# PUBLICACIONES DE PRIMER CURSO

**Grado: ADEi**

**Curso: 1st**

**Asignatura: Mathematics I**

**Grupos: 100**

**Tema: Theoretical Notes**

**Profesores: Miguel Puente Ajovin**

**Departamento de Análisis Económico**

**Curso Académico 2018/19**

Facultad de Economía y Empresa
Universidad Zaragoza

# Matrices

For more details of what follows in this Section see the *Capítulo 2* of Jarne et al. (1997) and the *Unidad didáctica 6* of `www.unizar.es/aragon_tres`

## Matrices

### Examples of matrices

$$\begin{aligned} A &= \begin{pmatrix} 1 & 8 & 3 \\ 2 & 7 & 4 \end{pmatrix} \in M_{2\times 3} \\ B &= \begin{pmatrix} 2 & 1 & 3 \\ -3 & 2 & 4 \\ 4 & 5 & 6 \end{pmatrix} \in M_{3\times 3} \equiv M_3 \\ C &= \begin{pmatrix} 3 & 7 & -2 \end{pmatrix} \in M_{1\times 3} \end{aligned}$$

$A$ is said to be *rectangular* (of order $2 \times 3$), $B$ is said to be *square* (of order 3) and $C$ is a *row matrix* (of three elements).

**Transposition**: To interchange rows by columns. With the examples above,

$$\begin{aligned} A^t &= \begin{pmatrix} 1 & 2 \\ 8 & 7 \\ 3 & 4 \end{pmatrix} \in M_{3\times 2} \\ B^t &= \begin{pmatrix} 2 & -3 & 4 \\ 1 & 2 & 5 \\ 3 & 4 & 6 \end{pmatrix} \in M_3 \\ C^t &= \begin{pmatrix} 3 \\ 7 \\ -2 \end{pmatrix} \in M_{3\times 1} \end{aligned}$$

We have that $C^t$ is a *column* matrix. It holds that $(A^t)^t = A, \forall A \in M_{m\times n}$.

**Symmetric matrix**: $A$ is symmetric if and only if $A^t = A$. Example:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & -7 \\ 4 & -7 & 8 \end{pmatrix}$$

**Diagonal matrix**: Example:

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

**Scalar matrix**: Example:

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

A very important particular case is $I_n$, namely the identity matrix of order $n$. Example for $n = 3$:

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

**Triangular matrices**: Examples:

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & -7 \\ 0 & 0 & 8 \end{pmatrix}$$

is an *upper triangular* matrix.

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & -7 & 8 \end{pmatrix} = A^t$$

is a *lower triangular* matrix.

**Examples of matrix operations**: Let

$$
\begin{aligned}
A &= \begin{pmatrix} 2 & 10 \\ -2 & 0 \\ 4 & 1 \end{pmatrix} \in M_{3\times 2} \\
B &= \begin{pmatrix} 0 & -10 \\ -5 & 3 \\ 2 & 9 \end{pmatrix} \in M_{3\times 2} \\
C &= \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in M_{2\times 3}
\end{aligned}
$$

Then we have the *sum*

$$
A + B = \begin{pmatrix} 2 & 10 \\ -2 & 0 \\ 4 & 1 \end{pmatrix} + \begin{pmatrix} 0 & -10 \\ -5 & 3 \\ 2 & 9 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ -7 & 3 \\ 6 & 10 \end{pmatrix} \in M_{3\times 2}
$$

Exercise: check that $(A+B)^t = A^t + B^t$ for this example. This property holds for all $A, B$ of the same order.

*Product by an escalar*: Example:

$$
2A = 2\begin{pmatrix} 2 & 10 \\ -2 & 0 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 20 \\ -4 & 0 \\ 8 & 2 \end{pmatrix} \in M_{3\times 2}
$$

Exercise: check that $(2A)^t = 2A^t$ for this example. In general it holds $(\lambda A)^t = \lambda A^t$ for any $\lambda \in \mathbb{R}$ and any matrix $A$.

Note that, with the previous matrices and operations, that the difference $A - B$ is just

$$
A - B = A + (-1)B = \begin{pmatrix} 2 & 10 \\ -2 & 0 \\ 4 & 1 \end{pmatrix} - \begin{pmatrix} 0 & -10 \\ -5 & 3 \\ 2 & 9 \end{pmatrix} = \begin{pmatrix} 2 & 20 \\ 3 & -3 \\ 2 & -8 \end{pmatrix} \in M_{3\times 2}
$$

Exercise: Check that $(A-B)^t = A^t - B^t$ for this example. This property is valid for any $A, B$ of the same order.

*Matrix product*: We have that

$$
\underset{3\times 2}{A} \cdot \underset{3\times 2}{B}
$$

does not exist.

But we have that

$$\underset{3\times 2}{A} \cdot \underset{2\times 3}{C} = \begin{pmatrix} 2 & 10 \\ -2 & 0 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2\cdot 1 + 10\cdot 0 & 2\cdot 2 + 10\cdot 1 & 2\cdot 1 + 10\cdot 0 \\ -2\cdot 1 + 0\cdot 0 & -2\cdot 2 + 0\cdot 1 & -2\cdot 1 + 0\cdot 0 \\ 4\cdot 1 + 1\cdot 0 & 4\cdot 2 + 1\cdot 1 & 4\cdot 1 + 1\cdot 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 14 & 2 \\ -2 & -4 & -2 \\ 4 & 9 & 4 \end{pmatrix} \in M_3$$

and also

$$\underset{3\times 2}{B} \cdot \underset{2\times 3}{C} = \begin{pmatrix} 0 & -10 \\ -5 & 3 \\ 2 & 9 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -10 & 0 \\ -5 & -7 & -5 \\ 2 & 13 & 2 \end{pmatrix} \in M_3$$

$$\underset{2\times 3}{C} \cdot \underset{3\times 2}{A} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 10 \\ -2 & 0 \\ 4 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 11 \\ -2 & 0 \end{pmatrix} \in M_2$$

Note that the matrix product **is not commutative**. In general, it holds $(A \cdot B)^t = B^t \cdot A^t$, for all matrices $A, B$ of appropriate orders.

Exercise: check that $I_3 \cdot A = A$, $A \cdot I_2 = A$, $(A \cdot C)^t = C^t \cdot A^t$ for this example.[1]

## Determinants

The determinant is a mapping from square matrices to real numbers, $A \in M_n \to \det(A) = |A| \in \mathbb{R}$ which is computed in the following way:

If $n = 1$, e.g., $A = 2$, then $\det(A) = 2$.

If $n = 2$, e.g., $A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$, then $\det(A) = \begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = 2 \cdot 2 - 3 \cdot 1 = 1$

[1]We will omit the $\cdot$ in the matrix product when there be no risk of confusion.

If $n = 3$, e.g., $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -5 & 6 \\ 4 & -1 & -2 \end{pmatrix}$, then

$$\begin{aligned} \det(A) &= \begin{vmatrix} 1 & 2 & 3 \\ 0 & -5 & 6 \\ 4 & -1 & -2 \end{vmatrix} \\ &= 1 \cdot (-5) \cdot (-2) + 2 \cdot 6 \cdot 4 + 0 \cdot (-1) \cdot 3 \\ &\quad -3 \cdot (-5) \cdot 4 - 6 \cdot (-1) \cdot 1 - 0 \cdot 2 \cdot (-2) \\ &= 10 + 48 + 0 + 60 + 6 + 0 = 124 \end{aligned}$$

If $n > 3$, we will need the following concepts:

**Complementary minor**. Example: If

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

the *complementary minor* of $a_{23} = 6$ is

$$\begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} = -1 - 8 = -9$$

**Cofactor element**. Example with the same matrix: The *cofactor element* $A_{23}$ is

$$A_{23} = (-1)^{2+3} \cdot \text{complementary minor of } a_{23} = (-1)^5 \cdot (-9) = 9$$

**Proposition 1** *Let $A \in M_n$. Then the following identities hold:*

- *Expansion by the $i$-th row:*

$$\det(A) = |A| = a_{i1}A_{i1} + \cdots + a_{in}A_{in}\,, \quad i \in \{1, \ldots, n\}$$

- *Expansion by the $j$-th column:*

$$\det(A) = |A| = a_{1j}A_{1j} + \cdots + a_{nj}A_{nj}\,, \quad j \in \{1, \ldots, n\}$$

- *Also, it holds:*

$$\begin{aligned} a_{i1}A_{i'1} + \cdots + a_{in}A_{i'n} = 0\,, &\quad i \neq i' \\ a_{1j}A_{1j'} + \cdots + a_{nj}A_{nj'} = 0\,, &\quad j \neq j' \end{aligned}$$

**Properties of the determinants**

1.

$$\begin{vmatrix} a_{11} & \cdots & a_{1j}+b_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j}+b_{2j} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nj}+b_{nj} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \cdots & b_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & b_{2j} & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_{nj} & \cdots & a_{nn} \end{vmatrix}$$

and analogously for any other column or row.

2. If in a square matrix they are interchanged two rows (columns), the determinant changes sign.

3. If in a square matrix a row (column) is multiplied by a real number, the determinant of the resulting matrix is equal to the determinant of the initial matrix times the cited real number.

4. If in a square matrix we add to a row (column) a real multiple of another row (column), the determinant does not change.

5. The determinant of a square matrix with a row (column) with null elements is zero.

6. The determinant of a square matrix with two equal rows (columns) is zero.

7. The determinant of a square matrix with two proportional rows (columns) is zero.

8. If $A$ and $B$ are two square matrices of the same order, then $|AB| = |A||B|$.

9. If $A$ is a square matrix, $|A^t| = |A|$.

10. If $A$ is a square matrix of order $n$ and $\lambda$ is a real number, then $|\lambda A| = \lambda^n |A|$.

11. If $A$ is an invertible matrix (see below), then $|A^{-1}| = |A|^{-1}$.

12. The determinant of a triangular matrix is equal to the product of the elements of the main diagonal.

The calculation of the determinants is greatly simplified using some of these properties, with the aim of obtaining the maximum number of zeros in the row (column) chosen in order to expand the determinant.

**Inverse matrix**

Let $A \in M_n$ be a square matrix of order $n$. Then $A$ is said to be *regular* or *invertible* if and only if $\det(A) = |A| \neq 0$.

Given $A \in M_n$ regular, its *inverse* $A^{-1}$ is the matrix such that

$$A^{-1} \cdot A = A \cdot A^{-1} = I_n$$

**Calculation of the inverse matrix by cofactors**

Let $A \in M_n$, then $\mathrm{Adj}(A) \in M_n$ is the matrix of the cofactors of $A$.

If $A \in M_n$ is regular, we have the following formula for the inverse:

$$A^{-1} = \frac{1}{|A|}(\mathrm{Adj}(A))^t \tag{1}$$

**Calculation of the inverse matrix by elementary transformations**

Let $A \in M_{m\times n}$. The elementary transformations are the following:

i) Interchange two rows or two columns.

ii) Multiply a row or column by a real number **different from zero**.

iii) Add to a row (resp. column) a linear combination of the other rows (resp. columns).

When performing any elementary transformation of the three cited types on $A \in M_{m\times n}$ it is obtained a matrix $B \in M_{m\times n}$, in general **different** from $A$, but called **equivalent** to $A$. It is written as follows:

$$A \sim B \in M_{m\times n}$$

The calculation of an inverse matrix by elementary transformations is based on the following equivalences, if $A \in M_n$ is regular:

Following row elementary transformations:

$$(A|I_n) \sim (I_n|A^{-1})$$

Following column elementary transformations:

$$\begin{pmatrix} A \\ - \\ I_n \end{pmatrix} \sim \begin{pmatrix} I_n \\ - \\ A^{-1} \end{pmatrix}$$

**Properties of the inverse matrix**

- If $A$ is regular, $|A^{-1}| = |A|^{-1}$.
- If $A^t$ is regular, $(A^t)^{-1} = (A^{-1})^t$.
- If $A, B \in M_n$ are regular, $(AB)^{-1} = B^{-1}A^{-1}$.
- If $A$ is regular and $\lambda \in \mathbb{R}$ with $\lambda \neq 0$, $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$.

## Rank of a matrix

**Definition 1 (Minor of order $h$)** *Let $A \in M_{m\times n}$. The determinant of a square submatrix of $A$ of order $h$ is called* minor of order $h$ *of $A$. If they are taken the $h$ first rows and the $h$ first columns of $A$ it is called* main minor *of order $h$.*

**Definition 2 (Rank of $A$ (by minors))** *Let $A \in M_{m\times n}$. If $\mathrm{rk}(A) = p$ we mean that:*

*i) There exists a non-vanishing minor of order $p$.*

*ii) All the minors of order greater than $p$ vanish or do not exist.*

Note that if $A \in M_{m\times n}$, then $\mathrm{rk}(A) \leq \min\{m, n\}$.

**Proposition 2** *The elementary transformations described previously leave invariant the rank of the matrix they are applied to.*

This proposition allows us to obtain from $A \in M_{m\times n}$ an equivalent matrix $B$, called *matrix in row echelon form of $A$*, which is characterized by the property of having in each row a minimum of one zero more at the beginning of the row than the preceding row.

A matrix in row echelon form $B$ of $A$ has the great advantage that its rank is directly the number of its non-vanishing rows. Since $A$ is equivalent to $B$ we have:

$$A \sim B \Rightarrow \mathrm{rk}(B) = \text{number of non vanishing rows of } B = \mathrm{rk}(A)$$

**Proposition 3 (Properties of the rank)**

*i)* $\mathrm{rk}(A) = \mathrm{rk}(A^t)$, $\forall A \in M_{m\times n}$.

*ii) Given $A \in M_n$, we have that $|A| \neq 0 \Leftrightarrow \mathrm{rk}(A) = n$, and that $|A| = 0 \Leftrightarrow \mathrm{rk}(A) < n$.*

*iii) If $A \in M_{m\times n}$, $\mathrm{rk}(A)$ is the number of linearly independent rows or columns of $A$.*

## Systems of linear equations

A linear system is written in *matrix form* as

$$AX = B$$

where $A \in M_{m\times n}$ is the *coefficient matrix*, $X \in M_{n\times 1}$ is the column matrix of *unknowns*, and $B \in M_{m\times 1}$ is the column matrix of *independent terms*. With these conditions, the *augmented matrix* of the system is

$$(A|B) \in M_{m\times(n+1)}$$

A *solution* of the previous system is $C \in M_{n\times 1}$ such that $AC \equiv B$.
Two systems are said to be *equivalent* if they have the same solutions.

### Classification of linear systems

i) According to the existence and uniqueness of solutions:

Inconsistent system (IS): there exists no solution.

Consistent system (CS): There exists at least one solution. It is determined (DCS) if the solution is unique, and indetermined (ICS) if there exists more than one solution.

ii) According to the independent terms:

Homogeneous system: $B = 0$ ($AX = 0$).

Inhomogeneous or complete system: $B \neq 0$.

**Theorem 1 (Rouché-Frobenius)** *The linear system $AX = B$ of $m$ equations with $n$ unknowns is consistent* $\Leftrightarrow \mathrm{rk}(A) = \mathrm{rk}(A|B) = r$. *Moreover if*

$r = n \Rightarrow$ *the system is DCS.*

$r < n \Rightarrow$ *the system is ICS.*

With the conditions of this Theorem, in the case ICS, the number $n - r$ is called *number of degrees of freedom of the system*, meaning that out of the $n$ unknowns, it can be chosen $r$ of them (*main unknowns*) in terms of the $n - r$ remaining ones (*degrees of freedom*), which serve as parameters for the solutions.

### Gaussian elimination

This method consists on associating to a linear system its augmented matrix, which by means of row elementary transformations are carried to a matrix in row echelon form, equivalent to the former.

In this row echelon form it is very easy to determine the rank of the coefficient matrix and of the augmented matrix, finding the consistent cases by using the Rouché-Frobenius Theorem.

The row echelon form can be associated to a linear system "in echelon form", equivalent to the given one and very easy to solve "in cascade".

**Cramer's Rule**

A system $AX = B$ of $n$ equations with $n$ unknowns is said to be *Cramer* $\Leftrightarrow$ $|A| \neq 0 \Leftrightarrow \text{rk}(A) = n$.

Then $\text{rk}(A|B) = n = \text{rk}(A) \Rightarrow$ the system is DCS. Also, since $|A| \neq 0$ there exists $A^{-1}$, and therefore

$$AX = B \Rightarrow A^{-1}AX = A^{-1}B \Rightarrow X = A^{-1}B$$

It can be proved, using the formula (1), that the components of $X$ are given by

$$x_i = \frac{1}{|A|} \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix}, \quad \forall i = 1, \ldots, n,$$

where $(a_{ij}) = A$, $(b_i) = B$. These formulae are known as *Cramer's Rule*.

**Homogeneous systems**

- An homogeneous system $AX = 0$ is **always consistent** because $\text{rk}(A) = \text{rk}(A|0)$.
- This system **has always the vanishing solution**: $A0 \equiv 0$.
- The remaining question is whether there exists any other solution apart from the vanishing solution, or the only solution is the vanishing one:
  - If $\text{rk}(A) = n = \text{number of unknowns} \Rightarrow$ the system is DCS (unique solution: the vanishing one).
  - If $\text{rk}(A) < n \Rightarrow$ the system is ICS (there exist at least one solution apart form the vanishing one).

# $\mathbb{R}^n$ set

**Definition 3** *Let $n$ be a natural number: $n \in \mathbb{N}$. Then we define*

$$\mathbb{R}^n = \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ times}}$$

*where* $\times$ *denotes the* Cartesian product. *In other words,*

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) \,|\, x_1, \ldots, x_n \in \mathbb{R}\}$$

If $u \in \mathbb{R}^n$, we can write $u = (x_1, \ldots, x_n)$ for some $x_1, \ldots, x_n \in \mathbb{R}$. The elements of $\mathbb{R}^n$ are called *vectors*.

**Example 1**

$(1, 1, 2) \in \mathbb{R}^3 \quad$ *(*$n = 3$*)*

$(2, 4, 5, 8) \in \mathbb{R}^4 \quad$ *(*$n = 4$*)*

**Definition 4 (Equality of vectors)** *Let* $u = (x_1, x_2, \ldots, x_n)$*,* $v = (y_1, y_2, \ldots, y_n)$ *be two vectors of* $\mathbb{R}^n$*. Then,*

$$u = v \Leftrightarrow x_i = y_i\,, \quad \forall i = 1, \ldots, n$$

**Example 2** *In* $\mathbb{R}^3$ *we have*

$(1, 1, 2) \neq (1, 2, 1)$

$(1, 2, 3) \neq (1, 3, 2)$

**Definition 5 (Sum of vectors and product of a vector by a scalar)**
*Let* $u = (x_1, x_2, \ldots, x_n)$*,* $v = (y_1, y_2, \ldots, y_n)$ *be two vectors of* $\mathbb{R}^n$*, and* $\lambda \in \mathbb{R}$*. Then we define the* sum *of* $u$ *and* $v$ *as follows*

$$\begin{aligned} u + v &= (x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) \\ &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \end{aligned}$$

*and the* product *of* $u$ by the scalar $\lambda$ *as follows*

$$\begin{aligned} \lambda u &= \lambda(x_1, x_2, \ldots, x_n) \\ &= (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \end{aligned}$$

Note: If $\lambda \in \mathbb{R}$, $u \in \mathbb{R}^n$ and $\lambda u = \overline{0}$ (the null vector $\overline{0} = \underbrace{(0, \ldots, 0)}_{n}$) then $\lambda = 0$ or $u = \overline{0}$.

**Example 3** *In* $\mathbb{R}^3$ *we have*

$(1, 1, 2) + (1, 2, 3) = (2, 3, 5)$

$4(1, 2, 3) = (4, 8, 12)$

**Definition 6 (Linear combinations)**
*Let* $u_1, \ldots, u_r \in \mathbb{R}^n$*. Then a* linear combination *of these vectors is of the form*

$$\lambda_1 u_1 + \cdots + \lambda_r u_r$$

*with* $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$*.*

**Definition 7 (Spanning sets or linear hulls and generating systems)**
*Let* $u_1, \ldots, u_r \in \mathbb{R}^n$*. Then the* set spanned *by these vectors or the* linear hull *of these vectors is the set of linear combinations of* $u_1, \ldots, u_r \in \mathbb{R}^n$*. The set* $\{u_1, \ldots, u_r\}$ *is called* generating system *of its linear hull.*

**Definition 8 (Linear dependence and linear independence of vectors)**
*Let* $u_1, u_2, \ldots, u_r \in \mathbb{R}^n$*.*

- *These* $r$ *vectors are called* linearly dependent *if it can be written the* null linear combination

  $$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_r u_r = \overline{0}$$

  *with some* $\lambda_i \neq 0$*, where* $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$*. (Note: The null vector* $\overline{0} = \underbrace{(0, \ldots, 0)}_{n}$ *is a linearly dependent vector).*

- *The mentioned* $r$ *vectors are called* linearly independent *if it is verified*

  $$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_r u_r = \overline{0} \Rightarrow \lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$$

Note: If $u_1, \ldots, u_r \in \mathbb{R}^n$ are linearly dependent, then at least one of them can be written as a linear combination of the other ones (relation of dependency).

**Definition 9 (Basis of** $\mathbb{R}^n$**)** *Let* $u_1, u_2, \ldots, u_n \in \mathbb{R}^n$ *be linearly independent. Then the ordered set* $\mathfrak{B} = \{u_1, u_2, \ldots, u_n\}$ *is called a* basis *of* $\mathbb{R}^n$*.*

**Example 4** *In* $\mathbb{R}^2$*,* $\mathfrak{B} = \{(1,0), (0,1)\}$ *is the* canonical basis*. In* $\mathbb{R}^3$*, the basis* $\mathfrak{B} = \{(1,0,0), (0,1,0), (0,0,1)\}$ *is the corresponding* canonical basis.

**Definition 10 (Coordinates with respect to a basis of** $\mathbb{R}^n$**)**
*Let* $\mathfrak{B} = \{u_1, u_2, \ldots, u_n\}$ *be a basis of* $\mathbb{R}^n$*. Let* $v \in \mathbb{R}^n$*. Then* $v$ *can be written as a linear combination of the vectors of the basis* $\mathfrak{B}$*:*

$$v = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n$$

*with unique* $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$*. These coefficients are called* coordinates *of* $v$ *with respect to the basis* $\mathfrak{B}$*.*

# Diagonalization of square matrices

For more details of what follows in this Section, see the *Capítulo 5* of Jarne et al. (1997).

## Eigenvalues and eigenvectors of a square matrix

**Definition 11** *Let $A \in M_n$. Then:*

*a)* $\lambda \in \mathbb{R}$ *is called* eigenvalue of $A$ *if there exists* $v \in \mathbb{R}^n$, $v \neq 0$ *such that* $Av = \lambda v$ *(the vector $v$ is written as a* column *matrix in this last expression).*

*b)* $v \in \mathbb{R}^n$, $v \neq 0$ *is called* eigenvector of $A$ *if there exists* $\lambda \in \mathbb{R}$ *such that* $Av = \lambda v$ *(the vector $v$ is written as a* column *matrix in this last expression).*

*In both cases, we will say that $v$ is an eigenvector of $A$ associated to the eigenvalue $\lambda$ of $A$.*

**Proposition 4**

- *It corresponds to each eigenvector $v$ of $A \in M_n$ a unique associated eigenvalue $\lambda$.*
- *It corresponds to each eigenvalue $\lambda$ of $A \in M_n$ an infinite number of associated eigenvectors.*

Note that if $A \in M_n$ and there exist $v \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ such that $Av = \lambda v$, then

$$Av = \lambda v \Leftrightarrow Av = \lambda I_n v \Leftrightarrow Av - \lambda I_n v = 0 \Leftrightarrow (A - \lambda I_n)v = 0$$

This last obtained linear system is homogeneous, so it has solutions $v \neq 0$ when the rank of the corresponding coefficient matrix $A - \lambda I_n$ is less than the number of unknowns $n$. That will happen if and only if $|A - \lambda I_n| = 0$. Note that this last fact is equivalent to $\mathrm{rk}(A - \lambda I_n) < n$ and then $n - \mathrm{rk}(A - \lambda I_n) \geq 1$.

## Characteristic polynomial of a square matrix

**Definition 12** *Let $A \in M_n$.*

- *We will call* characteristic polynomial of $A$ *to the determinant* $|A - \lambda I_n|$, *which is a polynomial in the variable $\lambda$ of degree $n$, with real coefficients.*
- *We will call* characteristic equation of $A$ *to the equation* $|A - \lambda I_n| = 0$.

**Proposition 5** *Let $A \in M_n$. Then $\lambda$ is an eigenvalue of $A \Leftrightarrow \lambda$ is a real root of the characteristic polynomial of $A$.*

In short:

- Eigenvalues of $A$: they are the real roots of $|A - \lambda I_n| = 0$.
- Eigenvectors of $A$ associated to an eigenvalue $\lambda$: they are the non-trivial (non-vanishing) solutions $X$ to the linear system $(A - \lambda I_n)X = 0$.

**Definition 13** *Let $A \in M_n$, and let $\lambda$ be an eigenvalue of $A$. We will call* multiplicity order of $\lambda$ *to the multiplicity of $\lambda$ as a root of $|A - \lambda I_n| = 0$.*

**Proposition 6** *Let $A \in M_n$, and let $\lambda$ be an eigenvalue of $A$. Then the number of* linearly independent *eigenvectors associated to $\lambda$ is*

$$n - \mathrm{rk}(A - \lambda I_n)$$

*Note that this number is the number of degrees of freedom of the indetermined consistent linear system $(A - \lambda I_n)X = 0$, and that this number is greater than or equal to $1$.*

**Proposition 7** *Let $A \in M_n$, and $\lambda$ eigenvalue of $A$. Then*

$$1 \leq n - \mathrm{rk}(A - \lambda I_n) \leq \text{multiplicity order of } \lambda$$

**Proposition 8** *Let $A \in M_n$. The eigenvectors associated to different eigenvalues are linearly independent.*

## Diagonalizable matrices

**Definition 14** *Let $A, B \in M_n$. Then $A$ and $B$ are called* similar $\Leftrightarrow$ *there exists $P \in M_n$ with $|P| \neq 0$ such that $P^{-1}AP = B$.*

**Definition 15** *Let $A \in M_n$. We will say that $A$* is diagonalizable *if there exists $P \in M_n$ with $|P| \neq 0$ such that $P^{-1}AP = D$, where $D \in M_n$ is a diagonal matrix.*

*In other words, $A$ is diagonalizable if it is similar to a diagonal matrix.*

**Proposition 9** *Let $A \in M_n$. Then*

$$A \text{ is diagonalizable} \Leftrightarrow \text{there exists a basis of } \mathbb{R}^n \text{ made up with eigenvectors of } A$$

**Proposition 10** *Let $A \in M_n$. Let $\lambda_1, \lambda_2, \ldots, \lambda_r$ be all of the different roots of $|A - \lambda I_n| = 0$. Then $A$ is diagonalizable if and only if they are* simultaneously *satisfied the three following conditions:*

*a)* $\lambda_i \in \mathbb{R}, \forall i = 1, \ldots, r$.

*b) The multiplicity order of $\lambda_i$ is equal to $n - \mathrm{rk}(A - \lambda_i I_n), \forall i = 1, \ldots, r$.*

*c) The sum of the multiplicity orders of all of the $\lambda_i$ is equal to the order of the matrix $n$.*

If $A$ is diagonalizable, with a basis of $\mathbb{R}^n$ made up with eigenvectors of $A$ given by $\{v_1, v_2, \ldots, v_n\}$, we have that $P^{-1}AP = D$ with:

$P$ is a matrix with the components of the eigenvectors written as the columns of the former.

$D$ is a diagonal matrix with the corresponding eigenvalues written in the main diagonal.

**Proposition 11** *Let $A \in M_n$. Then*

$$A \text{ is symmetric} \Rightarrow A \text{ is diagonalizable}$$

Note that the converse of this proposition is not true in general.

## Calculation of the powers of diagonalizable matrices

Let $A \in M_n$ be diagonalizable. Then there exists $P \in M_n$ with $|P| \neq 0$ such that $P^{-1}AP = D$, where $D$ is a diagonal matrix. Therefore,

$$\begin{aligned}
&A = PDP^{-1} \\
&A^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1} \\
&A^3 = PDP^{-1}PD^2P^{-1} = PD^3P^{-1}
\end{aligned}$$

and in general, it can be proved that

$$A^k = PD^kP^{-1}, k = 1, 2, 3, \ldots$$

This expression relates the natural powers of the matrix $A$ (in general, difficult to compute) with the natural powers of $D$ (very easy to compute since $D$ is diagonal).

# Real quadratic forms

For more details of what follows in this Chapter, see the *Capítulo 6* of Jarne et al. (1997).

**Definition 16** *A* real quadratic form $Q$ *is given by*

$$\underset{1\times 1}{Q(x_1,\ldots,x_n)} = \underset{1\times n}{X^t}\,\underset{n\times n}{A}\,\underset{n\times 1}{X}$$

*where*

$$X = \underset{n\times 1}{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}$$

*with $A \in M_n$ being a real symmetric matrix. The matrix $A$ is termed as* associated *to $Q$.*

**Example 5** *In the case of two variables, $n = 2$, we have as an example*

$$\begin{aligned} Q(x,y) &= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= x^2 + 3y^2 + 2xy \end{aligned}$$

*The first expression serves as an example of* matrix form of $Q$*, and the second as an example of* polynomial form of $Q$.



## Change of variables

If we perform the change

$$\underset{n\times 1}{X} = \underset{n\times n}{P}\,\underset{n\times 1}{\overline{X}}$$

where $P$ is regular ($|P| \neq 0$), we have

$$Q(X) = X^t A X = (P\overline{X})^t A P\overline{X} = \overline{X}^t P^t A P \overline{X} = Q(\overline{X})$$

so the associated matrix to the quadratic form $Q$ *in the variables of* $\overline{X}$ is $P^tAP$. Note that

$$(P^tAP)^t = P^tA^t(P^t)^t = P^tAP$$

where it has been used that $A^t = A$, because $A$ is symmetric, and that $(P^t)^t = P$, which holds for any matrix $P$. Therefore, $P^tAP$ is also symmetric.

**Definition 17** *Let $A, B \in M_n$, $A, B$ symmetric. $A, B$ are called* congruent *if they are associated matrices to the same quadratic form, namely, if there exists $P \in M_n$ with $|P| \neq 0$, such that $B = P^tAP$.*

**Definition 18** *If one associated matrix to one quadratic form $Q$ is diagonal,*

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

*it is said that*

$$Q(x_1, \dots, x_n) = X^tDX = d_1x_1^2 + d_2x_2^2 + \cdots + d_nx_n^2$$

*is a* diagonal expression *of $Q$.*

## Classification of a real quadratic form according to its sign

**Definition 19** *Let $Q$ be a real quadratic form given by $Q(\overline{x}) = X^tAX$ where $X$ are the coordinates of the vector $\overline{x}$ with respect to a basis $\{\overline{a}_1, \overline{a}_2, \dots, \overline{a}_n\}$ of $\mathbb{R}^n$. Then we will say that:*

1. *$Q$ is* positive definite *(PD) if $Q(\overline{x}) > 0\ \forall \overline{x} \in \mathbb{R}^n$ such that $\overline{x} \neq \overline{0}$.*
2. *$Q$ is* negative definite *(ND) if $Q(\overline{x}) < 0\ \forall \overline{x} \in \mathbb{R}^n$ such that $\overline{x} \neq \overline{0}$.*
3. *$Q$ is* positive semidefinite *(PSD) if $Q(\overline{x}) \geq 0\ \forall \overline{x} \in \mathbb{R}^n$ and there exists $\overline{x} \neq \overline{0}$ such that $Q(\overline{x}) = 0$.*
4. *$Q$ is* negative semidefinite *(NSD) if $Q(\overline{x}) \leq 0\ \forall \overline{x} \in \mathbb{R}^n$ and there exists $\overline{x} \neq \overline{0}$ such that $Q(\overline{x}) = 0$.*
5. *$Q$ is* indefinite *(I) if $Q(\overline{x}) > 0$ for certain $\overline{x} \in \mathbb{R}^n$ and $Q(\overline{x}^*) < 0$ for another $\overline{x}^* \in \mathbb{R}^n$.*

In general, the direct study of the sign of an arbitrary real quadratic form is not immediate.

However, if we have a diagonal expression of $Q$ it is very easy to determine its sign, which only depends on the sign of the coefficients $d_1, d_2, \ldots, d_n$:

**Proposition 12** *Let $Q(\overline{x}) = X^t D X = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2$. It holds:*

1. *$Q$ is PD $\Leftrightarrow$ $d_i > 0$, $i = 1, \ldots, n$.*

2. *$Q$ is ND $\Leftrightarrow$ $d_i < 0$, $i = 1, \ldots, n$.*

3. *$Q$ is PSD $\Leftrightarrow$ $d_i \geq 0$, $i = 1, \ldots, n$ and certain $d_j = 0$.*

4. *$Q$ is NSD $\Leftrightarrow$ $d_i \leq 0$, $i = 1, \ldots, n$ and certain $d_j = 0$.*

5. *$Q$ is I $\Leftrightarrow$ certain $d_i < 0$ and certain $d_j > 0$.*

A given real quadratic form admits different diagonal expressions, but all of them share a common feature: they have the same number of positive coefficients and the same number of negative coefficients (*Sylvester's Law of inertia*), so that we can classify a real quadratic form through any of its diagonal expressions.

## Study of the sign of a real quadratic form

In this subsection we will study three forms of finding the sign of a real quadratic form $Q$:

1. Finding a diagonal expression of $Q$:

   a) Method of the eigenvalues.

   b) Method of Lagrange or reducing to a sum of squares.

2. Without finding a diagonal expression of $Q$:

   c) Method of the main minors.

**a) Method of eigenvalues**. Example to be developed in the classroom sessions.

**b) Method of Lagrange or reducing to a sum of squares**. Examples to be developed in the classroom sessions.

**c) Method of the main minors**. It is based on the following:

**Proposition 13** *Let $Q(\overline{x}) = X^t AX$ be a real quadratic form, and let*

$$|A_1|, |A_2|, \ldots, |A_n|$$

*be the main minors of $A$. It holds:*

1. *$Q$ is PD $\Leftrightarrow |A_i| > 0, \forall i = 1, \ldots, n$.*
2. *$Q$ is ND $\Leftrightarrow (-1)^i |A_i| > 0, \forall i = 1, \ldots, n$.*
3. *$|A_i| > 0, \forall i = 1, \ldots, n-1$ and $|A_n| = 0 \Rightarrow Q$ is PSD.*
4. *$(-1)^i |A_i| > 0, \forall i = 1, \ldots, n-1$ and $|A_n| = 0 \Rightarrow Q$ is NSD.*
5. *$|A_n| \neq 0$ and it does not hold 1. neither 2. $\Rightarrow Q$ is I.*
6. *$|A_n| = 0$, $|A_i| \neq 0$, $i = 1, \ldots, n-1$ and it does not hold 3. neither 4. $\Rightarrow Q$ is I.*

## Constrained real quadratic forms

Let $Q(X) = \underset{1\times n}{X^t} \underset{n\times n}{A} \underset{n\times 1}{X}$ be constrained to (the solutions of) $\underset{m\times n}{B} \underset{n\times 1}{X} = 0$.

If $\mathrm{rk}(B) = r$, from $BX = 0$ we can isolate $r$ main variables in terms of $n - r$ degrees of freedom.

These main variables are substituted in $Q(X) = X^t AX$. The result is a real quadratic form *without constraints* with $n - r$ free variables (degrees of freedom).

The sign of this real cuadratic form without constraints is the same of that of $Q(X)$ constrained to $BX = 0$.

# Some topology notions in $\mathbb{R}^n$

**Definition 20 (Scalar product in $\mathbb{R}^n$)**
*Let $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ be two vectors of $\mathbb{R}^n$. Then the* scalar product *of $x$ and $y$ is the real number given by*

$$\langle x, y\rangle = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

**Proposition 14**

- $\langle x, y\rangle = \langle y, x\rangle, \forall x, y \in \mathbb{R}^n$.
- $\langle x, x\rangle = x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0, \forall x \in \mathbb{R}^n$.
- $\langle x, x\rangle = 0 \Leftrightarrow x = (0, \ldots, 0)$.
- $\langle tx, sy\rangle = ts\langle x, y\rangle \quad \forall t, s \in \mathbb{R}, \forall x, y \in \mathbb{R}^n$.

**Definition 21 (Norm in $\mathbb{R}^n$)**
*Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. We define the* norm *of $x$ as*

$$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

**Proposition 15**

- $\|x\| > 0 \quad \forall x \neq 0$ *and* $\|x\| = 0 \Leftrightarrow x = 0$.
- $\|tx\| = |t|\|x\| \quad \forall t \in \mathbb{R}, \forall x \in \mathbb{R}^n$.
- $\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n$ *(triangular inequality).*

**Definition 22** *Let $x \in \mathbb{R}^n$, $x \neq 0$, it is termed as* normalized or unitary *if $\|x\| = 1$.*

**Proposition 16** *Given a vector $x \neq 0$ of $\mathbb{R}^n$, the vector $y = \dfrac{x}{\|x\|}$ is normalized.*

**Definition 23** *Let $x, y \in \mathbb{R}^n$, with $x, y \neq 0$. These two vectors are called* orthogonal *if $\langle x, y \rangle = 0$. If moreover it holds $\|x\| = 1$, $\|y\| = 1$ they are called* orthonormal.

**Definition 24** *Let $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ be two vectors of $\mathbb{R}^n$. We define the* distance *between $x$ and $y$ as*

$$d(x, y) = \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$$

**Proposition 17** *Let $x, y, z \in \mathbb{R}^n$. It holds:*

- $d(x, y) \geq 0$, $d(x, y) = 0 \Leftrightarrow x = y$.
- $d(x, y) = d(y, x)$.
- $d(x, y) \leq d(x, z) + d(z, y)$.

## Particular case of $n = 1$

**Definition 25**

- *The "scalar product" of $x, y \in \mathbb{R}$ is the ordinary product $xy$.*
- *The "norm" of $x \in \mathbb{R}$ is the* absolute value *of $x$, namely*
$$|x| = \sqrt{xx} = \sqrt{x^2}.$$

**Proposition 18 (Properties of the absolute value)**

- $|x| > 0 \quad \forall x \neq 0$ *and* $|x| = 0 \Leftrightarrow x = 0$.
- $|tx| = |t||x| \quad \forall t, x \in \mathbb{R}$.
- $|x + y| \leq |x| + |y| \quad \forall x, y \in \mathbb{R}$ *(triangular inequality).*
- *It also holds*
$$|x| = \begin{cases} x & \text{if} \quad x \geq 0 \\ -x & \text{if} \quad x < 0 \end{cases}$$

**Definition 26 (Distance in $\mathbb{R}$)** *Let $x, y \in \mathbb{R}$. The* distance *between $x$ and $y$ is*

$$d(x, y) = |x - y|$$

**Proposition 19** *Let $x, y, z \in \mathbb{R}$. It holds:*

- $|x - y| \geq 0$, $|x - y| = 0 \Leftrightarrow x = y$.
- $|x - y| = |y - x|$.
- $|x - y| \leq |x - z| + |z - y|$.

## Some metric topology notions in $\mathbb{R}^n$

**Definition 27** *Let* $x_0 \in \mathbb{R}^n$*, and* $r > 0$*. It is called* open ball of center $x_0$ and radius $r$ *the set of elements of* $\mathbb{R}^n$ *whose distance to* $x_0$ *is less than* $r$*,*

$$B(x_0, r) = \{x \in \mathbb{R}^n \,|\, d(x, x_0) < r\} = \{x \in \mathbb{R}^n \,|\, \|x - x_0\| < r\}$$

**Definition 28** *Let* $x_0 \in \mathbb{R}^n$*, and* $r > 0$*. It is called* reduced open ball of center $x_0$ and radius $r$ *the set of elements of* $\mathbb{R}^n$ *whose distance to* $x_0$ *is less than* $r$ *and more than 0,*

$$B^*(x_0, r) = \{x \in \mathbb{R}^n \,|\, 0 < d(x, x_0) < r\} = \{x \in \mathbb{R}^n \,|\, 0 < \|x - x_0\| < r\}$$

**Definition 29** *Let* $S$ *be a subset of* $\mathbb{R}^n$*. It is said that* $x_0 \in S$ *is an* interior point *of* $S$ *if there exists* $r > 0$ *such that* $B(x_0, r) \subset S$.

**Definition 30** *Let* $S$ *be a subset of* $\mathbb{R}^n$*. It is said that* $x_0 \in S$ *is an* accumulation point *of* $S$ *if for all* $r > 0$*,* $B^*(x_0, r) \bigcap S \neq \emptyset$.

**Definition 31** *It is said that* $S \subset \mathbb{R}^n$ *is an* open set *if all of its points are interior.*

**Definition 32** *It is said that* $S \subset \mathbb{R}^n$ *is a* closed set *if its complementary set* $\mathbb{R}^n \setminus S$ *is open*.

**Definition 33** *It is said that* $S \subset \mathbb{R}^n$ *is a* bounded set *if there exist* $x_0 \in \mathbb{R}^n$ *and* $r > 0$ *such that* $S \subset B(x_0, r)$.

## Intervals and neighbourhoods on $\mathbb{R}$

This Section is taken from *Sección 7.1* of Jarne et al. (1997).

**Definition 34 (Open intervals)** *Let* $a, b \in \mathbb{R}$.

- *We will call* open intervals *to the following subsets of* $\mathbb{R}$*:*

$$\begin{aligned}(a, b) &= \{x \in \mathbb{R} \,|\, a < x < b\}\\ (-\infty, b) &= \{x \in \mathbb{R} \,|\, x < b\}\\ (a, +\infty) &= \{x \in \mathbb{R} \,|\, a < x\}\end{aligned}$$

- *We will call* closed intervals *to the following subsets of* $\mathbb{R}$*:*

$$\begin{aligned}[a, b] &= \{x \in \mathbb{R} \,|\, a \leq x \leq b\}\\ (-\infty, b] &= \{x \in \mathbb{R} \,|\, x \leq b\}\\ [a, +\infty) &= \{x \in \mathbb{R} \,|\, a \leq x\}\end{aligned}$$

- *We will call* half-closed or half-open *to the following subsets of* $\mathbb{R}$*:*

$$(a, b] = \{x \in \mathbb{R} \,|\, a < x \leq b\}$$
$$[a, b) = \{x \in \mathbb{R} \,|\, a \leq x < b\}$$

**Definition 35 (Neighbourhoods on $\mathbb{R}$)**

- *We will call* symmetric neighbourhood *of center* $x_0 \in \mathbb{R}$ *and radius* $\delta > 0$ *to the open interval*

$$E(x_0, \delta) = (x_0 - \delta, x_0 + \delta) = \{x \in \mathbb{R} \,|\, |x - x_0| < \delta\}$$

- *We will call* reduced symmetric neighbourhood *of center* $x_0 \in \mathbb{R}$ *and radius* $\delta > 0$ *to the set*

$$E^*(x_0, \delta) = E(x_0, \delta) \setminus \{x_0\}$$

- *We will consider as neighbourhoods of* $+\infty$ *the open intervals of the form* $(a, +\infty)$ *and as neighbourhoods of* $-\infty$ *those of the form* $(-\infty, b)$.

**Definition 36** *Given a point* $x_0 \in \mathbb{R}$ *and a subset* $A$ *of* $\mathbb{R}$ *we will say that:*

- $x_0$ *is an* interior point *of* $A$ *if there exists* $\delta > 0$ *such that* $E(x_0, \delta) \subset A$.
- $x_0$ *is a* frontier point *of* $A$ *if* $\forall \delta > 0$ *we have that* $E(x_0, \delta) \bigcap A \neq \emptyset$ *and* $E(x_0, \delta) \bigcap (\mathbb{R} \setminus A) \neq \emptyset$.
- $x_0$ *is an* accumulation point *of* $A$ *if* $\forall \delta > 0$ *we have that* $E^*(x_0, \delta) \bigcap A \neq \emptyset$.
- $A$ *is an* open set *of* $\mathbb{R}$ *if all of the points of* $A$ *are interior points of* $A$.

# Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

There are a number of basic concepts that, although contained in the material of the subject Mathematics I, will not be covered here but we will refer to the *Unidad 7* of

`www.unizar.es/aragon_tres`

More specifically, the students should study the following topics therein:

- *Conceptos básicos*.
- *Funciones elementales*.
- *Límite de una función*.
- *Continuidad*.

For more details of what follows in this Chapter, see the *Capítulos 7 y 8* of Jarne et al. (1997).

**Definition 37** *We will call* function *from* $\mathbb{R}^n$ *to* $\mathbb{R}^m$ *to any mapping* $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$. *The* domain *of* $f$ *is*

$$D = \{\overline{x} \in \mathbb{R}^n \,|\, \text{there exists } f(\overline{x})\}$$

*The* image *of* $f$ *is*

$$\operatorname{Im} f = \{f(\overline{x}) \,|\, \overline{x} \in D\}$$

- *If* $m = 1$ *the function is called* scalar *or* real-valued.
- *If* $m > 1$ *the function is called* vectorial *or* vector-valued.
- *If* $n = 1$ *the function is called of* real variable.
- *If* $n > 1$ *the function is called of* vectorial variable.

*If* $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, *with* $m > 1$, $f$ *can be decomposed in* $m$ *scalar functions* $f_1, f_2, \ldots, f_m$ *that are called* components *of* $f$*:*

$$f(\overline{x}) = (f_1(\overline{x}), f_2(\overline{x}), \ldots, f_m(\overline{x})) , \quad \forall \overline{x} = (x_1, \ldots, x_n) \in D$$

**Example 6**

$$n = 1, m = 1 \quad f(x) = \mathrm{e}^x, \quad D = \mathbb{R}$$

$$n = 2, m = 1 \quad f(x, y) = \sin(xy), \quad D = \mathbb{R}^2$$

$$n = 1, m = 2 \quad f(x) = (f_1(x), f_2(x)) = (\cos x, \ln x), \quad D = (0, +\infty)$$

*and also, for example, if* $n = 3, m = 2$ *we can define*

$$\begin{aligned} f(\overline{x}) &= f(x, y, z) = (f_1(x, y, z), f_2(x, y, z)) \\ &= (\mathrm{e}^{x+y+z}, 2x^2 + y + \sqrt{z}) \end{aligned}$$

*with domain* $D = \underset{x}{\mathbb{R}} \times \underset{y}{\mathbb{R}} \times \underset{z}{[0, +\infty)}$.

**Definition 38** *Given* $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ *and* $c \in \mathbb{R}$, *we will call* level set *of value* $c$ *to the set*

$$S_c = \{\overline{x} \in D \mid f(\overline{x}) = c\}$$

*For* $n = 2$ *the level sets are* level curves *or* contour curves *of the plane* $\mathbb{R}^2$.

Note: The level sets corresponding to *different* values do never intersect. The level sets may be empty sets.

## Limits and continuity

This topic will be explained in the classroom sessions.

## Derivability

**Definition 39** *Given* $f : D \to \mathbb{R}$ *and* $x_0$ *an interior point of* $D \subset \mathbb{R}$, *we will say that* $f$ *is* derivable *at* $x_0$ *if there exists and it is a real number the limit*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

*If it does exist, the limit is denoted as* $f'(x_0) = \dfrac{df(x_0)}{dx}$, *and it is called* derivative *of* $f$ *at* $x_0$.

If it exists, $f'(x_0)$ is the slope of the straight line tangent to the graph of $f$ at $(x_0, f(x_0))$, which has the analytic expression

$$y - f(x_0) = f'(x_0)(x - x_0)$$

**Proposition 20** *If $f$ is a derivable function at $x_0 \Rightarrow f$ is continuous at $x_0$.*

Note that the converse is not true in general. For example, it may be considered the function

$$f(x) = |x|$$

at $x_0 = 0$. We have that this function $f$ is continuous at $x_0 = 0$ but it is not derivable at the same point.

**Definition 40** *Let $f : D \to \mathbb{R}$. We will call* derivative function[2] *of $f$ to the function $f' : D_1 \to \mathbb{R}$ that associates to each real number $x \in D_1$ the value of the derivative of $f$ at such a point. $D_1$ is defined as $D_1 = \{x \in D \,|\, \text{there exists } f'(x)\}$.*

**Table of elementary derivatives**

| $f(x)$ | $f'(x)$ |
|---|---|
| $k$ | $0$ |
| $x^n$ | $n\,x^{n-1}$ |
| $\ln x$ | $\dfrac{1}{x}$ |
| $\log_a x, a > 0$ | $\dfrac{1}{x\ \ln a}$ |
| $\mathrm{e}^x$ | $\mathrm{e}^x$ |
| $a^x, a > 0$ | $a^x\ \ln a$ |
| $\sin x$ | $\cos x$ |
| $\cos x$ | $-\sin x$ |
| $\tan x$ | $1 + \tan^2 x = \dfrac{1}{\cos^2 x}$ |
| $\arcsin x$ | $\dfrac{1}{\sqrt{1 - x^2}}$ |
| $\arccos x$ | $-\dfrac{1}{\sqrt{1 - x^2}}$ |
| $\arctan x$ | $\dfrac{1}{1 + x^2}$ |

**Definition 41** *We will say that a function $f$ is derivable on an open interval $(a, b)$ if it is derivable for all $x \in (a, b)$.*

**Proposition 21 (Properties of derivable functions)** *Let $f, g$ be derivable functions on $(a, b)$. Then, $\forall x \in (a, b)$ we have:*

[2]Or simply derivative if it is clear from the context what we are referring to.

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

2. $(\lambda f)'(x) = \lambda f'(x)$, $\lambda \in \mathbb{R}$.

3. $(f\,g)'(x) = f'(x)g(x) + f(x)g'(x)$ *(Leibnitz's Rule)*.

4. $\left(\dfrac{f}{g}\right)'(x) = \dfrac{f'(x)g(x) - f(x)g'(x)}{g^2(x)} = \dfrac{f'(x)}{g(x)} - \dfrac{f(x)g'(x)}{g^2(x)}$, *if* $g(x) \neq 0$.

**Proposition 22 (Derivative of the composition of functions)** *If $f$ is derivable at $x_0$ and $g$ is derivable at $f(x_0)$, then the composition $g \circ f$ is derivable at $x_0$ and its derivative is*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$$

*This relation is known as the* Chain Rule.

**Proposition 23 (Derivative of the inverse function)** *If $f$ is derivable at $x_0$, $f'(x_0) \neq 0$ and $f$ has inverse function, then $f^{-1}$ is derivable at $f(x_0)$, with*

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}$$

**Proposition 24 (L'Hôpital's Rule)** *Let $f, g$ be continuous and derivable functions in a reduced neighbourhood of $a$, $(a-\delta, a+\delta) \setminus \{a\}$, with $\delta > 0$, such that there exists*

$$\lim_{x \to a} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{-\infty, \infty\},$$

*$g'(x) \neq 0$ for all $x$ in the mentioned reduced neighbourhood of $a$ and it is verified i) or ii):*

*i)* $\lim_{x\to a} f(x) = \lim_{x \to a} g(x) = 0$.

*ii)* $\lim_{x \to a} |g(x)| = +\infty$.

*Then, it holds*

$$\text{there exists } \lim_{x \to a} \frac{f(x)}{g(x)} = L$$

**Example 7 (Partial derivatives on a point)** *Let the function of two variables*

$$f(x,y) = \begin{cases} \dfrac{x^3 + 3y^3}{x^2 + y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$$

*Then the partial derivatives of $f$ with respect to $x$ and $y$ at the point $(0,0)$ are computed as follows:*

$$\begin{aligned}
D_x f(0,0) &= f_x(0,0) = \frac{\partial f(0,0)}{\partial x} = \lim_{\lambda\to 0}\frac{f((0,0)+\lambda(1,0))-f(0,0)}{\lambda}\\
&= \lim_{\lambda\to 0}\frac{f(\lambda,0)-f(0,0)}{\lambda} = \lim_{\lambda\to 0}\frac{1}{\lambda}\frac{\lambda^3}{\lambda^2} = 1\\
D_y f(0,0) &= f_y(0,0) = \frac{\partial f(0,0)}{\partial y} = \lim_{\lambda\to 0}\frac{f((0,0)+\lambda(0,1))-f(0,0)}{\lambda}\\
&= \lim_{\lambda\to 0}\frac{f(0,\lambda)-f(0,0)}{\lambda} = \lim_{\lambda\to 0}\frac{1}{\lambda}3\frac{\lambda^3}{\lambda^2} = 3
\end{aligned}$$

*The case of functions of $n$ variables and $m$ components is treated analogously.*

**Definition 42** *If there exist the $n$ partial derivatives of a scalar function $f$ at $\overline{x}_0$, we will call* gradient vector *of $f$ at $\overline{x}_0$ to the vector*

$$\nabla f(\overline{x}_0) = \left(\frac{\partial f(\overline{x}_0)}{\partial x_1}, \frac{\partial f(\overline{x}_0)}{\partial x_2}, \ldots, \frac{\partial f(\overline{x}_0)}{\partial x_n},\right)$$

**Example 8** *Let the previous function of two variables*

$$f(x,y) = \begin{cases} \dfrac{x^3+3y^3}{x^2+y^2} & \text{if} \quad (x,y)\neq(0,0)\\ 0 & \text{if} \quad (x,y)=(0,0)\end{cases}$$

*Then*

$$\nabla f(0,0) = \left(\frac{\partial f(0,0)}{\partial x}, \frac{\partial f(0,0)}{\partial y}\right) = (1,3)$$

**Definition 43** *If there exist the $n$ partial derivatives of a vector-valued function $f=(f_1,f_2,\ldots,f_m)$ at $\overline{x}_0$, we will call* Jacobian matrix *of $f$ at $\overline{x}_0$ to the following matrix of order $m\times n$:*

$$Jf(\overline{x}_0) = \begin{pmatrix} \dfrac{\partial f_1(\overline{x}_0)}{\partial x_1} & \cdots & \dfrac{\partial f_1(\overline{x}_0)}{\partial x_n}\\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m(\overline{x}_0)}{\partial x_1} & \cdots & \dfrac{\partial f_m(\overline{x}_0)}{\partial x_n}\end{pmatrix} = \begin{pmatrix}\nabla f_1(\overline{x}_0)\\ \vdots \\ \nabla f_m(\overline{x}_0)\end{pmatrix}$$

*If $n=m$ we will call* Jacobian *of $f$ at $\overline{x}_0$ to the determinant $|Jf(\overline{x}_0)|$.*

**Definition 44** *If a scalar or vector-valued function admits $\dfrac{\partial f(\overline{x})}{\partial x_i}$, $\forall \overline{x}\in D$, then it is naturally defined the* partial derivative function *of $f$ with respect to $x_i$, $\forall\overline{x}\in D$.*

**Definition 45** *If $f$ is a scalar function which admits all of its partial derivatives on a domain $D \subset \mathbb{R}^n$, then there exists the* gradient function of $f$*, given by*

$$\nabla f(\overline{x}) = \left( \frac{\partial f(\overline{x})}{\partial x_1}, \frac{\partial f(\overline{x})}{\partial x_2}, \ldots, \frac{\partial f(\overline{x})}{\partial x_n}, \right), \forall \overline{x} \in D$$

*For vector-valued functions and the* Jacobian matrix function *it is proceeded analogously.*

With these conditions, **when taking a partial derivative with respect to one variable, the other variables are kept fixed (constant)**:

**Example 9** *Let $f(x, y, z) = x^3y^2 + x^2z + xyz^4$. Then,*

$$\begin{aligned} \frac{\partial f(x,y,z)}{\partial x} &= 3x^2y^2 + 2xz + yz^4 \\ \frac{\partial f(x,y,z)}{\partial y} &= 2x^3y + xz^4 \\ \frac{\partial f(x,y,z)}{\partial z} &= x^2 + 4xyz^3 \end{aligned}$$

**Example 10** *Consider now a vector-valued function $f : \mathbb{R}^3 \to \mathbb{R}^2$, given by*

$$f(\overline{x}) = (f_1(x, y, z), f_2(x, y, z)) = (2xy - z, xz)$$

*Then*

$$\begin{aligned} Jf(x,y,z) &= \begin{pmatrix} \frac{\partial f_1(x,y,z)}{\partial x} & \frac{\partial f_1(x,y,z)}{\partial y} & \frac{\partial f_1(x,y,z)}{\partial z} \\ \frac{\partial f_2(x,y,z)}{\partial x} & \frac{\partial f_2(x,y,z)}{\partial y} & \frac{\partial f_2(x,y,z)}{\partial z} \end{pmatrix} \\ &= \begin{pmatrix} 2y & 2x & -1 \\ z & 0 & x \end{pmatrix} \end{aligned}$$

*If we want to evaluate it at a specific point, for example $(1, 1, 0)$, we have*

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

## Differentiability of a function

From an informal point of view, a scalar function of several variables is *differentiable* at $\overline{x}_0$ if there exists a well-defined tangent "plane" to its graph at $(\overline{x}_0, f(\overline{x}_0))$.

This allows to approximate the values of the function around $\overline{x}_0$ by the values obtained through the tangent "plane". If there exists the mentioned tangent "plane" at $(\overline{x}_0, f(\overline{x}_0))$, its analytic expression is

$$z = f(\overline{x}_0) + \nabla f(\overline{x}_0) \cdot (\overline{x} - \overline{x}_0)$$

Of course, there exists a formal and rigourous definition of differentiability of a function at a point, see for example the *Sección 8.6* of Jarne et al. (1997).

**Proposition 25 (Sufficient condition for differentiability (not necessary))** *Let $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, $D$ open and $\overline{x}_0 \in D$. If there exists $Jf(\overline{x})$ on an open neighbourhood of $\overline{x}_0$ and it is continuous (all the partial derivatives are continuous) at $\overline{x}_0$, then $f$ is differentiable at $\overline{x}_0$.*

- The converse to this Proposition is not true in general.
- If $f$ is differentiable at $\overline{x}_0$, we can take $Jf(\overline{x}_0) \cdot d\overline{x}$ as the *differential* of $f$ at $\overline{x}_0$ (in the variables given).
- For scalar functions, $Jf(\overline{x}), Jf(\overline{x}_0)$ reduce, respectively, to $\nabla f(\overline{x}), \nabla f(\overline{x}_0)$.

**Proposition 26** *If $f : D \to \mathbb{R}^m$, where $D \subset \mathbb{R}^n$ is an open set, is differentiable at $\overline{x}_0 \in D$, then $f$ is continuous at $\overline{x}_0$.*

The converse to this Proposition is not true in general.

**Definition 46 (Directional derivatives for differentiable functions)**
*If $f : D \to \mathbb{R}^m$, with $D \subset \mathbb{R}^n$ open set, is differentiable at $\overline{x}_0 \in D$ and $\overline{v} \in \mathbb{R}^n$ is unitary (i.e., $\|\overline{v}\| = 1$), we define the* directional derivative *of $f$ with respect to $\overline{v}$ at $\overline{x}_0$ as*

$$D_{\overline{v}} f(\overline{x}_0) = Jf(\overline{x}_0) \cdot \overline{v}$$

*where in the right hand side $\cdot$ denotes matrix product.*[3] *If we have a scalar function, namely $m = 1$, we have*

$$D_{\overline{v}} f(\overline{x}_0) = \nabla f(\overline{x}_0) \cdot \overline{v}$$

[3] For functions not necessarily differentiable at $\overline{x}_0$, the directional derivative is defined, if it exists and it is a real number, as the following limit:

$$D_{\overline{v}} f(\overline{x}_0) = \lim_{\lambda \to 0} \frac{f(\overline{x}_0 + \lambda \overline{v}) - f(\overline{x}_0)}{\lambda}$$

For functions $f$ differentiable at $\overline{x}_0$, both concepts coincide and the corresponding expressions provide the same result. However, the first given expression is generally easier to compute. In any case, if they exist, the partial derivatives are directional derivatives with respect to the vectors of the canonical basis of $\mathbb{R}^n$, $n$ being the number of variables of the function.

**Proposition 27** *Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set, $f$ differentiable at $\overline{x}_0 \in D$. Then it holds:*

- *$\nabla f(\overline{x}_0)$ is perpendicular to the level set passing through $\overline{x}_0$ at this point.*
- *The maximum value of the directional derivative of $f$ at $\overline{x}_0$ is attained in the direction of $\nabla f(\overline{x}_0)$.*

**Example 11 (Directional derivatives)**

- *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by $f(x, y, z) = x^2yz$. Let us compute its directional derivative with respect to $\overline{v} = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right)$ at $\overline{x}_0 = (1, 1, 1)$. We find first the partial derivatives of $f$:*

$$\begin{aligned} \frac{\partial f(x,y,z)}{\partial x} &= 2xyz \\ \frac{\partial f(x,y,z)}{\partial y} &= x^2z \\ \frac{\partial f(x,y,z)}{\partial z} &= x^2y \end{aligned}$$

  *We observe that they are all continuous at all points of $\mathbb{R}^3$ and in particular at the point $(1, 1, 1)$, thus $f$ is differentiable at $(1, 1, 1)$. Then,*

$$D_{\overline{v}}f(\overline{x}_0) = \nabla f(\overline{x}_0) \cdot \overline{v} = \begin{pmatrix} 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{pmatrix} = \frac{3}{2} + \frac{1}{\sqrt{2}}$$

- *Let $f : \mathbb{R}^2 \to \mathbb{R}^2$ be given by $f(x, y) = (x^2 + y^2, \mathrm{e}^{x+y})$. Let us compute its directional derivative with respect to $\overline{v} = (0, 1)$ at $\overline{x}_0 = (1, -1)$. We find first the partial derivatives of $f$:*

$$\begin{aligned} \frac{\partial f_1(x,y)}{\partial x} &= 2x & \frac{\partial f_1(x,y)}{\partial y} &= 2y \\ \frac{\partial f_2(x,y)}{\partial x} &= \mathrm{e}^{x+y} & \frac{\partial f_2(x,y)}{\partial y} &= \mathrm{e}^{x+y} \end{aligned}$$

  *which are all continuous at all points of $\mathbb{R}^2$ and in particular at $\overline{x}_0 = (1, -1)$. Therefore, $f$ is differentiable at $\overline{x}_0 = (1, -1)$. The desired directional derivative is*

$$D_{\overline{v}}f(\overline{x}_0) = Jf(\overline{x}_0) \cdot \overline{v} = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

*Note that the result coincides with the value of the partial derivative of $f$ with respect to $y$ because the chosen unitary vector $\overline{v}$ is nothing but the second vector of the canonical basis of $\mathbb{R}^2$:*

$$\frac{\partial f(1,-1)}{\partial y} = \begin{pmatrix} \dfrac{\partial f_1(1,-1)}{\partial y} \\ \dfrac{\partial f_2(1,-1)}{\partial y} \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

## Properties of differentiable functions

**Proposition 28** *Let $f, g : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set, be differentiable functions at $\overline{x}_0 \in D$. Then $f+g$, $\lambda f$ $\forall \lambda \in \mathbb{R}$ and $f \cdot g$ are differentiable at $\overline{x}_0$ with*

$$\begin{aligned} &\nabla(f+g)(\overline{x}_0) = \nabla f(\overline{x}_0) + \nabla g(\overline{x}_0) \\ &\nabla(\lambda f)(\overline{x}_0) = \lambda \nabla f(\overline{x}_0) \\ &\nabla(f \cdot g)(\overline{x}_0) = f(\overline{x}_0)\nabla g(\overline{x}_0) + g(\overline{x}_0)\nabla f(\overline{x}_0) \end{aligned}$$

**Proposition 29** *Let $f : D_1 \to \mathbb{R}^m$, $D_1 \subset \mathbb{R}^n$ open set, $g : D_2 \to \mathbb{R}^p$, $D_2 \subset \mathbb{R}^m$ open set, $h : D_1 \to \mathbb{R}^p$ such that $\forall \overline{x} \in D_1$, $h(\overline{x}) = (g \circ f)(\overline{x})$.*

*If $f$ is differentiable at $\overline{x}_0$ and $g$ is differentiable at $\overline{y}_0 = f(\overline{x}_0)$ then $h$ is differentiable at $\overline{x}_0$ and we have*

$$Jh(\overline{x}_0) = J(g \circ f)(\overline{x}_0) = Jg(\overline{y}_0) \cdot Jf(\overline{x}_0)$$

In this last relation the Jacobian matrices are matrix multiplied, and it is the expression of the *Chain Rule* for (vector-valued) functions of several variables. This Chain Rule generalizes the analogous one for derivable functions of one variable.

**Example**

Let the function $f : \mathbb{R}^2 \to \mathbb{R}^3$ be given by

$$\begin{aligned} f(\overline{x}) &= (f_1(x_1,x_2), f_2(x_1,x_2), f_3(x_1,x_2)) \\ &= (x_1^3, x_1^2 - x_2, 3x_1 + x_2^2) = (y_1, y_2, y_3) \end{aligned}$$

Let the function $g : \mathbb{R}^3 \to \mathbb{R}^2$ be given by

$$\begin{aligned} g(\overline{y}) &= (g_1(y_1,y_2,y_3), g_2(y_1,y_2,y_3)) \\ &= (y_1 y_2^2 y_3, y_1 + y_3^2) = (z_1, z_2) \end{aligned}$$

Then, we have the derivatives

$$\begin{array}{ll} \dfrac{\partial f_1}{\partial x_1} = 3x_1^2 & \dfrac{\partial f_1}{\partial x_2} = 0 \\ \dfrac{\partial f_2}{\partial x_1} = 2x_1 & \dfrac{\partial f_2}{\partial x_2} = -1 \\ \dfrac{\partial f_3}{\partial x_1} = 3 & \dfrac{\partial f_3}{\partial x_2} = 2x_2 \end{array}$$

which are clearly continuous $\forall (x_1, x_2) \in \mathbb{R}^2$, and therefore $f$ is differentiable on $\mathbb{R}^2$.

Moreover, we have the partial derivatives of $g$:

$$\begin{array}{lll} \dfrac{\partial g_1}{\partial y_1} = y_2^2 y_3 & \dfrac{\partial g_1}{\partial y_2} = 2y_1 y_2 y_3 & \dfrac{\partial g_1}{\partial y_3} = y_1 y_2^2 \\ \dfrac{\partial g_2}{\partial y_1} = 1 & \dfrac{\partial g_2}{\partial y_2} = 0 & \dfrac{\partial g_2}{\partial y_3} = 2y_3 \end{array}$$

which are continuous $\forall (y_1, y_2, y_3) \in \mathbb{R}^3$, thus $g$ is differentiable on $\mathbb{R}^3$.

We have the Jacobian matrices of $f$ and $g$:

$$Jf(x_1, x_2) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \\ \dfrac{\partial f_3}{\partial x_1} & \dfrac{\partial f_3}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 3x_1^2 & 0 \\ 2x_1 & -1 \\ 3 & 2x_2 \end{pmatrix}$$

$$Jg(y_1, y_2, y_3) = \begin{pmatrix} \dfrac{\partial g_1}{\partial y_1} & \dfrac{\partial g_1}{\partial y_2} & \dfrac{\partial g_1}{\partial y_3} \\ \dfrac{\partial g_2}{\partial y_1} & \dfrac{\partial g_2}{\partial y_2} & \dfrac{\partial g_2}{\partial y_3} \end{pmatrix} = \begin{pmatrix} y_2^2 y_3 & 2y_1 y_2 y_3 & y_1 y_2^2 \\ 1 & 0 & 2y_3 \end{pmatrix}$$

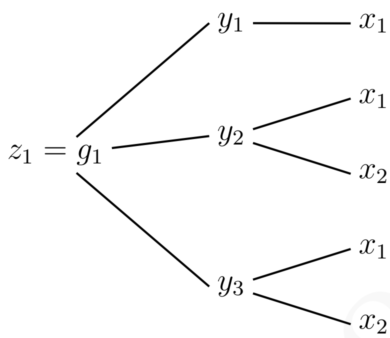
The composite function $g \circ f$ is therefore differentiable $\forall (x_1, x_2) \in \mathbb{R}^2$, with Jacobian matrix which can be computed by means of the Chain Rule:

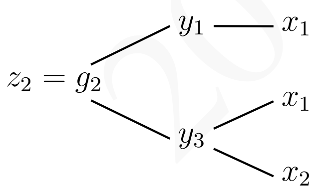
$$\begin{aligned} J(g \circ f)(x_1, x_2) &= Jg(y_1, y_2, y_3) Jf(x_1, x_2) \\ &= \begin{pmatrix} y_2^2 y_3 & 2y_1 y_2 y_3 & y_1 y_2^2 \\ 1 & 0 & 2y_3 \end{pmatrix} \begin{pmatrix} 3x_1^2 & 0 \\ 2x_1 & -1 \\ 3 & 2x_2 \end{pmatrix} \end{aligned} \tag{2}$$

With this relation they are obtained *all* of the (first) partial derivatives of the composite function $g \circ f$ in terms of the (first) partial derivatives of $f$ and $g$.

But there exists a way of computing a (first) partial derivative of the composite function $g \circ f$ individually, something which can be of interest in practical situations.

For that, they are established the *dependences* amongst the variables in *trees* like the ones that follow for the previous example:




If we want to compute the first partial derivative $\dfrac{\partial z_1}{\partial x_1}$, we establish in the first of the previous trees all the possible paths joining the "final variable" $z_1$ with the "initial variable" $x_1$. We obtain in this case three paths. Each path adds up a term to the expression of the derivative, and each of these terms has as many factors as intermediate lines each path has. The expression of our derivative is

$$\frac{\partial z_1}{\partial x_1} = \frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial x_1} + \frac{\partial z_1}{\partial y_2}\frac{\partial y_2}{\partial x_1} + \frac{\partial z_1}{\partial y_3}\frac{\partial y_3}{\partial x_1}$$

Observe how the "intermediate variables" $y_1, y_2, y_3$ are inserted between the final one $z_1$ and the initial one $x_1$ in the expression of the derivative, giving place each one of them to one term of the derivative. Upon substitution of the values of the derivatives computed before, we simply have

$$\frac{\partial z_1}{\partial x_1} = 3y_2^2 y_3 x_1^2 + 4y_1 y_2 y_3 x_1 + 3y_1 y_2^2$$

This value of the derivarive is the one which is obtained from the element in the first row and first column of the matrix product of (2). If we want to evaluate it at a specific point, for example $(x_1, x_2) = (1, 2)$, we **have to** evaluate the values of the intermediate variables according to the function to be evaluated on the given point, in this case $f$ at $(1, 2)$, namely

$$(y_1, y_2, y_3) = f(1, 2) = (1, -1, 7)$$

Thus, we have

$$\frac{\partial z_1(1,2)}{\partial x_1} = 3 \cdot (-1)^2 \cdot 7 \cdot 1^2 + 4 \cdot 1 \cdot (-1) \cdot 7 \cdot 1 + 3 \cdot 1 \cdot (-1)^2 = -4$$

All of these results and the associated techniques can be generalized immediately to the case of more than two composed functions. We leave as an exercise to the reader to check that we have, with the same procedure, the derivatives

$$\begin{aligned}
\frac{\partial z_1}{\partial x_2} &= -2y_1y_2y_3 + 2y_1y_2^2x_2 \\
\frac{\partial z_2}{\partial x_1} &= 3x_1^2 + 6y_3 \\
\frac{\partial z_2}{\partial x_2} &= 4y_3x_2
\end{aligned}$$

We propose the reader to relate these results with the elements of the matrix product (2) and to evaluate these derivatives at the point $(x_1, x_2) = (1, 2)$ or at another one of his/her choice as it has been done before.

## Higher order derivatives

Assuming derivability on each step, from $f(x)$ it is obtained by derivation $f'(x)$; from this last $f''(x), f'''(x), \ldots, f^{(n)}(x), \ldots$, where $n \in \mathbb{N}$. By convention, it is taken $f^{(0)}(x) = f(x)$.

**Definition 47** *Given $f : D \to \mathbb{R}$, $D$ open set of $\mathbb{R}$ and $n \in \mathbb{N}$, it is said that $f$ is of class $C^{(n}$ on $D$ if there exist all of its derivatives to the order $n$ and they are all continuous on $D$, and it is denoted $f \in C^{(n}(D)$. It is said that $f \in C^{(\infty}(D)$ if there exist all of the successive derivatives of $f$ and they are all continuous on $D$.*

**Theorem 2 (Taylor)** *Let $f : D \to \mathbb{R}$, $D$ open set, $x_0 \in D$, such that $f \in C^{(n}(D)$. Then we have the Taylor formula of order $n$ of $f$ at $x_0$,*

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x)$$

*where $R_n(x)$ is the* Taylor's remainder *of order $n$ which has the following properties:*

$$\lim_{x\to x_0} \frac{R_n(x)}{(x-x_0)^k} = \begin{cases} 0 & k = 0, 1, \ldots, n \\ l \in \mathbb{R} & k = n+1 \\ \pm\infty & k > n+1 \end{cases}$$

*If $x_0 = 0$ the previous formula and the remainder are called after* McLaurin.

The version for functions of several variables is as follows. Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set. Let us assume that there exists $\dfrac{\partial f(\overline{x})}{\partial x_i}$ on all of $D$. If they do exist, we can compute also

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f(\overline{x})}{\partial x_i}\right) = \frac{\partial^2 f(\overline{x})}{\partial x_i^2}$$
$$\frac{\partial}{\partial x_j}\left(\frac{\partial f(\overline{x})}{\partial x_i}\right) = \frac{\partial^2 f(\overline{x})}{\partial x_j \partial x_i}$$

which are known as *second-order partial derivatives*. There are a maximum of $n^2$ second-order partial derivatives of a scalar function of $n$ variables, and when we take the derivative with respect to two different variables ($i \neq j$ in the previous expressions) we speak about *second-order crossed partial derivatives*. The process can be iterated in order to obtain third-, fourth-, etc. order derivatives, but their use in Economics is not usual. Nevertheless, the following notion makes use of the concept of partial derivatives of order $r \in \mathbb{N}$:

**Definition 48** *Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set. We say that:*

- *$f \in C^{(r}(D)$ if there exist on $D$ $f$, all of its partial derivatives to the order $r$, being all of $f$ and the mentioned partial derivatives continuous on $D$.*
- *$f \in C^{(\infty}(D)$ if there exist on $D$ $f$, all of its partial derivatives of any order, being all of $f$ and the mentioned partial derivatives continuous on $D$.*
- *If $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$ open set, with $f = (f_1, \ldots, f_m)$, we say that $f \in C^{(r}(D)$ if $f_j \in C^{(r}(D)$, $\forall j = 1, \ldots, m$.*

**Example**
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x^2y + xy^2 + y^2$. Then, we have

$$\frac{\partial f}{\partial x} = 2xy + y^2\,, \quad \frac{\partial f}{\partial y} = x^2 + 2xy + 2y$$

and

$$\frac{\partial^2 f}{\partial x^2} = 2y \qquad \frac{\partial^2 f}{\partial x \partial y} = 2x + 2y$$
$$\frac{\partial^2 f}{\partial y \partial x} = 2x + 2y \quad \frac{\partial^2 f}{\partial y^2} = 2x + 2$$

We observe that the crossed second-order partial derivatives are equal in this case:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 2x + 2y$$

The following result provides a sufficient condition, not neccessary, for this to happen:

**Theorem 3 (Schwarz)** *Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set. If $f \in C^{(2}(D)$, then*

$$\frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\overline{x})}{\partial x_j \partial x_i}, \quad \forall \overline{x} \in D, \quad \forall i, j = 1, \dots, n$$

**Definition 49** *Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set, $\overline{x}_0 \in D$. If there exist all the second-order partial derivatives of $f$ at $\overline{x}_0$, we will call* Hessian matrix *of $f$ at $\overline{x}_0$ to the following $n \times n$ matrix:*

$$Hf(\overline{x}_0) = \begin{pmatrix} \dfrac{\partial^2 f(\overline{x}_0)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 f(\overline{x}_0)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\overline{x}_0)}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f(\overline{x}_0)}{\partial x_n^2} \end{pmatrix}$$

*The determinant $|Hf(\overline{x}_0)| = \det(Hf(\overline{x}_0))$ is called* Hessian *of $f$ at $\overline{x}_0$. If $f \in C^{(2}(D)$, $Hf(\overline{x}_0)$ is a symmetric matrix $\forall \overline{x}_0 \in D$.*

**Example**

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x^2 y$. In this case $D = \mathbb{R}^2$, and as we will see afterwards, $f \in C^{(2}(\mathbb{R}^2)$. Then we have

$$\frac{\partial f}{\partial x} = 2xy, \quad \frac{\partial f}{\partial y} = x^2$$

and

$$\frac{\partial^2 f}{\partial x^2} = 2y \quad \frac{\partial^2 f}{\partial x \partial y} = 2x$$
$$\frac{\partial^2 f}{\partial y \partial x} = 2x \quad \frac{\partial^2 f}{\partial y^2} = 0$$

thus

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

that is symmetric. We also see that $f$, its first and second order derivatives are all continuous on $\mathbb{R}^2$.

We can choose a point in order to evaluate the Hessian matrix on it, for example $(1,1)$ (in a general case, it can be chosen any other point where all of the derivatives are well defined):

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

If we want to compute the Hessian of $f$ at $(1,1)$ we simply have

$$|Hf(1,1)| = -4$$

## Second differential

Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set. If $f \in C^{(2}(D)$ it is defined the *second differential* of $f$ at $\overline{x}_0 \in D$ as the map

$$d^2 f(\overline{x}_0) : \mathbb{R}^n \to \mathbb{R}$$

given by $d^2 f(\overline{x}_0)(\overline{v}) = \overline{v}^t Hf(\overline{x}_0)\overline{v}$, i.e., the quadratic form associated to the symmetric matrix $Hf(\overline{x}_0)$.

**Example**

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = x^2 y$ as in the previous example. We have $D = \mathbb{R}^2$ and $f \in C^{(2}(\mathbb{R}^2)$. We know that

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

and that

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

Thus, we can compute

$$d^2 f(1,1)(\overline{v}) = \overline{v}^t Hf(1,1)\overline{v} = (v_1\ v_2) \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 2v_1^2 + 4v_1 v_2$$

These quadratic forms are intensively used in the analysis of the second-order sufficient conditions of the optimization theory of functions of several variables.

## Taylor formula

Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set, $f \in C^{(3}(D)$. If $\overline{x}_0, \overline{x} \in D$ are such that the segment that joins $\overline{x}_0$ and $\overline{x}$ is contained in $D$, then we can write the second-order Taylor formula of $f$ at $\overline{x}_0$ as follows:

$$\begin{aligned} f(\overline{x}) &= f(\overline{x}_0) + df(\overline{x}_0)(\overline{x} - \overline{x}_0) + \frac{1}{2}d^2(\overline{x}_0)(\overline{x} - \overline{x}_0) + R_2(\overline{x}_0) \\ &= f(\overline{x}_0) + \nabla f(\overline{x}_0) \cdot (\overline{x} - \overline{x}_0) + \frac{1}{2}(\overline{x} - \overline{x}_0)^t Hf(\overline{x}_0)(\overline{x} - \overline{x}_0) + R_2(\overline{x}_0) \end{aligned}$$

where $R_2(\overline{x}_0)$ is the second-order (Taylor) remainder of $f$ at $\overline{x}_0$. If we wanted to obtain the first-order Taylor formula of $f$ at $\overline{x}_0$, it would be simply

$$f(\overline{x}) = f(\overline{x}_0) + \nabla f(\overline{x}_0) \cdot (\overline{x} - \overline{x}_0) + R_1(\overline{x}_0)$$

where $R_1(\overline{x}_0)$ is the first-order (Taylor) remainder of $f$ at $\overline{x}_0$. If $\overline{x}_0 = \overline{0}$ the formulae and the remainders are called after McLaurin. These Taylor formulae are key tools in the study of second-order sufficient conditions of the optimization theory of functions of several variables without constraints and with equality constraints to be seen in the subject Mathematics II.

**Example**

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x, y) = x^2 y$ as in the previous two examples. We have $D = \mathbb{R}^2$ and $f \in C^{(3}(\mathbb{R}^2)$. Let us compute the second-order Taylor formula of $f$ at the point $(1, 1)$.

We know that

$$\nabla f(x, y) = (2xy \quad x^2)$$

thus

$$\nabla f(1, 1) = (2 \quad 1)$$

Also, we know that

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

and therefore

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

In this way, we have

$$\begin{aligned} f(x, y) &= f(1, 1) + \nabla f(1, 1) \cdot (x - 1 \;\; y - 1) \\ &\quad + \frac{1}{2}(x - 1 \;\; y - 1) Hf(1, 1) \begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix} + R_2(1, 1) \\ &= 1 + 2(x - 1) + y - 1 + \frac{1}{2}\left[2(x - 1)^2 + 4(x - 1)(y - 1)\right] + R_2(1, 1) \end{aligned}$$

## Homogeneous functions

**Definition 50** *Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set. We will say that $f$ is homogeneous of degree $p \in \mathbb{R}$ if*

$$f(\lambda\overline{x}) = \lambda^p f(\overline{x})\,, \quad \forall \overline{x} \in D\,, \quad \forall \lambda > 0 \quad \text{such that } \lambda\overline{x} \in D.$$

**Example**
Any real linear function of $n$ variables

$$f(x_1, \dots, x_n) = c_1x_1 + \cdots + c_nx_n\,, \quad D = \mathbb{R}^n$$

where $c_1, \dots, c_n$ are real constants, is homogeneous of degree $p = 1$. In fact, for all $(x_1, \dots, x_n) \in \mathbb{R}^n$ and $\lambda > 0$

$$\begin{aligned} f(\lambda x_1, \dots, \lambda x_n) &= c_1\lambda x_1 + \cdots + c_n\lambda x_n \\ &= \lambda(c_1x_1 + \cdots + c_nx_n) = \lambda f(x_1, \dots, x_n) \end{aligned}$$

**Example**
Any real quadratic form in $n$ variables:

$$Q(X) = X^tAX\,, \quad A \in M_n\,, \quad D = \{X \in M_{n\times 1}\}$$

is homogeneous of degree $p = 2$. In fact, for all $X \in M_{n\times 1}$ and $\lambda > 0$:

$$Q(\lambda X) = (\lambda X)^tA(\lambda X) = \lambda^2X^tAX = \lambda^2Q(X).$$

**Example**
Let the function

$$f(x, y) = \frac{x+y}{x-y}\,, \quad D = \{(x, y) \in \mathbb{R}^2 \,|\, x - y \neq 0\}$$

Then we have for all $(x, y) \in D$ and $\lambda > 0$:

$$f(\lambda x, \lambda y) = \frac{\lambda x + \lambda y}{\lambda x - \lambda y} = \frac{\lambda(x+y)}{\lambda(x-y)} = \frac{x+y}{x-y} = f(x, y) \tag{3}$$

thus $f$ is homogeneous of degree $p = 0$ in $D$.

**Example** Cobb-Douglas function.
Let the function defined for $K > 0$ and $L > 0$:

$$\begin{aligned} &f(K, L) = AK^\alpha L^\beta\,, \quad A > 0\,, \quad 0 < \alpha < 1\,, \quad 0 < \beta < 1 \\ &D = \{(K, L) \in \mathbb{R}^2 \,|\, K > 0, L > 0\} \end{aligned}$$

We have that for all $(K, L) \in D$ and $\lambda > 0$,

$$\begin{aligned} f(\lambda K, \lambda L) &= A(\lambda K)^\alpha (\lambda L)^\beta \\ &= A\lambda^\alpha K^\alpha \lambda^\beta L^\beta = \lambda^{\alpha+\beta} A K^\alpha L^\beta = \lambda^{\alpha+\beta} f(K, L) \end{aligned}$$

thus $f(K, L)$ is homogeneous of degree $p = \alpha + \beta$ in $D$. The Cobb-Douglas functions are intensively used in Economics.

The definition of homogeneous function relates the values of the function at points of its domain $\overline{x}$ and $\lambda\overline{x}$, with $\lambda > 0$.

There exist a characterization of homogeneous functions that relates the values of their derivatives with the values of the function itself at the same points, due to Euler, as follows:

**Theorem 4 (Euler)** *Let $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$ open set, such that $f$ has first-order partial derivatives on $D$. Then it holds that $f$ is homogeneous of degree $p$ if and only if*

$$\nabla f(\overline{x}) \cdot \overline{x} = p f(\overline{x}) \,, \quad \forall \overline{x} \in D$$

*condition that can be written in expanded form as*

$$x_1 \frac{\partial f(x_1, \ldots, x_n)}{\partial x_1} + \cdots + x_n \frac{\partial f(x_1, \ldots, x_n)}{\partial x_n} = p f(x_1, \ldots, x_n) \,, \quad \forall (x_1, \ldots, x_n) \in D$$

**Example**

Let the function

$$f(x, y) = \frac{x^3}{y} \,, \quad D = \{(x, y) \in \mathbb{R}^2 \mid y \neq 0\}$$

We have for all $(x, y) \in D$ and $\lambda > 0$:

$$f(\lambda x, \lambda y) = \frac{(\lambda x)^3}{\lambda y} = \frac{\lambda^3 x^3}{\lambda y} = \lambda^2 \frac{x^3}{y} = \lambda^2 f(x, y)$$

thus $f$ is homogeneous of degree $p = 2$. Let us check how Euler's Theorem is verified in this case. Firstly, let us compute the first-order partial derivatives:

$$\frac{\partial f}{\partial x} = 3\frac{x^2}{y} \,, \quad \frac{\partial f}{\partial y} = -\frac{x^3}{y^2}$$

Then,

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

so we see in another way that $f$ is homogeneous of degree $p = 2$ on $D$.

**Example**

Let the function

$$f(x, y) = \frac{x}{y}, \quad D = \{(x, y) \in \mathbb{R}^2 \,|\, y \neq 0\}$$

We have in this case, for all $(x, y) \in D$ and $\lambda > 0$,

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

so $f$ is homogeneous of degree $p = 0$. Let us see how Euler's Theorem is verified also in this case. The first-order partial derivatives are

$$\frac{\partial f}{\partial x} = \frac{1}{y}, \quad \frac{\partial f}{\partial y} = -\frac{x}{y^2}$$

Thus,

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = \frac{x}{y} - \frac{x}{y} = 0 = 0f(x, y)$$

so we see in another way that $f$ is homogeneous of degree $p = 0$ on $D$.

# Implicit functions

Ocassionally, one or several equations do not define explicitly one or more variables in terms of the rest of them, but the relation is *implicit* instead. Although it could not be possible to write explicitly these functions, it is instead possible to find their derivatives in terms of the implicit functions and the independent variables.

We will develop this topic through showing three examples which we hope would allow the reader to deal with the general case.

The formal Theorems of the implicit function can be found, for example, in the *Sección 8.8* of Jarne et al. (1997).

**Example**: One equation and two variables

Consider the equation

$$y^2 + x\mathrm{e}^{x+y} = 2$$

From it, it is not possible to isolate $y = y(x)$ in an explicit way. The equation is of the form $F(x, y) = C$, where $F(x, y) = y^2 + x\mathrm{e}^{x+y}$ is a function of two variables and $C = 2$ is a constant. We wonder about whether the equation defines implicitly a unique $y = y(x)$ that fulfills the equation, continuous and derivable with continuous derivative on an open neighbourhood of the point $(x, y) = (1, -1)$ and with $y(1) = -1$. The conditions that must be met for that are the three following ones:

i) $F(1, -1) = 2$. It is verified immediately.

ii) $F \in C^{(1}(U)$, where $U$ is an open neighbourhood of $(1, -1)$: we see that

$$\frac{\partial F(x, y)}{\partial x} = \mathrm{e}^{x+y} + x\mathrm{e}^{x+y} = (1 + x)\mathrm{e}^{x+y}$$
$$\frac{\partial F(x, y)}{\partial y} = 2y + x\mathrm{e}^{x+y}$$

and $F$ are continuous on $\mathbb{R}^2$ and thus on any open neighbourhood $U$ of $(1, -1)$.

iii) $\dfrac{\partial F(1, -1)}{\partial y} \neq 0$. In fact, we have $\dfrac{\partial F(1, -1)}{\partial y} = -2 + \mathrm{e}^{1-1} = -1 \neq 0$.

Then it holds that the previous equation defines implicitly $y = y(x)$ on an open neighbourhood of the point $(1, -1)$, with $y(1) = -1$ as indicated.

Let us compute the derivative of $y = y(x)$ implicitly. Assuming this relation, the equation reads

$$(y(x))^2 + x\mathrm{e}^{x+y(x)} = 2$$

Taking the derivative with respect to $x$, we have

$$2y(x)y'(x) + \mathrm{e}^{x+y(x)} + x\mathrm{e}^{x+y(x)}(1 + y'(x)) = 0$$

rearranging, we obtain

$$[2y(x) + x\mathrm{e}^{x+y(x)}]y'(x) = -\mathrm{e}^{x+y(x)}(1 + x)$$

and isolating $y'(x)$,

$$y'(x) = -\frac{\mathrm{e}^{x+y(x)}(1 + x)}{2y(x) + x\mathrm{e}^{x+y(x)}}$$

We know that $y(1) = -1$, so upon substitution,

$$y'(1) = -\frac{\mathrm{e}^{1-1}(1 + 1)}{-2 + \mathrm{e}^{1-1}} = 2$$

**Example**: One equation and three variables

Consider now the equation

$$x^2y + y^2z + z^3x = 3 \tag{4}$$

We wonder about whether (4) defines implicitly a unique $z = z(x, y)$ on an open neighbourhood of the point $(1, 1, 1)$, that verifies the equation, be continuous and derivable with continuous derivatives, such that $z(1, 1) = 1$. In this case we have $F(x, y, z) = C$, with $F(x, y, z) = x^2y + y^2z + z^3x$ and $C = 3$. The three corresponding conditions to be verified in this case are the following ones:

i) $F(1, 1, 1) = 3$. It is immediately verified.

ii) $F \in C^{(1}(U)$, where $U$ is an open neighbourhood of the point $(1, 1, 1)$: we see that

$$\frac{\partial F(x,y,z)}{\partial x} = 2xy + z^3$$
$$\frac{\partial F(x,y,z)}{\partial y} = x^2 + 2yz$$
$$\frac{\partial F(x,y,z)}{\partial z} = y^2 + 3xz^2$$

and $F$ are continuous on $\mathbb{R}^3$ and therefore on any open neighbourhood $U$ of the point $(1, 1, 1)$.

iii) $\dfrac{\partial F(1,1,1)}{\partial z} \neq 0$. In fact, we have $\dfrac{\partial F(1,1,1)}{\partial z} = 1 + 3 = 4 \neq 0$.

Then it holds that the equation (4) defines implicitly $z = z(x, y)$ on an open neighbourhood of the point $(1, 1, 1)$, with $z(1, 1) = 1$ as indicated.

Let us compute the partial derivatives of $z = z(x, y)$ implicitly. Assuming this last relation, the equation (4) reads

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

Let us take first the derivative with respect to $x$. We have

$$2xy + y^2\frac{\partial z(x,y)}{\partial x} + (z(x, y))^3 + 3x(z(x, y))^2\frac{\partial z(x,y)}{\partial x} = 0$$

and rearranging

$$[y^2 + 3x(z(x, y))^2]\frac{\partial z(x,y)}{\partial x} = -2xy - (z(x, y))^3$$

so that

$$\frac{\partial z(x,y)}{\partial x} = -\frac{2xy + (z(x, y))^3}{y^2 + 3x(z(x, y))^2}$$

Similarly, taking the derivative with respect to $y$ it can be found that

$$\frac{\partial z(x,y)}{\partial y} = -\frac{x^2 + 2yz(x,y)}{y^2 + 3x(z(x,y))^2}$$

Let us evaluate these derivatives at $(1,1)$, recalling that $z(1,1) = 1$:

$$\begin{aligned} \frac{\partial z(1,1)}{\partial x} &= -\frac{2+1}{1+3} = -\frac{3}{4} \\ \frac{\partial z(1,1)}{\partial y} &= -\frac{1+2}{1+3} = -\frac{3}{4} \end{aligned}$$

**Example**: Two equations and four variables
Consider now the system

$$\left\{ \begin{array}{l} x^2 + xy + 2u - v = 0 \\ xy - y^2 + u + 2v = 0 \end{array} \right. \tag{5}$$

We wonder about whether (5) defines implicitly unique functions $x = x(u,v)$, $y = y(u,v)$ such that they verify the system, they are continuous and derivable with continuous derivatives on an open neighbourhood of the point $(x,y,u,v) = \left(0, 1, \frac{1}{5}, \frac{2}{5}\right)$ with $x\left(\frac{1}{5}, \frac{2}{5}\right) = 0$, $y\left(\frac{1}{5}, \frac{2}{5}\right) = 1$. Now we have $F_1(x,y,u,v) = x^2 + xy + 2u - v$, $F_2(x,y,u,v) = xy - y^2 + u + 2v$, $C_1 = 0$, $C_2 = 0$. The three conditions to be verified in this case are the following ones:

i) $F_1\left(0, 1, \frac{1}{5}, \frac{2}{5}\right) = 0$, $F_2\left(0, 1, \frac{1}{5}, \frac{2}{5}\right) = 0$. They are verified by direct substitution.

ii) $F_1, F_2 \in C^{(1}(U)$, where $U$ is an open neighbourhood of the point $\left(0, 1, \frac{1}{5}, \frac{2}{5}\right)$: we see that

$$\begin{array}{llll} \dfrac{\partial F_1}{\partial x} = 2x + y & \dfrac{\partial F_1}{\partial y} = x & \dfrac{\partial F_1}{\partial u} = 2 & \dfrac{\partial F_1}{\partial v} = -1 \\ \dfrac{\partial F_2}{\partial x} = y & \dfrac{\partial F_2}{\partial y} = x - 2y & \dfrac{\partial F_2}{\partial u} = 1 & \dfrac{\partial F_2}{\partial v} = 2 \end{array}$$

and $F_1, F_2$ are all continuous on $\mathbb{R}^4$ and thus on any open neighbourhood $U$ of the point $\left(0, 1, \frac{1}{5}, \frac{2}{5}\right)$.

iii) $\det \begin{pmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{pmatrix}_{|(x,y,u,v)=\left(0,1,\frac{1}{5},\frac{2}{5}\right)} \neq 0.$

In fact, this determinant is

$$\det\begin{pmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{pmatrix}_{|(x,y,u,v)=\left(0,1,\frac{1}{5},\frac{2}{5}\right)} = \det\begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix} = -2 \neq 0$$

Then, it holds that (5) defines implicitly $x = x(u,v)$, $y = y(u,v)$ on an open neighbourhood of the point $(x,y,u,v) = \left(0,1,\frac{1}{5},\frac{2}{5}\right)$, with $x\left(\frac{1}{5},\frac{2}{5}\right) = 0$, $y\left(\frac{1}{5},\frac{2}{5}\right) = 1$ and with the indicated properties.

Let us compute the partial derivatives of these implicit functions with respect to $u, v$. For that, we write $x = x(u,v)$ and $y = y(u,v)$ so that the system (5) can be written

$$\left\{\begin{array}{l} F_1(x(u,v),y(u,v),u,v) = 0 \\ F_2(x(u,v),y(u,v),u,v) = 0 \end{array}\right. \tag{6}$$

We take the derivatives with respect to $u$ and $v$ and applying the Chain Rule:

$$\begin{array}{ll} \dfrac{\partial F_1}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial F_1}{\partial y}\dfrac{\partial y}{\partial u} + \dfrac{\partial F_1}{\partial u} = 0 & \dfrac{\partial F_1}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial F_1}{\partial y}\dfrac{\partial y}{\partial v} + \dfrac{\partial F_1}{\partial v} = 0 \\ \dfrac{\partial F_2}{\partial x}\dfrac{\partial x}{\partial u} + \dfrac{\partial F_2}{\partial y}\dfrac{\partial y}{\partial u} + \dfrac{\partial F_2}{\partial u} = 0 & \dfrac{\partial F_2}{\partial x}\dfrac{\partial x}{\partial v} + \dfrac{\partial F_2}{\partial y}\dfrac{\partial y}{\partial v} + \dfrac{\partial F_2}{\partial v} = 0 \end{array}$$

and written in matrix form

$$\begin{pmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{pmatrix}\begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix} = -\begin{pmatrix} \dfrac{\partial F_1}{\partial u} & \dfrac{\partial F_1}{\partial v} \\ \dfrac{\partial F_2}{\partial u} & \dfrac{\partial F_2}{\partial v} \end{pmatrix}$$

Evaluating this last equality at $(x,y,u,v) = \left(0,1,\frac{1}{5},\frac{2}{5}\right)$ we have

$$\begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix}_{|(u,v)=\left(\frac{1}{5},\frac{2}{5}\right)} = -\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$$

so that

$$\begin{aligned} \begin{pmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{pmatrix}_{|(u,v)=\left(\frac{1}{5},\frac{2}{5}\right)} &= -\begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix}^{-1}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} -2 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -4 & 2 \\ -1 & 3 \end{pmatrix} \end{aligned}$$

# Bibliography

Jarne, G., Pérez-Grasa, I., and Minguillón, E. (1997). *Matemáticas para la Economía. Álgebra lineal y cálculo diferencial*. McGraw-Hill.