CM_a(q) = 2q - 10$$

be the marginal cost of the production of a good of consumption A in which the variable $q > 0$ stands for the number of output units. If the unit sales price of A is that of €520 and that all the output is sold: (1) Determine the function of cost of A if the total profit of 10 units amounts to €5100. (2) Calculate the output of A that maximizes the total profit.

SOLUTION:

(1) Recall that marginal cost $CM_a(q)$ is the derivative of the function of total cost $C(q)$. So, if we integrate the former function we get as a total cost function:

$$\frac{dC(q)}{dq} = CM_a(q) \text{ implies: } C(q) = \int CM_a(q)dq = \int (2q-10)dq = q^2 - 10q + C.$$

We must figure out the value of the integration constant through the hypotheses of the problem. Since the income function in euros is equal to:

$$I(q) = 520 \cdot q$$

then the profit function will be:

$$B(q) = I(q) - C(q) = 530q - q^2 - C.$$

Considering that total profit of 10 units of A amounts to €5100, we deduce that:

$$5100 = B(10) = 530 \cdot 10 - 10^2 - C \text{ implies: } C = €100.$$

Thus, the function of total cost of A is:

$$C(q) = q^2 - 10q + 100 \blacksquare$$

(2) Bearing in mind that:

$$0 = \frac{dB(q)}{dq} = 530 - 2q \text{ implies: } q = 265 \text{ and: } \frac{d^2B(265)}{dq^2} = -2 < 0$$

the maximum benefit is obtained by producing and selling 265 units of A with a maximum profit of:

$$B(265) = 530 \cdot 265 - 265^2 - 100 = €70{,}125 \blacksquare$$

3.2.4.2. The Consumer and Producer Surplus

Another interesting application involves finding both the *consumer* and the *producer surplus* of a good. Let us take a look at this example:

Example: If the demand and supply functions of a product P are given by:

$$p_D = 25 - q^2 \text{ and } p_S = 2q + 1$$

1. Determine the equilibrium price of P as well as the quantity demanded at that price.[88]
2. Find the consumer and the producer surplus of P associated with this price.

SOLUTION:

(1) We know that equilibrium price is given by the equality $p_S = p_D$. Thus, the equilibrium quantity $q_0 \geq 0$ demanded of P will be of:

$$p_S = p_D \text{ implies: } q^2 + 2q - 24 = 0 \text{ implies: } q_0 = 4 \text{ units.}$$ [89]

Hence the associated equilibrium price $p_0 \geq 0$ of P must be of:

$$p_0 = p_D(4) = p_S(4) \text{ implies: } p_0 = €9 ■$$

(2) Generally, the consumer surplus ε_D and the producer surplus ε_S of P associated with both the equilibrium price $p_0 \geq 0$ and quantity $q_0 \geq 0$ are given by the formulas:

$$\varepsilon_D = \left(\int_0^{q_0} p_D dq \right) - p_0 q_0 \text{ and: } \varepsilon_S = p_0 q_0 - \left(\int_0^{q_0} p_S dq \right).$$

Thus in our case we have:

$$\varepsilon_D = \left(\int_0^{q_0} p_D dq \right) - p_0 q_0 = \left(\int_0^{4} (25 - q^2) dq \right) - 9 \cdot 4 = €42.66 ■$$

and:

$$\varepsilon_S = p_0 q_0 - \left(\int_0^{q_0} p_S dq \right) = 9 \cdot 4 - \left(\int_0^{4} (2q + 1) dq \right) = €16 ■$$ [90]

[88] Equilibrium price appears when demand and supply matches. So, the quantity demanded at that price is called the "equilibrium quantity".

[89] Obviously, we do not consider the negative solution of this equation.

[90] It is not necessary to apply integrals in order to evaluate this last producer surplus since it is a triangle area.

3.2.4.3. Accumulating an Economic Variable

The third of these applications deals with the *stock levels* of a good and it specifically consists in calculating the time this product takes to run out under a strong increasing demand.

Example: It is known that the demand in tonnes for a rare mineral over the time is:

$$D(t) = 2 \cdot 10^6 \cdot e^{0.04}$$

where $t \geq 0$ denotes the passing years since the start of its demand. If the initial reserves amounted to 20,000 million tonnes estimate how long it will take for this rare mineral to run out.

SOLUTION:

Assuming that the global demand for this mineral is continuous in time,[91] the "total amount" $Q(t_0)$ demanded from the initial instant $t = 0$ until the year $t_0 > 0$ will be given by the definite integral of the "instantaneous" demand:

$$Q(t_0) = \int_0^{t_0} D(t)dt.$$

Given that the initial reserves amounted to 20,000 million tonnes, we need to find the moment $t_0 > 0$ in which the "total amount" demanded $Q(t_0)$ is exactly 20,000 million tonnes. Therefore:

$$20{,}000 \cdot 10^6 = Q(t_0) = \int_0^{t_0} D(t)dt = \int_0^{t_0} (2 \cdot 10^6 \cdot e^{0.04t})dt = \frac{2 \cdot 10^6}{0.04} \cdot (e^{0.04t_0} - 1)$$

from which, simplifying and taking logarithms, we deduce that:

$$400 = e^{0.04\ _0} - 1 \text{ implies: } e^{0.04t_0} = 401 \text{ implies: } t_0 = \frac{\ln(401)}{0.04} \cong 150.$$

Thus this rare mineral will take almost 150 years to run out■

[91] This assumption is essential to solve this problem by definite integrals.

3.3. First-Order Differential Equations

In order to put under the spotlight the importance of first-order differential equations, let us consider the famed example on "14-carbon radiactive isotop" C^{14} method of dating biological remains. This procedure is based on the fact that the ratio of C^{14} in every living organism remains constant throughout its life and, being no replaced, it decreases after death. We experimentally know that at any moment $t > 0$ after death ($t = 0$), the rate of decay of C^{14} satisfies the "first-order differential equation":[92]

$$\frac{dC^{14}(t)}{dt} = -1.24486 \cdot 10^{-4} \cdot C^{14}(t).$$

It can be proved that the "solution" of this equation is the one-variable function depending on time as independent variable:[93]

$$C^{14}(t) = C^{14}(0) \cdot e^{-1.24486 \cdot 10^{-4} \cdot t}.$$

The values $C^{14}(0)$ is the quantity of 14-carbon isotop that lies inside the organism at the moment of death.

As application, suppose that the remains of a bear found in the Pyrenees mountains contain only fifth of the C^{14} present in a living bear. How long this bear is dead? In this case the time passed after the bear's death $t_0 > 0$ satisfies the equation:

$$C^{14}(t_0) = \frac{1}{5} \cdot C^{14}(0).$$

Combining it with the previous formula we have:

$$\frac{1}{5} \cdot C^{14}(0) = C^{14}(t_0) = C^{14}(0) \cdot e^{-1.24486 \cdot 10^{-4} \cdot t_0}$$

that simplifying and taking logarithms we deduce as a value of $t_0 > 0$:

$$t_0 = \frac{\ln\left({}^{1}/_{5}\right)}{-1.24486 \cdot 10^{-4}} \cong 12{,}928.66.$$

Thus the bear died approximately 12,929 years ago. As a remarkable fact, observe that this time does not depend on the initial amount $C^{14}(0)$ of 14-carbon isotop.

[92] $C^{14}(t)$ is the quantity of 14-carbon isotope C^{14} at moment $t \geq 0$ after death. Note that this equation involves the derivative of a function (i.e., the 14-carbon isotope amount) and hence the name of differential equation.

[93] To solve this type of equations is what we are going to do here.

3.3.1. *First-Order Differential Equations and Solutions*

Given that this section is intended to serve as just an introduction to this topic, we will only study the first-order differential equations and among these the simplest. We begin by considering the following definition:

Definition: A **first-order differential equation** (**differential equation** from now onwards) is an equation that depends on a variable independent x, a variable dependent $y = f(x)$ and the first derivative of it:

$$y' = \frac{dy}{dx}.$$

Any differential equation can be expressed in:

1. "Implicit" form: $F(x, y, y') = 0$.
2. "Explicit" form: $y' = f(x, y)$.
3. "Continuous" form: $p(x, y)dx + q(x, y)dy = 0$.

Example: Determine the three forms of the following differential equations:

1. $\ln(y') + 2y = 0$
2. $y' = 1 + x + y$
3. $xdx + ydy = 0$

SOLUTION:

(1) In this case we have a differential equation expressed in implicit form. Taking logarithms, the explicit and continuous forms would be:

$$y' = e^{-2y} \text{ and: } e^{-2y} \cdot dx - dy = 0 \blacksquare$$

(2) This differential equation comes expressed in explicit form since the derivative y' is isolated. The implicit and continuous forms will be:

$$y' - y - x - 1 = 0 \text{ and: } dx - \left(\frac{1}{1 + x + y}\right) \cdot dy = 0 \blacksquare$$

(3) Now this differential equation is in continuous form. The implicit and explicit associated forms would be:

$$y \cdot y' + x = 0 \text{ and: } y' = -\frac{x}{y} \blacksquare$$

3.3.1.1. Solution of a First-Order Differential Equation

The concept of the solution of a differential equation is crucial. In fact, the aim here is to find the solutions of the most basic differential equations and, as we will see, indefinite integrals will be essential.

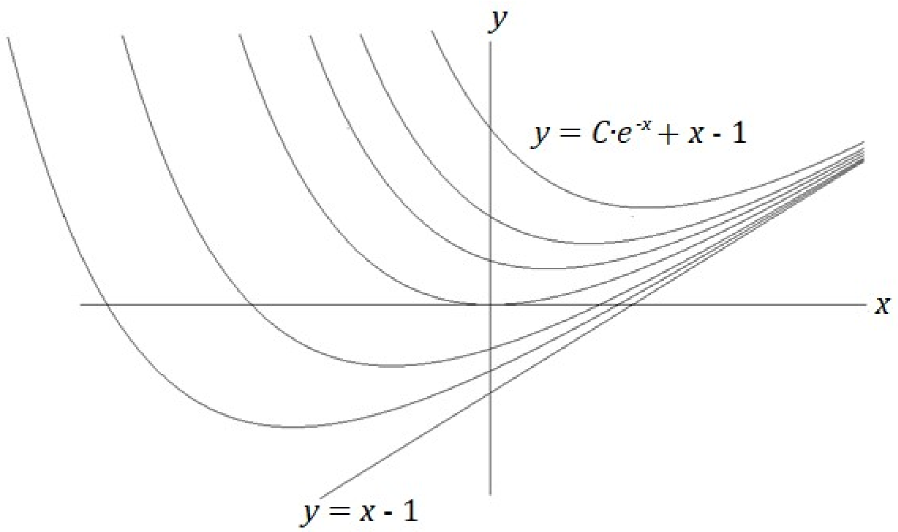
Definition: A **solution** of a first-order differential equation is any differentiable function that satisfies it. There are basically two types of solutions:

1. **General solution**: A solution depending on one real parameter.
2. **A particular solution**: Any solution obtained from the general solution when numerical values are given to its parameter.[94]

Example: Prove that the function $y = C \cdot e^{-x} + x - 1$, where $C \in \mathbb{R}$, is the general solution of the differential equation $y' + y = x$. Is the straight line $y = x - 1$ a particular solution?
SOLUTION:
Since we have a function depending on one parameter, we only need to check that it satisfies the differential equation, and this is the case. Indeed:

$$y = C \cdot e^{-x} + x - 1 \text{ implies: } y' + y = (-C \cdot e^{-x} + 1) + (C \cdot e^{-x} + x - 1) = x\blacksquare$$

Note that if we put $C = 0$ in the general solution we can conclude that the line $y = x - 1$ is one of their particular solutions (integral curves). The integral curves are graphically:



[94] Geometrically speaking these solutions are also called **integral curves**. As we are going to see in the next exemple, the general solution is the family formed by their integral curves.

3.3.2. *Integrating First-Order Differential Equations*

3.3.2.1. Separated Differential Equations

This type of differential equations represents the easiest case.[95]

Definition: **Separated differential equations** are differential equations that take as a continuous form the expression:

$$p(x)dx + q(y)dy = 0$$

where $p(x)$ and $q(y)$ are functions of one real variable.

Separated differential equations can be solved directly as long as both functions $p(x)$ and $q(y)$ accept indefinite integral. Let us look at the following example:

Example: Given the differentiable equation $xdx + ydy = 0$ find:

1. The general solution.
2. The particular solution (i.e., the integral curve) passing through the point (1,1).

SOLUTION:

(1) As we can see we have a separated differential equation with the identity functions:

$$p(x) = x \text{ and } q(y) = y.$$

Taking integrals in the continuous form of this differential equation we have:

$$\int xdx + \int ydy = C \text{ implies: } \frac{x^2}{2} + \frac{y^2}{2} = C \blacksquare$$ [96]

Thus the general solution is the family of the circumferences centered at point (0,0).

(2) So the integral curve passing through point (1,1) will be the circumference centered at (0,0) with radius $r = \sqrt{2}$. Indeed:

$$(x, y) = (1,1) \text{ implies: } \frac{1^2}{2} + \frac{1^2}{2} = C \text{ implies: } C = 1 \text{ implies: } x^2 + y^2 = 2 = \left(\sqrt{2}\right)^2 \blacksquare$$

[95] In fact, the differential equations that we are going to solve will have to be transformed into equations of this type since they are the only ones that can be directly integrated.

[96] The parameter C is the primitive of 0 and should be considered a "dummy" constant, namely a constant that may take any real value. In this case however only positive values.

3.3.2.2. Separable Differential Equations

Definition: **Separable differential equations** are those first-order differential equations that take as a continuous form:

$$\big(p_1(x)\cdot p_2(y)\big)dx+\big(q_1(x)\cdot q_2(y)\big)dy=0$$

where $p_1(x), p_2(y), q_1(x)$ and $q_2(y)$ are functions of one real variable.

Observe that these equations can be transformed into separated differential equations if we divide the whole equation by the product $q_1(x)\cdot p_2(y)$. Indeed:

$$\big(p_1(x)\cdot p_2(y)\big)dx+\big(q_1(x)\cdot q_2(y)\big)dy=0 \text{ implies: } \left(\frac{p_1(x)}{q_1(x)}\right)dx+\left(\frac{q_2(y)}{p_2(y)}\right)dy=0.$$

Consider the following example:

Example: Determine the general solution of:

$$ydx+xdy=0$$

as well as the associated integral curves.

SOLUTION:

We have a separable differential equation where:

$$p_1(x)=1, p_2(y)=y, q_1(x)=x \text{ and: } q_2(y)=1.$$

Dividing by $q_1(x)\cdot p_2(y)=xy$ we obtain the separated differential equation:

$$\left(\frac{1}{x}\right)dx+\left(\frac{1}{y}\right)dy=0$$

with general solution:

$$\int\left(\frac{1}{x}\right)dx+\int\left(\frac{1}{y}\right)dy=C \text{ implies: } \ln x+\ln y=C \text{ implies: } x\cdot y=e^C \blacksquare$$ [97]

Note that the associated integral curves (particular solutions) are the hyperbolas with the two coordinated axes as asymptotes■

[97] In this case we could substitute the positive constant e^C by another constant regardless of the sign.

3.3.2.3. Example of Application

Example: Find the general solution of the following first-order differential equations:

1. $y' - 3y = 17.$
2. $y' - \frac{2y}{x} = 0.$
3. $x^2 y' - y = 5.$

SOLUTION:

(1) Since:

$$\frac{dy}{dx} = y' = 3y + 17 \text{ implies: } \int \frac{dy}{3y+17} = \int dx \text{ implies: } \frac{1}{3} \cdot \ln(3y+17) = x + C$$

the general solution is:

$$\frac{1}{3} \cdot \ln(3y+17) = x + C \text{ implies: } 3y + 17 = e^{3(x+C)} \text{ implies: } y = \frac{e^{3(x+C)} - 17}{3} \blacksquare$$

(2) In this case we have:

$$\frac{dy}{dx} = y' = \frac{2y}{x} \text{ implies: } \int \frac{dy}{y} = \int \frac{2dx}{x} \text{ implies: } \ln y = 2 \ln x + C.$$

Then the general solution is:

$$\ln y = 2 \ln x + C = \ln x^2 + C \text{ implies: } y = e^{\ln x^2 + C} = e^{\ln\ \ ^2} \cdot e^C = x^2 \cdot e^C \blacksquare$$ [98]

(3) Now dividing by x^2 we have the separated differential equation:

$$y' - \left(\frac{1}{x^2}\right) y = \frac{5}{x^2} \text{ implies: } \frac{dy}{dx} = y' = \frac{y+5}{x^2} \text{ implies: } \frac{dy}{y+5} = \frac{dx}{x^2}$$

with general solution:

$$\int \frac{dy}{y+5} = \int \frac{dx}{x^2} \text{ implies: } \ln(y+5) = -\left(\frac{1}{x}\right) + C \text{ implies: } y = e^{-\left(\frac{1}{x}\right)+C} - 5 \blacksquare$$ [99]

---

[98] Note that the integral curves of this differential equation are all the parabolas with vertex at point (0,0).

[99] Observe that in this case the constant function $y = -5$ is a solution that cannot be deduced from the general solution given values to the constant C. Some differential equations have solutions that are neither the general solution nor any particular one; they are called **singular** solutions.

3.3.3. *Applications in Economics*

3.3.3.1. Instant Compound Interest

The next mathematical model formalizes a financial scenario in which monetary interest becomes capital in every instant of time.

Example: Supose that the instant rate of increase of a certain amount of money $C = C(t)$, dependig on the temporal variable $t \geq 0$ measured in years, is proportional to the capital $C(t)$ at that moment according to the differential equation:

$$\frac{dC(t)}{dt} = i \cdot C(t)$$

where $0 < i < 1$ is a constant (called "instantaneous interest rate"). Find: (1) The general expression of the capital $C(t)$ depending on the initial amount of money $C_0 = C(0)$. (2) If $i = 3\%$, how long must we wait for an initial capital C_0 to double?

SOLUTION:[100] (1) Since:

$$\frac{dC(t)}{dt} = i \cdot C(t) \text{ implies: } \int \frac{dC(t)}{C(t)} = \int i \cdot dt \text{ implies: } \ln C(t) = it + C$$

the general solution is:

$$C(t) = e^{it+C}.$$

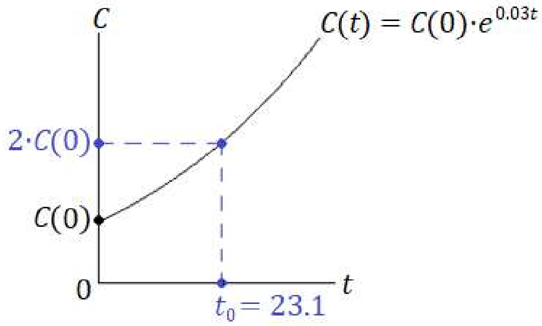
Knowing that $C_0 = C(0)$ we obtain as a general solution depending on the initial amount:

$$C_0 = C(0) = e^{i\cdot 0+C} = e^C \text{ implies: } C(t) = e^{it+C} = e^C \cdot e^{it} = C_0 \cdot e^{i\cdot t}\blacksquare$$

(2) In the particular case of $i = 3\%$, the time $t_0 > 0$ we must wait until $C(t_0) = 2 \cdot C_0$ is:

$$2 \cdot C_0 = C(t_0) = C_0 \cdot e^{0.03\cdot t_0} \text{ implies: } 2 = e^{0.03\cdot t_0} \text{ implies: } t_0 = \frac{\ln 2}{0.03} = 23.1 \text{ years}\blacksquare$$

Graphically:



[100] This differential equation is of the same type as the "14-carbon radiactive isotop" C^{14} case.

3.3.3.2. Temporary Evolution of Prices: The Walras's Linear Model

In this paragraph we analyze this model of price evolution using a specific example.

Example: Suppose that at moment $t \geq 0$ the rate of growth of a price $p(t)$ of a good A satisfies the differential equation:

$$\frac{dp(t)}{dt} = 0.5 \cdot \big(D(t) - S(t)\big)$$

being the functions:

$$D = D(t) = 21 - 2 \cdot p(t) \text{ and: } S = S(t) = 10 \cdot p(t) - 3$$

the demand and the supply of A at moment $t \geq 0$. Study the temporary evolution of $p(t)$.
SOLUTION:
Observe that the formal expression of $p(t)$ satisfies the differential equation on $p(t)$:

$$\frac{dp(t)}{dt} = 0.5 \cdot \big(D(t) - S(t)\big) = 0.5 \cdot \Big(\big(21 - 2p(t)\big) - \big(10p(t) - 3\big)\Big) = 12 - 6p(t)$$

with general solution:

$$\int \frac{dp(t)}{p(t) - 2} = -\int 6dt \text{ implies: } \ln(p(t) - 2) = -6t + C \text{ implies: } p(t) = 2 + e^{-6t+C}.$$

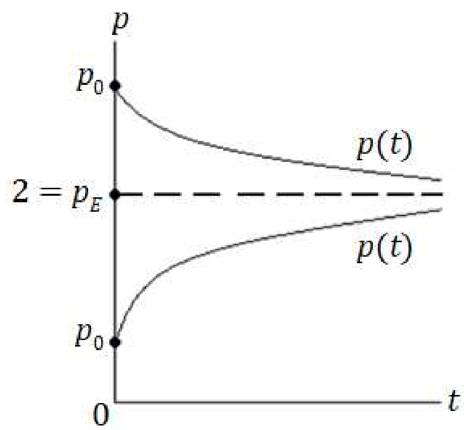
We can easily see that the price $p(t)$ is stable over time in the sense that it converges as time goes by:

$$\lim_{t \to +\infty} p(t) = 2 + e^{-6\cdot\infty + C} = 2 + e^{-\infty} = \{e^{-\infty} = 0\} = 2 \blacksquare$$

It is worth noting that this limit value is precisely the "equilibrium price" p_E of A. Indeed:

$$D = S \text{ implies: } 21 - 2 \cdot p_E = 10 \cdot p_E - 3 \text{ implies: } p_E = 2.$$

Graphically:



3.3.3.3. Application on Demography: the Logistic Growth

The last example is related to the growth of one variable that does not depen only on its amount at any moment of time.

Example: Suppose that at time $t \geq 0$ the number $N = N(t)$ in thousands of individuals of a certain population of 160,000 people awared of a rumor, is proportional to both its amount and that of the unawared individuals according to the differential equation:

$$N' = \frac{dN}{dt} = 0.02 \cdot N \cdot (160 - N).$$

Determine the population that will be aware of the rumor within half a year ($t = 0.5$) if it was initially ($t = 0$) of 20,000 individuals. What will eventually happen?

SOLUTION:[101] Dividing this differential equation by N^2 we obtain a differential equation depending on the variable $Y = {}^{1}\!/_{N}$ that we know how to solve:

$$Y' = \frac{dY}{dt} = -3.2 \cdot Y + 0.02.$$

The solution of this differential equation helps us to obtain the particular solution of the initial one with the condition $N(0) = 20$. It can be proved that this solution is:

$$N(t) = \frac{160}{1 + 7 \cdot e^{-3.2t}}.$$

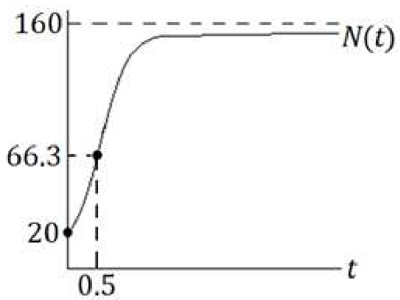
Thus the population aware of the rumor within half a year will be approximately of:

$$N(0.5) = \frac{160}{1 + 7 \cdot e^{-3.2 \cdot 0.5}} = 66.3 \text{ thousand individuals} \blacksquare$$

Eventually the entire population will be aware of the rumor since:

$$\lim_{t \to +\infty} N(t) = \frac{160}{1 + 7 \cdot e^{-\infty}} = \{e^{-\infty} = 0\} = 160 \blacksquare$$

Graphically:[102]



[101] We are goint to offer only an outline of the resolution.

[102] This curve is known as a *logistic curve.*

3.4. Exercises

1. Calculate the following indefinite integrals:

$$(1) \int \left(\frac{x^2-7}{\sqrt{x}}\right) dx. \ (2) \int \frac{dx}{1-2x}. \ (3) \int \left(\frac{5x}{e^{x^2}}\right) dx. \ (4) \int \frac{2dx}{x \ln x}. \ (5) \int \left(\frac{x^2+2x}{x^3+3x^2-1}\right) dx.$$

$$(6) \int \left(\frac{x^2-x}{1+x}\right) dx. \ (7) \int \frac{dx}{(x-3)^5}. \ (8) \int \left(\frac{e^{2x}}{1+e^{2x}}\right) dx. \ (9) \int \frac{dx}{\sqrt{3x+5}}. \ (10) \int \left(\frac{2^{1/x}}{x^2}\right) dx.$$

2. Calculate the following definite integrals:

$$(1) \int_1^4 \left(\frac{1+\sqrt{x}}{x^2}\right) dx. \ (2) \int_e^{e^2} \left(\frac{\ln x}{2x}\right) dx. \ (3) \int_0^{\ln} \ (1+e^{2x}) dx. \ (4) \int_0^{0.25} \left(\frac{0.5}{\sqrt{1-x}}\right) dx.$$

3. Calculate the area A determined by the function $f(x) = x \cdot e^{x^2}$ and the x-axis between the two lines $x = -1$ and $x = 1$.
4. Find the area A closed by the function $y = \frac{1}{2+x}$ and the straight line $15y + x = 6$.
5. Calculate the area A determined by the "consecutive" function:

$$f(x) = \begin{cases} \left(\frac{3}{2}\right) x + 3, x < 0 \\ \frac{3}{\sqrt{1+x}}, x \geq 0 \end{cases}$$

   between the lines $x = -1$ and $x = 3$.
6. The marginal cost function in euros associated with a manufacturer of electric cars is $CM_a(q) = 0.8q + 4$. If this manufacturer produces 50 units calculate how much it would cost to double the production.
7. Find the general solution of the following differential equations:

$$(1)\ xdx = (1-x^2)dy. \ (2)\ xdy + (1+y)^2 dx = 0. \ (3)\ exdx = 2ye^{-x^2} dy.$$

$$(4)\ y' = e^{y-x}. \ (5)\ y = \ln(y'). \ (6)\ xy' + y^2 = 0.$$

$$(7)\ 2y' - 3y = 12. \ (8)\ y' + 3x^{-2} y = 0. \ (9)\ 6y' + 3x^2 y = 2x^2.$$

8. Find the demand function $Q = f(P)$ if the elasticity ϵ is -1 for all prices $P > 0$.[103]

[103] Dowling, E. T. (2010), exercise 16.38, page 382. The **elasticity** (or **elastic derivative**) of a differentiable function $y = f(x)$ is defined by:

$$\epsilon = \frac{x}{f(x)} \cdot \frac{df(x)}{dx}.$$

In this case we want to obtain those demand functions associated to a good for which an increase of $r\%$ in price generates approximately a decrease of $r\%$ in its demand.

SOLUTIONS:

1.

$$(1)\ \frac{2x^{5/2}}{5} - 14x^{1/2} + C.\ (2)\ -\frac{1}{2}\cdot\ln(1-2x) + C.\ (3)\ -\left(\frac{5e^{-x^2}}{2}\right) + C.\ (4)\ 2\ln(\ln x) + C.$$

$$(5)\ \frac{1}{3}\ln(x^3+3x^2-12) + C.\ (6)\ \frac{x^2}{2} - 2x + 2\ln(x+1) + C.\ (7)\ -\left(\frac{1}{4(x-3)^4}\right) + C.$$

$$(8)\ \frac{1}{2}\cdot\ln(1+e^{2x}) + C.\ (9)\ \frac{2}{3}\sqrt{3x+5} + C.\ (10)\ -\left(\frac{2^{1/x}}{\ln 2}\right) + C.$$

2.

$$(1)\ \frac{7}{4}.\ (2)\ \frac{3}{4}.\ (3)\ \ln 2 + \frac{3}{2} \cong 2.19.\ (4)\ 1 - \frac{\sqrt{3}}{2} \cong 0.134.$$

3. $A = e - 1 \cong 1.718.$
4. $A = \frac{8}{15} + \ln\left(\frac{3}{5}\right) \cong 0.0225.$
5. $A = \frac{33}{4} = 8.25.$
6. €3200.
7.

$$(1)\ y = C - 0.5\cdot\ln(1-x^2).\ (2)\ y = \left(\frac{1}{\ln x + C}\right) - 1.\ (3)\ y^2 = \frac{e^{1-x^2}}{2} + C.$$

$$(4)\ y = \ln\left(\frac{1}{e^{-x}+C}\right).\ (5)\ y = \ln\left(\frac{1}{C-x}\right).\ (6)\ y = \frac{1}{\ln x + C}.$$

$$(7)\ y = \left(\frac{1}{3}\right)\cdot e^{1.5x+C} - 4.\ (8)\ y = e^{3/x+C}.\ (9)\ y = \frac{2-e^{-0.5\left(x^3/3+C\right)}}{3}$$

8. The demand function is $Q = f(P) = \frac{C}{P}$, with $C > 0$ as a constant.

REFERENCES

ADILLON, R.; ALVAREZ, M.; GIL, D.; JORBA, L. (2015) *Mathematics for Economics and Business*. Economy UB.

ALEGRE, P.; GONZÁLEZ, L.; ORTÍ, F.; RODRÍGUEZ, G.; SÁEZ, J.; SANCHO, T. (1995). *Matemáticas Empresariales* (AC, Madrid).

DOWLING, E. T. (2010). *Introduction to Mathematical Economics* (Schaum's Outline Series, 3rd. ed.) McGraw Hill.

LIPSCHUTZ, S.; LIPSON, M. L. (2009). *Linear Algebra* (Schaum's Outline Series, 4th. ed.) McGraw Hill.

SYDSAETER, K.; HAMMOND, P.J. (1995). *Mathematics for Economic Analysis*. Prentice Hall.

SYDSAETER, K.; HAMMOND, P.; SEIERSTAD, A.; STRØM, A. (2008). *Further Mathematics for Economic Analysis*. Prentice Hall.

GLOSSARY

Angle between vectors, 14
Barrow's rule, 61
Basis of a vector space, 7
Consumer/producer surplus, 66
Definite integral, 58
Direct-case method, 47
Dimension of a vector space, 7
Distance between vectors, 15
Domain of a function, 25
Elasticity of a differentiable function, 77
Extreme points of a function, 38
Euclidian vector space, 11
Extreme value of a function, 38
First-order differential equation, 69
Function of several variables, 25
Gradient vector, 34
Hessian matrix, 37
Hyperplane, 16
Immediate integral, 54
Indefinite integral, 53
Instantaneous compound interes, 74
Integral curve, 70
Integral function, 60
Level curve, 27
Linear combination of vectors, 6
Linear (in)dependence of vectors, 7
Linear regression, 46
Logarithm integration, 55
Logistic curve, 76
Marginal economic functions, 65
Marginalism, 32
Marshall's model of consumption, 48
Norm of a vector, 13
Maximum/minimum of a function, 38
Optima of a function, 38
Optimal value of a function, 38
Orthogonal vectors, 14
Partial derivative of a function, 29
Partial elasticity of a function, 33
Position-vector of a point, 11
Primitive of a function, 52
Principal minor, 19
Quadratic form, 17
Radiactive isotop method, 68
Saddle point of a function, 40
Scalars of a vector space, 5
Scalar product, 11
Second-order partial derivative, 36
Separated differential equation, 71
Separable differential equation, 72
Sign of a quadratic form, 18
Solution of a differential equation, 70
Stationary point of a function, 40
Stock-level variation, 67
Unit vector, 13
Vector space, 5
Vectors of a vector space, 5
Walras's linear price model, 75
Zero-vector, 5