Unit 2. Convexity and concavity

The convexity and concavity concepts are very useful for finding the maximum and minimum of the functions since these concepts make it considerably easy to reach a mathematical solution to optimisation problems. We will be able to confirm that convex functions have special qualities that simplify the search for minimums while concave functions produce the same results in the search for maximums.

It is important to point out that all functions referred to throughout this chapter will be scalar functions, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, that is, functions the result of which is a real value and not a vector.

1 Convex sets. Convex envelope

The convexity and concavity concept in functions is closely related to the convexity of sets. This is the reason why we begin by defining and studying the convex sets of $\mathbb{R}^n$.

Definition 1. *A set A in $\mathbb{R}^n$ is a* convex set *when any segment connecting a couple of points in A is fully contained in A, that is, all the points of the segment belong to A .*

Example 2. *The set*

$$A = \{(x, y) \in \mathbb{R}^2 \ / \ y \leq 3 - x\}$$

is a convex set, since it is easy to confirm graphically (see figure 2) that any segment connecting two points of A is completely included in A.

Example 3. *The set formed by the graphic of any non-linear function is not convex. This could be seen, for example, in the figure 3, where the set $B = \{(x, y) \in \mathbb{R}^2 \ / \ y = e^x\}$ is represented.*

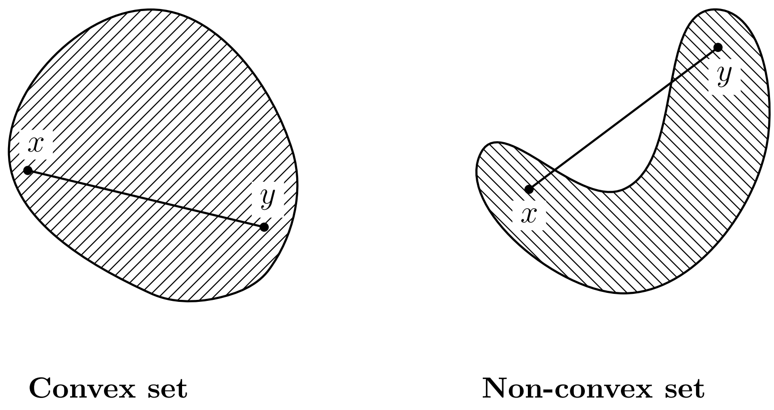

Figure 1: Convex set and non-convex set

Property 4. *The* $I\!R^n$ *space is a convex set.*

Property 5. *Intersection of convex sets is a convex set.*

Example 6. *The set union may or may not be a convex set. For example, consider sets* $A = \{(x, y) \in I\!R^2 \,/\, x^2 + y^2 = 1\}$, $B = \{(x, y) \in I\!R^2 \,/\, (x-1)^2 + (y-1)^2 = 1\}$ *and* $C = \{(x, y) \in I\!R^2 \,/\, x^2 + y^2 = 4\}$, *which are circles in* $I\!R^2$. A, B *and* C *are convex sets.* $A \cup B$ *is not convex, while* $A \cup C$ *is convex.*

Definition 7. *We will call* hyperplane *in* $I\!R^n$ *the set of points that verify that the linear equation in* $I\!R^n$, *that is,*

$$H = \{(x_1, x_2, \ldots, x_n) \in I\!R^n \,/\, c_1x_1 + c_2x_2 + \cdots + c_nx_n = \alpha\}$$

where $c_1, c_2, \ldots, c_n, \alpha \in I\!R$ *and at least* $c_i \neq 0$.

Example 8. *Straight lines in* $I\!R^2$ *and planes in* $I\!R^3$ *are hyperplanes.*

Example 9. *The set* $\left\{(x, y, z) \in I\!R^3 \,/\, 3x + y - z = -3\right\}$ *is a hyperplane in* $I\!R^3$.

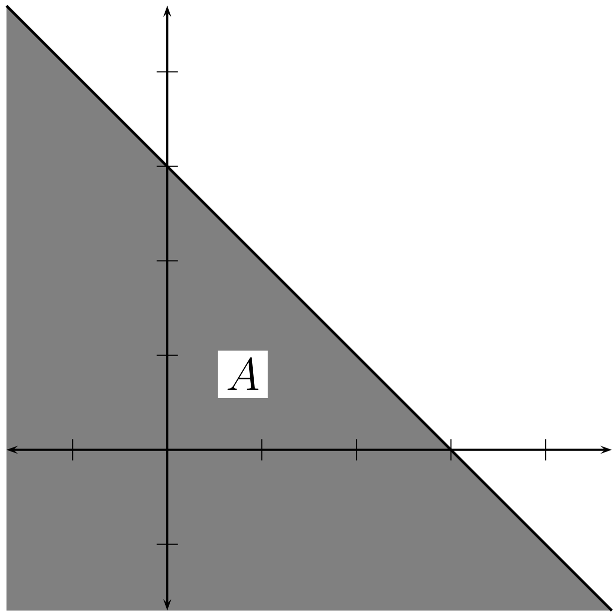

Figure 2: $A = \{(x, y) \in \mathbb{R}^2 \ / \ y \leq 3 - x\}$.

Example 10. *The sets*

$$\left\{(x, y, z) \in \mathbb{R}^3 \ / \ x^2 + y - z = 1\right\} \ \textit{and} \ \left\{(x, y, z) \in \mathbb{R}^3 \ / \ 2xy + 3z = 1\right\}$$

aren't hyperplanes in $\mathbb{R}^3$.

Definition 11. *Let H be a hyperplane in $\mathbb{R}^n$. It is possible to define new sets which are called* semi-spaces *in* $\mathbb{R}^n$:

$$H^+ = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \ / \ c_1x_1 + c_2x_2 + \cdots + c_nx_n \geq \alpha\} \ \textbf{(Upper semi-space)}$$

$$H^- = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \ / \ c_1x_1 + c_2x_2 + \cdots + c_nx_n \leq \alpha\} \ \textbf{(Lower semi-space)}$$

Example 12. *The sets*

$$\left\{(x, y, z) \in \mathbb{R}^3 \ / \ 3x + y - z \geq -3\right\} \ \textit{and} \ \left\{(x, y, z) \in \mathbb{R}^3 \ / \ 3x + y - z \leq -3\right\}$$

are semi-spaces $\mathbb{R}^3$.

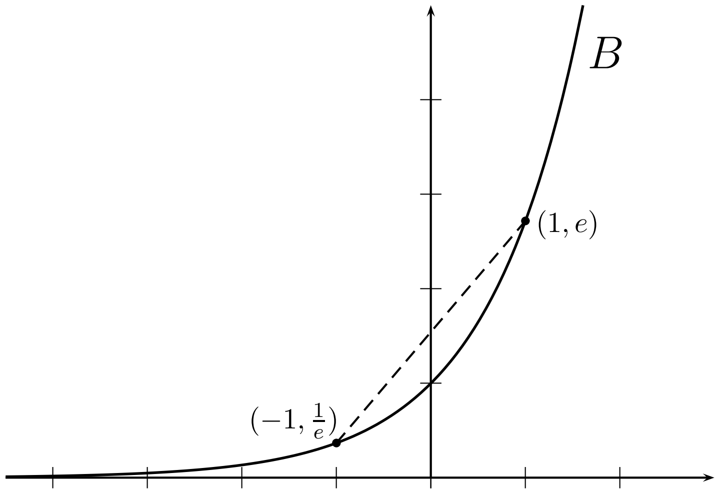

Figure 3: $B = \{(x, y) \in \mathbb{R}^2 \,/\, y = e^x\}$.

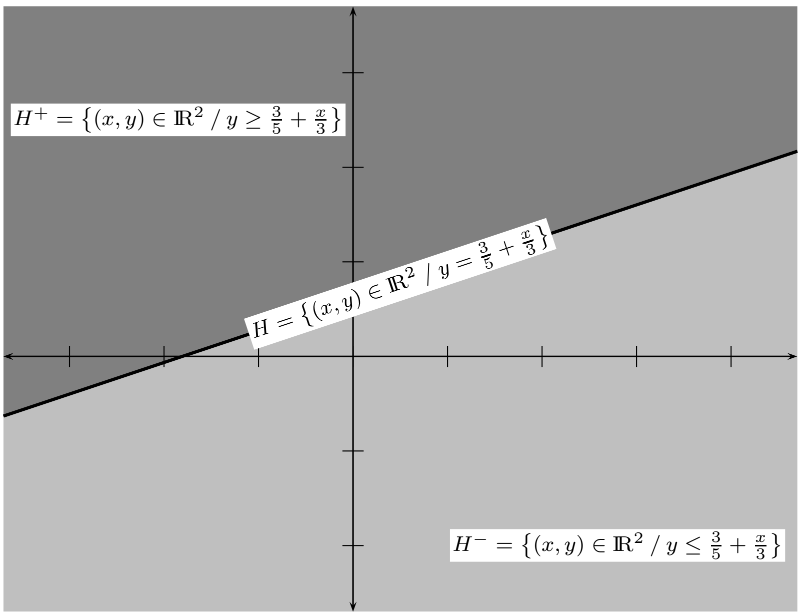

Figure 4: Each hyperplane divides the space $\mathbb{R}^n$ into two semi-spaces.

Definition 13. *The intersection of a finite number of semi-spaces in $\mathbb{R}^n$ is called* **polyhedron** *(See figure 5).*

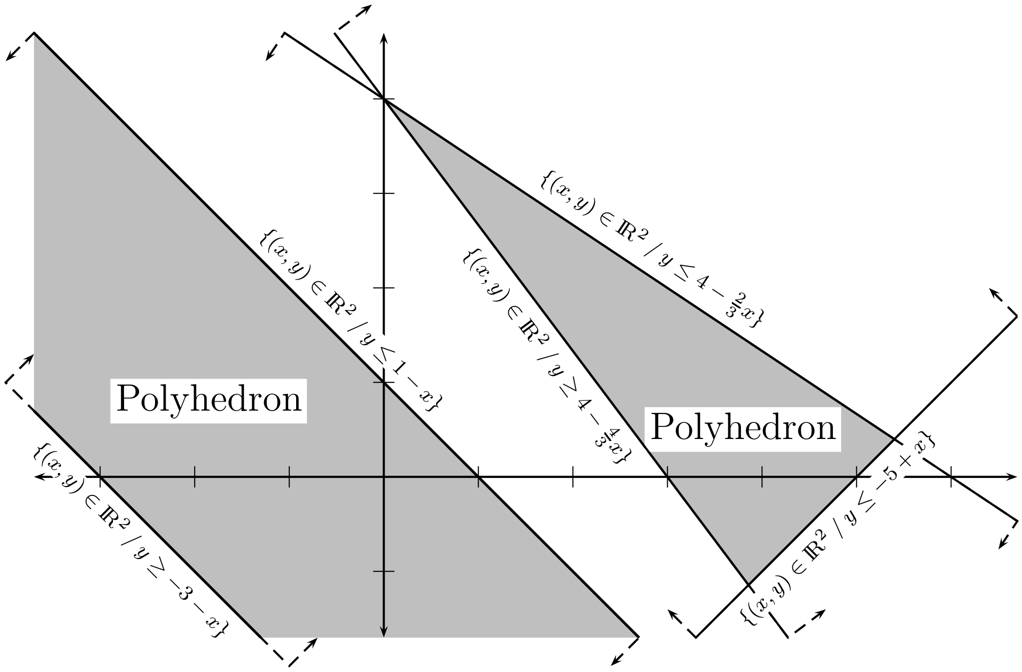

Figure 5: Two polyhedrons in the plane.

Property 14. *Hyperplanes, semi-spaces and polyhedrons in $\mathbb{R}^n$ are all convex sets.*

Example 15. *A very useful type of set in Linear Programming (which we will study in Unit 3) are those formed by points of $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ the coordinates of which are all greater or equal to zero and which also fulfil one or several linear equalities or inequalities, for example, the formed by the points $(x_1, x_2, x_3) \in \mathbb{R}^3$ which verify the following conditions:*

$$3x_1 + \frac{4}{5}x_2 - 6x_3 \leq 5$$
$$\frac{1}{3}x_1 - 3x_2 + 7x_3 = 2$$
$$x_1 \geq 0, \;\; x_2 \geq 0, \;\; x_3 \geq 0$$

These types of sets are always convex. To confirm this, let us assume that set A is the one formed by points $(x_1, x_2, \ldots, x_n) \in I\!R^n$ *which verify the following conditions (where* a_{ij} *are parameters):*

$$
\begin{array}{c}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\
\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m \\
x_1 \geq 0, \ x_2 \geq 0, \ \ldots, \ x_n \geq 0
\end{array}
$$

We shall assign a name to the set of points that verify each of the individual conditions:

$$
\begin{array}{rcl}
A_1 & = & \{(x_1, x_2, \ldots, x_n) \,/\, a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1\} \\
A_2 & = & \{(x_1, x_2, \ldots, x_n) \,/\, a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2\} \\
\vdots & & \vdots \qquad\qquad\qquad \vdots \\
A_m & = & \{(x_1, x_2, \ldots, x_n) \,/\, a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m\} \\
A_{m+1} & = & \{(x_1, x_2, \ldots, x_n) \,/\, x_1 \geq 0\} \\
A_{m+2} & = & \{(x_1, x_2, \ldots, x_n) \,/\, x_2 \geq 0\} \\
\vdots & & \vdots \\
A_{m+n} & = & \{(x_1, x_2, \ldots, x_n) \,/\, x_n \geq 0\}
\end{array}
$$

The points of A are the ones that comply with all conditions simultaneously and as such, they belong to all the previous sets, and accordingly

$$A = A_1 \cap A_2 \cap \cdots \cap A_m \cap A_{m+1} \cap \cdots \cap A_{m+n},$$

and given that each set A_i is convex (or a hyperplane or a semi-space of $I\!R^n$) it can be deducted that A is a convex intersection of convex sets.

Definition 16. *Given a finite set of points $x_1, x_2, \ldots, x_k$ in $I\!R^n$, a linear combination of $x_1, x_2, \ldots, x_k$ in which all scalars are positive and add up to 1, is called a* convex combination*, that is, x is a convex combination of $x_1, x_2, \ldots, x_k$ if there are $\lambda_1, \lambda_2, \ldots, \lambda_k \in I\!R$ such as*

$$x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k$$

and it fulfils $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$ and $\lambda_1, \lambda_2, \ldots, \lambda_k \geq 0$.

The convex combination of two points x_1, x_2 can be written using a single parameter λ:

$$x = \lambda x_1 + (1 - \lambda) x_2 \,,$$

being $\lambda \in [0, 1]$.

Definition 17. *Given a finite amount of points $x_1, x_2, \ldots, x_k$ in $I\!R^n$, the set formed by all the possible convex combinations of these points is called* convex envelope *and is represented by $CE[x_1, x_2, \ldots, x_k]$.*

$$\mathrm{EC}(x_1, x_2, \ldots, x_k) = \left\{ \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k \;\Big/\; \begin{array}{l} \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1 \\ \lambda_1, \lambda_2, \ldots, \lambda_k \geq 0 \end{array} \right\}.$$

The convex envelope of two points x_1, x_2 is the segment connecting x_1 and x_2*. Besides calling it $\mathrm{EC}(x_1, x_2)$ it is usually simply denoted $[x_1, x_2]$.*

$$\mathrm{EC}(x_1, x_2) = [x_1, x_2] = \{\lambda x_1 + (1 - \lambda) x_2 \,/\, \lambda \in [0, 1]\} \,.$$

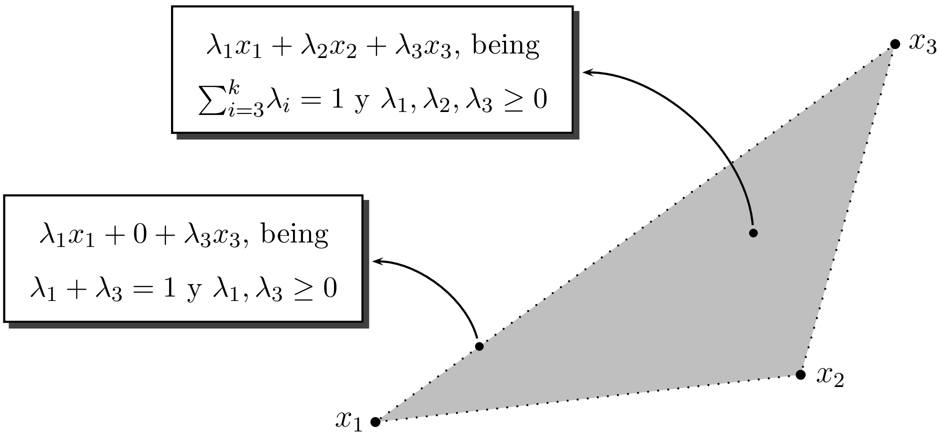

Figure 6: Convex envelope of three non-aligned points in a plane.

Example 18. *Let us consider three points* $(0,0),(1,0),(0,1) \in I\!R^2$. *The convex envelope of these points is:*

$$\mathrm{EC}[(0,0),(1,0),(0,1)] = \left\{ \lambda_1(0,0) + \lambda_2(1,0) + \lambda_3(0,1) \;\Big/\; \begin{array}{c} \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0 \end{array} \right\}$$

$$= \left\{ (\lambda_2, \lambda_3) \;\Big/\; \begin{array}{c} \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0 \end{array} \right\}.$$

Determining a point belonging to the convex envelope is as simple as assigning values to the multipliers $\lambda_1, \lambda_2, \lambda_3$ *that meet the established requisites. Hence, taking* $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$ *it can be deduced that*

$$\frac{1}{3}(0,0) + \frac{1}{3}(1,0) + \frac{1}{3}(0,1) = \left(\frac{1}{3}, \frac{1}{3}\right) \in \mathrm{EC}[(0,0),(1,0),(0,1)]\,.$$

In the same manner, given a point of $I\!R^2$, *it is possible to decide whether or not this point belongs to the convex envelope. If we take point* $(1,1)$, *it belongs to the convex envelope if*

$$\lambda_1(0,0)+\lambda_2(1,0)+\lambda_3(0,1)=(\lambda_2,\lambda_3)=(1,1),\ \textit{with}\ \left.\begin{array}{r}\lambda_1+\lambda_2+\lambda_3=1\\ \lambda_1,\lambda_2,\lambda_3\geq 0\end{array}\right\}$$

which is impossible since if $\lambda_2=\lambda_3=1$, then $\lambda_1+\lambda_2+\lambda_3\neq 1$ being $\lambda_1\geq 0$. Therefore,

$$(1,1)\notin \mathrm{EC}[(0,0),(1,0),(0,1)]\,.$$

Example 19. *Consider $x_1=(1,0,-1)$ and $x_2=(0,1,2)$. The convex envelope of these two points, or the segment connecting them, is:*

$$[x_1,x_2]=\{\lambda(1,0,-1)+(1-\lambda)(0,1,2)\ /\ \lambda\in[0,1]\}=\{(\lambda,1-\lambda,2-3\lambda)\ /\ \lambda\in[0,1]\}\,.$$

To calculate a point of this segment, we just have to give a value of between 0 and 1 to λ. For example, if we take $\lambda=1/5$ we have that $x=(\frac{1}{5},\frac{4}{5},\frac{7}{5})$ is a point in the segment $[x_1,x_2]$. Whereas, if we want to know whether or not another point $y=(\frac{2}{3},0,1)$ belongs to then convex envelope, we can check if $(\lambda,1-\lambda,2-3\lambda)$, for any $\lambda\in[0,1]$. To do this, we equalize each component:

$$\left.\begin{array}{r}\lambda=\frac{2}{3}\\ 1-\lambda=0\\ 2-3\lambda=1\end{array}\right\}$$

There is no solution for this system and, therefore $y=(\frac{2}{3},0,1)\notin[x_1,x_2]$.

Let us now see if point $z=(2,-1,-4)$ belongs to $[x_1,x_2]$. The system

$$\left.\begin{array}{r}\lambda=2\\ 1-\lambda=-1\\ 2-3\lambda=-4\end{array}\right\}$$

does have a solution $\lambda=2$, however, it is $\lambda\notin[0,1]$, which means that z is not a point in segment $[x_1,x_2]$.

2 Convex and concave functions

Definition 20. *A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be* convex *in a set $A \subseteq \mathbb{R}^n$ if two conditions are met:*

1. *A is a convex set.*
2. *Any segment connecting two points of the graph of the function has all its points on the graph of the function or on top of it.*

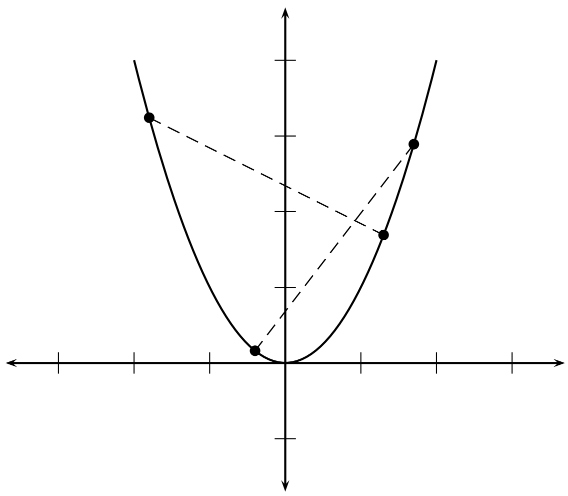

Figure 7: Function $f(x) = x^2$.

Example 21. *Function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^2$ is a convex function in $\mathbb{R}$ since it meets the two conditions provided:*

1. *$\mathbb{R}$ is a convex set.*
2. *Any segment connecting two points in the graph of the function has all its points in or above the function's graph and never below it (See figure 7).*

Example 22. *Function* $f : I\!R^2 \longrightarrow I\!R$ *defined by* $f(x, y) = x^2 + y^2 - 5$ *is a convex function in* $I\!R^2$ *since it meets the necessary requirements:*

1. $I\!R^2$ *is a convex set.*

2. *Any segment connecting two points of the function graph has all its points in the function graph or above it, never below the graph (See figure 8).*

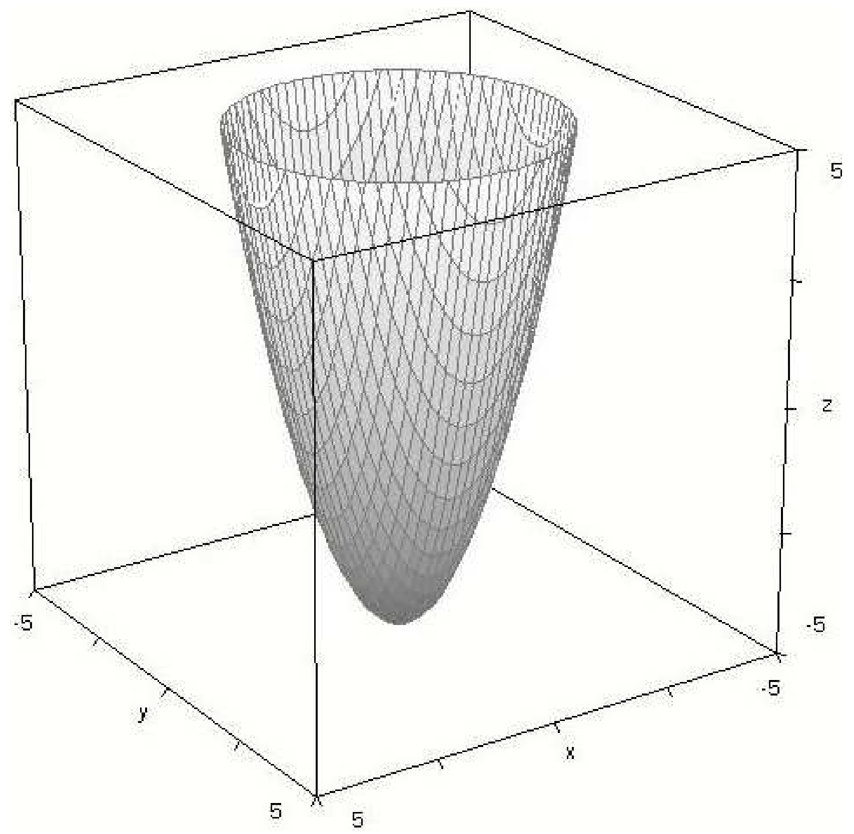

Figure 8: Function $f(x, y) = x^2 + y^2 - 5$.

Definition 23. *A function* $f : I\!R^n \longrightarrow I\!R$ *is said to be* concave *over a set* $A \subseteq I\!R^n$ *if it meets two conditions:*

1. *A is a convex set.*

2. *Any segment connecting two points of the function graph has all its points on the function graph or below it.*

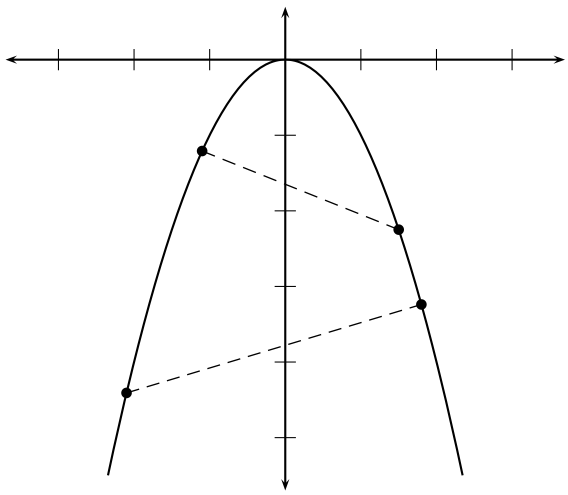

Figure 9: Function $f(x) = -x^2$.

Example 24. *Function $f : \mathbb{R} \to \mathbb{R}$ given as $f(x) = -x^2$ is a concave function in $\mathbb{R}$ since it meets the two conditions provided:*

1. *$\mathbb{R}$ is a convex set.*

2. *Any segment connecting two points of the function graph has all its points on or below the graph and never above it (See figure 9).*

The convexity and concavity concepts of a function are symmetrical, but not opposites, that is, a function could be concave or convex or neither one. In fact, in the same function the two concepts could be combined in different ways. For example, the $f(x) = \sin(x)$ is convex in $A = [\pi, 2\pi]$ and concave $B = [0, \pi]$ (note that A and B are convex sets), but is neither concave nor convex over the entire $\mathbb{R}$ (see figure 10).

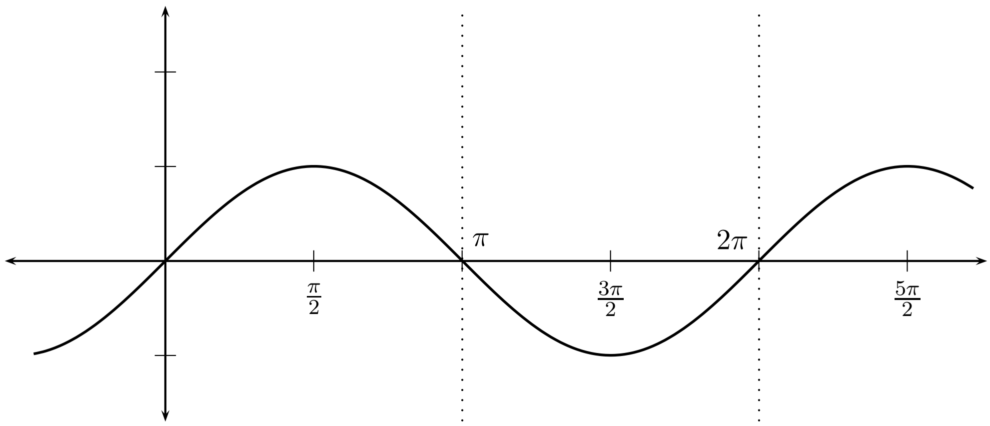

Figure 10: Function $f(x) = \sin(x)$.

Likewise, a function could be convex and concave simultaneously in the same set. This, however, only happens with linear functions.

Property 25. *The only convex and concave functions over an A set are the linear functions defined in A.*

Example 26. *Function $f : I\!R^2 \to I\!R$ defined as $f(x, y) = 3x - y + 2$ is a linear function in $I\!R^2$ (which is a convex set) and, therefore, f is convex and concave in $I\!R^2$.*

It is evident that if a function is convex over a certain set A, then it is also convex over a subset of A, provided that this subset is also a convex set. The same applies to concave functions.

Property 27. *If a function is convex or concave over an A set, then this is also convex over any $B \subseteq A$, provided that B is also a convex set.*

The symmetry existing between convexity and concavity concepts means that for each result over convex functions there is also a symmetrical result over concave functions so that henceforth, all results will present one property for convex functions and the equivalent property for concave functions.

♣ **Avoid this frequent mistake:** Note that whereas there are two concepts when we speak of functions: *concavidad* and *convexidad*, if we refer to a set, only the *convexity* concept must be used. A set can or cannot be convex, but there is no such thing as a *conjunto cóncavo*, which is frequently used mistakenly.

The following results are useful on certain occasions to determine the convexity or concavity of a function over a set.

Property 28. *If a function f is convex over an A set, then:*

1. *When multiplied by a positive number $\alpha \geq 0$, the resulting function αf is also convex over A.*
2. *When multiplied by a negative number $\alpha \leq 0$, the resulting function αf is* concave *over A*

Property 29. *If a function f is concave over an A set, then:*

1. *When multiplied by a positive number $\alpha \geq 0$, the resulting function αf is also concave over A.*
2. *When multiplied by a negative number $\alpha \leq 0$, the resulting function αf is* convex *over A*

Property 30. *If a set of functions is convex over a set A, then the addition of these functions is convex over A.*

Property 31. *If a set of functions is concave over a set A, then the addition of these functions is concave over A.*

Example 32. *Let us confirm if the* $f(x,y) = 3x^2 + \frac{3y^2}{2}$ *function is convex over* $\mathbb{R}^2$*. Firstly, we observe that the* $f_1(x,y) = x^2$ *and* $f_2(x,y) = y^2$ *are convex in* $\mathbb{R}^2$ *(see example 21). Applying the above-mentioned properties, we obtain that* $3x^2$ *and* $\frac{3y^2}{2}$ *are also convex functions over* $\mathbb{R}^2$ *(as they're convex functions multiplied by positive numbers) and so, its addition,* $3x^2 + \frac{3y^2}{2}$ *is also convex. So we can conclude that* $f(x,y) = 3x^2 + \frac{3y^2}{2}$ *is convex in* $\mathbb{R}^2$*.*

Example 33. *Let us consider the following functions:*

$$\begin{aligned} g_1(x) &= \frac{5}{2}x^2 + 3x + \frac{1}{2} \\ g_2(x) &= \frac{5}{2}x^2 - 3x - \frac{1}{2} \\ g_3(x) &= -\frac{5}{2}x^2 + 3x + \frac{1}{2} \end{aligned}$$

Since in example 21 we confirmed that $f(x) = x^2$ *is a convex function in* $\mathbb{R}$ *and* $g(x) = 3x + \frac{1}{2}$ *is a linear function and, therefore, convex over* $\mathbb{R}$ *(see property 25), we can ensure that* g_1 *is convex in* $\mathbb{R}$*. Likewise, we can conclude that* g_2 *is also a convex function in* $\mathbb{R}$*.*

On the other hand, $-\frac{5}{2}x^2$ *is concave in* $\mathbb{R}$*, and since* $3x + \frac{1}{2}$ *is a linear function, it's also concave in* $\mathbb{R}$*. and, accordingly, the sum of these functions, that is* g_3*, is also concave in* $\mathbb{R}$*.*

In practice, however, the best method to determine the convexity or concavity of a function over a convex set A, is to study the sign of the Hessian matrix of the function over all the points of A.

Algorithm 1 (to study the convexity/concavity of a function)

Let f be defined over a set A.

- **Step 1.** Confirm if A is a convex set. If not, it does not make any sense to talk about the concavity or convexity of f in A (although we could study it for subsets of A).

- **Step 2.** Calculate the Hessian matrix $Hf(x)$, and study its sign for any $x \in A$. Then:
  - If $Hf(x)$ positive semidefinite or definite for all points $x \in A$, then f is a convex function in A.
  - If $Hf(x)$ negative semidefinite or definite for all points $x \in A$, then f is a concave function in A.
  - Otherwise, f is not concave or convex A.

It is important to point out that the study should be carried out over all the points in the set. If we find any point where the Hessian matrix isn't positive semidefinite or definite or negative semidefinite or definite, then the function would not be convex or concave over the aforementioned set.

Example 34. *In example 22 we showed that $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, defined as $f(x, y) = x^2 + y^2 - 5$, is convex in $\mathbb{R}^2$. We would reach the same conclusion using the above-mentioned algorithm.*

The Hessian matrix of f at any point $(x, y) \in \mathbb{R}^2$ is:

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

Since $Hf(x, y)$ is positive definite at any point $(x, y) \in \mathbb{R}^2$ and $\mathbb{R}^2$ is a convex set, the conclusion is that f is convex in $\mathbb{R}^2$.

Example 35. *Let's consider function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, defined as $f(x, y) = (x - 2)^4 + (x - y)^2 + 300$. Its Hessian matrix at any point $(x, y) \in \mathbb{R}^2$ is*

$$Hf(x,y) = \begin{pmatrix} 12(x-2)^2+2 & -2 \\ -2 & 2 \end{pmatrix}.$$

The minors of $Hf(x,y)$ are:

$$H_1 = 12(x-2)^2 + 2$$
$$H_2 = 24(x-2)^2$$

Since $H_1 > 0$ and $H_2 \geq 0$ for any $(x,y) \in \mathbb{R}^2$, it can be asserted that the matrix is positive semidefinite in all the points of $\mathbb{R}^2$. Lastly, since $\mathbb{R}^2$ is a convex set, we can conclude that f is a convex function over $\mathbb{R}^2$.

Example 36. *We will study the convexity or concavity of the function $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$, defined as $f(x,y,z) = -(x-2)^2 - e^{2y} + z^3$. To this end, we calculate its Hessian matrix at any point in de $\mathbb{R}^3$ is:*

$$Hf(x,y,z) = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -4e^{2y} & 0 \\ 0 & 0 & 6z \end{pmatrix}$$

The eigenvalues of $Hf(x,y,z)$ are $\lambda_1 = -2$, $\lambda_2 = -4e^{2y}$ and $\lambda_3 = 6z$. For any point (x,y,z), the eigenvalues λ_1 and λ_2 are negative; nevertheless, the λ_3 sign depends on the z value:

- *If $z > 0$ then $\lambda_3 > 0$. In this case, the Hessian matrix is indefinite.*
- *If $z \leq 0$ then $\lambda_1, \lambda_2, \lambda_3 \leq 0$, which means that in this case Hessian matrix is negative semidefinite.*

The $A = \{(x, y, z)I\!R^3 \,/\, z \leq 0\}$ set is convex es convex, since it is a hyperplane in $I\!R^3$, and the Hessian matrix is negative semidefinite for all the points of A, therefore, f is concave over set A. Other than in set A, f is not concave or convex in any other set.

3 Applications to optimisation

The study of the convexity and concavity of functions is mainly justified because of their interesting applications in the field of optimisation. One of the most relevant aspects is that all the results presented below refer to global optimums rather than to local extremes as previously studied in Unit 1. This is of great importance in economic situations; the global maximum or minimum points are always sought while local maximums and minimums are practically irrelevant. When we study the minimum of a function over a set, knowing if that function is convex over such set is very useful since, if that is the case, it simplifies the search enormously. The same holds true in the search of the maximum of a function over a set when the function is concave over the set. We should bear in mind that the search procedure for the local minimums and maximums of a function, which we studied in the previous chapter, was based on the following steps:

1. Finding the critical points of the function.
2. Substituting each critical point in the Hessian matrix of the function. Depending on its sign:
   - If the Hessian matrix is positive definite in that point, then the point is a local minimum.
   - If the Hessian matrix is negative definite in that point, then the point is a local maximum.

- If the Hessian matrix is indefinite in that point, then it is a saddle point.
- If the Hessian matrix is positive or negative semidefinite in that point, then the process is not able to determine whether it is a local minimum, a local maximum or a saddle point.

This method poses two major problems: firstly, it only determines the local optimums of the function, but not its global optimums, and, secondly, there are often critical points for which it is not possible to determine if they are local minimums or maximums or saddle points of the function. The following theorems solve these two problems when we search for the minimums of a convex function and the maximums of a concave function.

Theorem 37. *When a function is differentiable and convex over a convex set A, the concepts* critical point of f in A *and* global minimum of f in A *are equivalent concepts. That is, if f is convex in A, all the critical points of f in A are global minimums of f in A.*

Theorem 38. *When a function is differentiable and concave over a convex set A, the concepts* critical point of f in A *and* global maximum of f in A *are equivalent concepts. That is, if f is concave in A, all the critical points of f in A are global maximums of f in A.*

The convexity or concavity of a function also enables us to determine the number of global maximums and minimums that a function may have in a given set.

Property 39. *A convex function over a convex set A has 0, 1 or infinite global minimums over A. If f is convex and it has more than one global minimum $x_1, x_2, \ldots, x_n$, then all the points of its convex envelope are also global minimums of the function.*

Property 40. *A concave function over a convex set A has 0, 1 or infinite global maximums over A. If f is concave and it has more than one global maximum $x_1, x_2, \ldots, x_n$, then all the points of its convex envelope are also global maximums of the function.*

Summarising the foregoing results, we can establish the steps to study the local and global optimums of a function in the following algorithm.

Algorithm 2 (to study the local and global optimums of a function)

Let f be a function defined on a set A.

- **Step 1.** Calculate the critical points of f in A. If there is none, then f does not have local or global optimums in A. If any, go to step 2.
- **Step 2.** Calculate the Hessian matrix of f, $Hf(x)$. Go to step 3.
- **Step 3.** For any critical point x_0, calculate the Hessian matrix of f at that point, $Hf(x_0)$. Then:
  - If $Hf(x_0)$ is positive definite then x_0 is a *local* minimum. Go to step 4.
  - If $Hf(x_0)$ negative definite then x_0 is a *local* maximum. Go to step 4.
  - If $Hf(x_0)$ is indefinite, then x_0 is a saddle point (that is, it isn't a local minimum not maximum). Repeat the procedure with the other critical points (if any).
  - If $Hf(x_0)$ is positive semidefinite then we cannot determine yet whether or not x_0 is a local minimum. Go to step 4.
  - If $Hf(x_0)$ is negative semidefinite then we cannot determine yet whether or not x_0 is a local maximum. Go to step 4.
- **Step 4.** Study if f is convex or concave in A. For it, study if A is a convex set. If not, you can look for a convex subset. Then, consider the Hessian matrix at any point, $H(x)$, and then use Algorithm 1. Then:
  - If f convex on A, x_0 is a critical point and $x_0 \in A$, then x_0 is a *global* minimum of f in A. Moreover, if $Hf(x)$ is positive definite for all $x \in A$, then x_0 is the *only* global minimum of f in A.

– If f concave on A, x_0 is a critical point and $x_0 \in A$, then x_0 is a *global* maximum of f in A. Moreover, if $Hf(x)$ is negative definite for all $x \in A$, then x_0 is the *only* global maximum of f in A.

Example 41. *Let us assume that the cost function of a firm is $C(x,y) = (x-2)^4 + (x-y)^2 + 300$, where x, y are two amounts of raw materials. To determine the minimum cost of the company, we need to find the global minimum of C over the set of points where the variables can take on values, that is, the set $A = \{(x,y) \in \mathbb{R}^2 \,/\, x \geq 0, y \geq 0\}$.*

Since we are looking for the minimums of the function over A, we want to know if the function is convex over the set A and, accordingly, the first thing we must do is determine whether or not A is convex. We can observe that A is the intersection of two hyperplanes, $x \geq 0$, $y \geq 0$, so A is a convex set. We calculate the critical points of C:

$$
\begin{aligned}
C'_x(x,y) &= 4(x-2)^3 + 2(x-y) = 0 \\
C'_x(x,y)(x,y) &= 2(x-y) \;\; = 0
\end{aligned}
$$

The only critical point of C is $(x,y) = (2,2)$, which is a point in A. In the event that the critical point did not belong to set A, we would have concluded that C does not have any minimum in A.

Now we calculate the Hessian matrix and then replace the critical point:

$$
Hf(x,y) = \begin{pmatrix} 12(x-2)^2 + 2 & -2 \\ -2 & 2 \end{pmatrix} \Rightarrow Hf(2,2) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix},
$$

This matrix is positive semidefinite. So we can't say if $(2,2)$ is a local optimum of C (in A) or not (we can only say it isn't a local maximum). To study it, we must study the convexity of the function in A (which we already showed it is a convex set). For it, we study the sign of

$HC(x,y)$. Its minors are $H_1 = 12(x-2)^2 + 2 > 0$ and $H_2 = 24(x-2)^2 \geq 0$. So $HC(x,y)$ is positive semidefinite in $\mathbb{R}^2$ and C is a convex function in $\mathbb{R}^2$, which means that it is also a convex function on A (since A is included in $\mathbb{R}^2$). So, $(2,2)$ is a global minimum of C in A.

4 Applications to Linear Programming

Definition 42. Linear Programming *is a part of mathematics that studies the solution of linear optimisation problems, that is, problems with the following features:*

1. *The objective is to find the global minimums or maximums (known as* optimal solutions*) of a function (known as the* objective function*) on a set (known as* feasible region*).*
2. *The objective function is linear.*
3. *The feasible region is determined by one or several linear equations or inequations (known as* constraints*).*
4. *Variables are all greater or equal to zero.*

The feasible region of a linear problem can be generally described as the set of points $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ that verify a set of conditions of the following type:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\quad \vdots \qquad\quad & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\geq b_m \\ x_1 \geq 0,\ x_2 \geq 0,\ \ldots,\ x_n \geq 0 & \end{aligned}$$

As we confirmed in example 15, these set are intersection of hyperplanes and semi-spaces in $\mathbb{R}^n$, and, accordingly, convex sets in $\mathbb{R}^n$. so we can conclude that:

Property 43. *The feasible region of a linear problem is always a convex set.*

The objective of a linear problem is to find the global maximums and minimums of a linear function over the feasible region. Since linear functions are simultaneously convex and concave in $\mathbb{R}^n$ and given that the feasible region is a convex set of $\mathbb{R}^n$, the above-mentioned properties 39 and 40 are verified. So:

Property 44. *A linear problem has 0, 1 or infinite optimal solutions. If $x_1, x_2, \ldots, x_n$ are optimal solutions of a linear problem, then all the points of its convex envelope are also optimal solutions.*

The Simplex algorithm is the solution method that we will use in Linear Programming. If, by applying this method, we obtain more than one optimum, the foregoing result will enable us to assure that in that case there will be infinite optimal solutions that we arrive at by calculating the points of the convex envelope of the known optimal solutions.

This is of particular interest from an economic standpoint, since each time that we obtain an optimal solution we are determining new variable combination possibilities in order to reach the best possible value of the function. For example, let us assume that f is a cost function dependent on the capital and work variables and that, in solving the problem, we obtain more than one global maximum. Then each one of these represents a new method for combining the capital and work values so that they reach the minimum possible cost. Mathematically, each of these solutions is also as valid but, in real economic situations, there could be differences among them with respect to items that were not incorporated in the mathematical model. Specifically, let us assume that two solutions were found for the problem one of which means that company workers must work 9 hours per day and another one where, for the same cost, they would only have to work 7 hours a day. Evidently, one of these will be more beneficial for the staff's morale than the other one. In any case, it is always interesting to learn about other solutions since it could happen that in the future the set of possible values is reduced, eliminating some of these alternative optimums. For example, if our problem is to find the best transport route between different cities, it is always of interest to have alternative routes in the event that one of the routes is closed due to snow, an accident, etc.

Example 45. *Given a linear problem*

$$\begin{array}{ll} \max & f(x_1, x_2) = x_1 + x_2 \\ \\ s.\ a & -x_1 + x_2 \leq 3 \\ & \quad x_1 + x_2 \leq 5 \\ & x_1 \geq 0,\ x_2 \geq 0 \end{array}$$

let's suppose that two solutions $(x_1, x_2) = (1, 4)$ *and* $(x_1, x_2) = (5, 0)$ *are known. Applying property 44 we reach the conclusion that there are infinite solutions for the linear problem which are precisely the points of the convex envelope of the two last given optimums:*

$$CE[(1,4),(5,0)] = \left\{ \lambda(1,4) + (1-\lambda)(5,0) \;\Big/\; \lambda \in [0,1] \right\}$$

This implies that any point calculated in this manner is also an optimum of the linear problem. For example,

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

is also an optimal solution to the problem.