



SAPIENZA
UNIVERSITÀ DI ROMA

Convex Optimization

Exercises 1 — Preliminaries

Worked exercises on convex sets, functions of several variables, convex and concave functions and quadratic functions: convexity is always proved from the definition.

Fabio Furini

fabiofurini.github.io

Department of Computer, Control and Management Engineering — Sapienza University of Rome

Online version and downloadable material:
fabiofurini.github.io/convex-optimization/
github.com/fabiofurini/convex-optimization

Academic year 2026–2027

Contents

1	Convex sets	2
2	Functions of several variables	14
3	Convex and concave functions	20
4	Quadratic functions	33



Convex sets

Exercise 1.1 — Convex sets (E1)

Consider the following two points in $\mathbb{R}^2$:

$$\mathbf{p} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} 7 \\ 1 \end{pmatrix}$$

and the following three points in $\mathbb{R}^2$:

$$\mathbf{q}^{(1)} = \begin{pmatrix} 2 \\ \frac{7}{2} \end{pmatrix}, \quad \mathbf{q}^{(2)} = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \quad \text{and} \quad \mathbf{q}^{(3)} = \begin{pmatrix} 9 \\ 0 \end{pmatrix}.$$

Questions

1. Compute the convex combinations of $\mathbf{p}$ and $\mathbf{w}$ with $\lambda = \frac{1}{3}$, $\lambda = \frac{1}{2}$ and $\lambda = \frac{2}{3}$.
2. Draw the points $\mathbf{p}$ and $\mathbf{w}$, the three convex combinations and the set of all convex combinations of $\mathbf{p}$ and $\mathbf{w}$.
3. Determine which of the points $\mathbf{q}^{(1)}$, $\mathbf{q}^{(2)}$ and $\mathbf{q}^{(3)}$ are convex combinations of $\mathbf{p}$ and $\mathbf{w}$.

Solution

1. For every $\lambda \in [0, 1]$, the convex combination of $\mathbf{p}$ and $\mathbf{w}$ with the scalar λ is the point:

$$\lambda \begin{pmatrix} 1 \\ 4 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} 7 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda + 7 - 7\lambda \\ 4\lambda + 1 - \lambda \end{pmatrix} = \begin{pmatrix} 7 - 6\lambda \\ 1 + 3\lambda \end{pmatrix}.$$

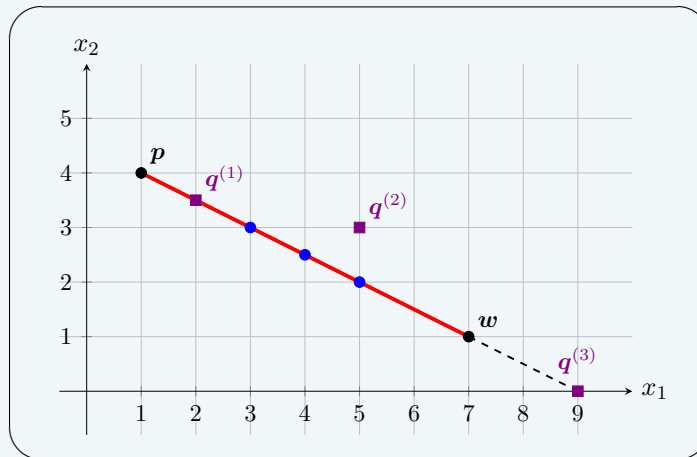
Therefore, we have:

$$\lambda = \frac{1}{3} : \quad \begin{pmatrix} \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 7 \\ \frac{1}{3} \cdot 4 + \frac{2}{3} \cdot 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

$$\lambda = \frac{1}{2} : \quad \begin{pmatrix} \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 7 \\ \frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ \frac{5}{2} \end{pmatrix}$$

$$\lambda = \frac{2}{3} : \quad \begin{pmatrix} \frac{2}{3} \cdot 1 + \frac{1}{3} \cdot 7 \\ \frac{2}{3} \cdot 4 + \frac{1}{3} \cdot 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

2. The set of all convex combinations of $\mathbf{p}$ and $\mathbf{w}$ is the segment joining the two points (red). The three convex combinations of the previous question are the blue points: with $\lambda = \frac{1}{3}$ the point is closer to $\mathbf{w}$, with $\lambda = \frac{2}{3}$ it is closer to $\mathbf{p}$.



3. A point $\mathbf{q} \in \mathbb{R}^2$ is a convex combination of $\mathbf{p}$ and $\mathbf{w}$ if and only if there exists $\lambda \in [0, 1]$ such that

$$\begin{pmatrix} 7 - 6\lambda \\ 1 + 3\lambda \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

The first equation gives $\lambda = \frac{7-q_1}{6}$; then we check the second equation and the condition $\lambda \in [0, 1]$.

- For $\mathbf{q}^{(1)}$, the first equation gives $\lambda = \frac{7-2}{6} = \frac{5}{6}$ and the second equation holds: $1 + 3 \cdot \frac{5}{6} = \frac{7}{2}$. Since $\frac{5}{6} \in [0, 1]$, the point $\mathbf{q}^{(1)}$ is the convex combination of $\mathbf{p}$ and $\mathbf{w}$ with $\lambda = \frac{5}{6}$.
- For $\mathbf{q}^{(2)}$, the first equation gives $\lambda = \frac{7-5}{6} = \frac{1}{3}$, but $1 + 3 \cdot \frac{1}{3} = 2 \neq 3$. The point $\mathbf{q}^{(2)}$ is not even on the line through $\mathbf{p}$ and $\mathbf{w}$: it is not a convex combination of $\mathbf{p}$ and $\mathbf{w}$.
- For $\mathbf{q}^{(3)}$, the first equation gives $\lambda = \frac{7-9}{6} = -\frac{1}{3}$ and the second equation holds: $1 + 3 \cdot (-\frac{1}{3}) = 0$. The point $\mathbf{q}^{(3)}$ is on the line through $\mathbf{p}$ and $\mathbf{w}$ (dashed in the figure), but $-\frac{1}{3} \notin [0, 1]$: it is not a convex combination of $\mathbf{p}$ and $\mathbf{w}$.

Exercise 1.2 — Convex sets (E2)

Consider the following three points in $\mathbb{R}^2$:

$$\mathbf{p}^{(1)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{p}^{(2)} = \begin{pmatrix} 7 \\ 2 \end{pmatrix} \quad \text{and} \quad \mathbf{p}^{(3)} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

and the following two points in $\mathbb{R}^2$:

$$\mathbf{q}^{(1)} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} \quad \text{and} \quad \mathbf{q}^{(2)} = \begin{pmatrix} 6 \\ 5 \end{pmatrix}.$$

Questions

1. Compute the convex combination of $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(3)}$ with $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{1}{4}$ and $\lambda_3 = \frac{1}{4}$.
2. Compute the convex combination of $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(3)}$ with $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$.
3. Show that $\mathbf{q}^{(1)}$ is a convex combination of $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(3)}$, and find the scalars.
4. Show that $\mathbf{q}^{(2)}$ is not a convex combination of $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(3)}$.
5. Draw the convex hull of $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(3)}$ and all the points of the previous questions.

Solution

1. The convex combination with $\lambda_1 = \frac{1}{2}$, $\lambda_2 = \frac{1}{4}$ and $\lambda_3 = \frac{1}{4}$ (non-negative scalars with sum 1) is the point:

$$\frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 7 \\ 2 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \frac{7}{4} + \frac{3}{4} \\ \frac{1}{2} + \frac{1}{2} + \frac{3}{2} \end{pmatrix} = \begin{pmatrix} 3 \\ \frac{5}{2} \end{pmatrix}.$$

2. The convex combination with $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$ (non-negative scalars with sum 1) is the point:

$$\frac{1}{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 7 \\ 2 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \begin{pmatrix} \frac{1+7+3}{3} \\ \frac{1+2+6}{3} \end{pmatrix} = \begin{pmatrix} \frac{11}{3} \\ 3 \end{pmatrix}.$$

3. A point $\mathbf{q} \in \mathbb{R}^2$ is a convex combination of the three points if and only if there exist $\lambda_1, \lambda_2, \lambda_3 \geq 0$ such that

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 &= 1 \\ \lambda_1 + 7\lambda_2 + 3\lambda_3 &= q_1 \\ \lambda_1 + 2\lambda_2 + 6\lambda_3 &= q_2 \end{cases}$$

Substituting $\lambda_1 = 1 - \lambda_2 - \lambda_3$ in the last two equations, we obtain the system

$$\begin{cases} 6\lambda_2 + 2\lambda_3 = q_1 - 1 \\ \lambda_2 + 5\lambda_3 = q_2 - 1 \end{cases}$$

whose determinant is $6 \cdot 5 - 2 \cdot 1 = 28 \neq 0$: the scalars are unique. For $\mathbf{q}^{(1)}$ we have $q_1 - 1 = 2$ and $q_2 - 1 = 3$, hence

$$\lambda_2 = \frac{2 \cdot 5 - 2 \cdot 3}{28} = \frac{1}{7}, \quad \lambda_3 = \frac{6 \cdot 3 - 2 \cdot 1}{28} = \frac{4}{7}, \quad \lambda_1 = 1 - \frac{1}{7} - \frac{4}{7} = \frac{2}{7}.$$

The three scalars are non-negative and their sum is 1, so $\mathbf{q}^{(1)}$ is a convex combination of the three points:

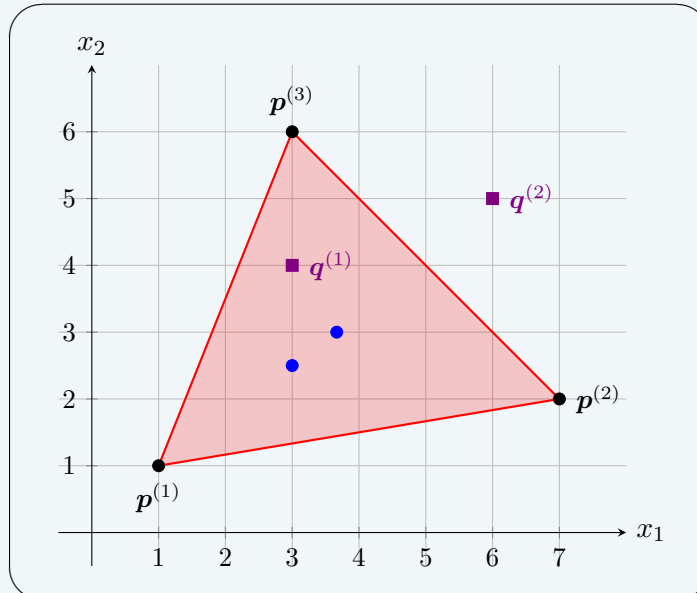
$$\frac{2}{7} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{1}{7} \begin{pmatrix} 7 \\ 2 \end{pmatrix} + \frac{4}{7} \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \begin{pmatrix} \frac{2+7+12}{7} \\ \frac{2+2+24}{7} \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

4. For $\mathbf{q}^{(2)}$ we have $q_1 - 1 = 5$ and $q_2 - 1 = 4$, hence

$$\lambda_2 = \frac{5 \cdot 5 - 2 \cdot 4}{28} = \frac{17}{28}, \quad \lambda_3 = \frac{6 \cdot 4 - 5 \cdot 1}{28} = \frac{19}{28}, \quad \lambda_1 = 1 - \frac{17}{28} - \frac{19}{28} = -\frac{2}{7}.$$

These are the only scalars with sum 1 that give $\mathbf{q}^{(2)}$, and $\lambda_1 < 0$: the point $\mathbf{q}^{(2)}$ is not a convex combination of the three points.

5. The convex hull of the three points is the triangle with vertices $\mathbf{p}^{(1)}$, $\mathbf{p}^{(2)}$ and $\mathbf{p}^{(3)}$ (red). The points of the first two questions are blue, $\mathbf{q}^{(1)}$ (inside the triangle) and $\mathbf{q}^{(2)}$ (outside the triangle) are violet.



Exercise 1.3 — Convex sets (E3)

Consider the following set:

$$F = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1 + 2x_2 \leq 8, \quad 3x_1 + x_2 \leq 9, \quad x_1 \geq 0, \quad x_2 \geq 0 \right\}$$

and the following two points in F and scalar:

$$\mathbf{q}^{(1)} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \quad \mathbf{q}^{(2)} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \quad \text{and} \quad \lambda = \frac{1}{3}.$$

Questions

1. Prove that F is a convex set.
2. Find the vertices of F and draw the set.
3. Compute the convex combination of $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$ with the scalar λ and verify that it belongs to F .

Solution

1. Each of the four inequalities that define F has the form $a_1 x_1 + a_2 x_2 \leq b$, with

$$(a_1, a_2, b) \in \{ (1, 2, 8), (3, 1, 9), (-1, 0, 0), (0, -1, 0) \}$$

(the last two are $x_1 \geq 0$ and $x_2 \geq 0$). For every pair of points $\mathbf{p}, \mathbf{w} \in F$ and every $\lambda \in [0, 1]$, for each of the four inequalities we have:

$$\begin{aligned} & a_1 (\lambda p_1 + (1 - \lambda)w_1) + a_2 (\lambda p_2 + (1 - \lambda)w_2) \\ &= \lambda (a_1 p_1 + a_2 p_2) + (1 - \lambda) (a_1 w_1 + a_2 w_2) \\ &\leq \lambda b + (1 - \lambda) b \\ &\quad (\text{since } \mathbf{p}, \mathbf{w} \in F, \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\ &= b \end{aligned}$$

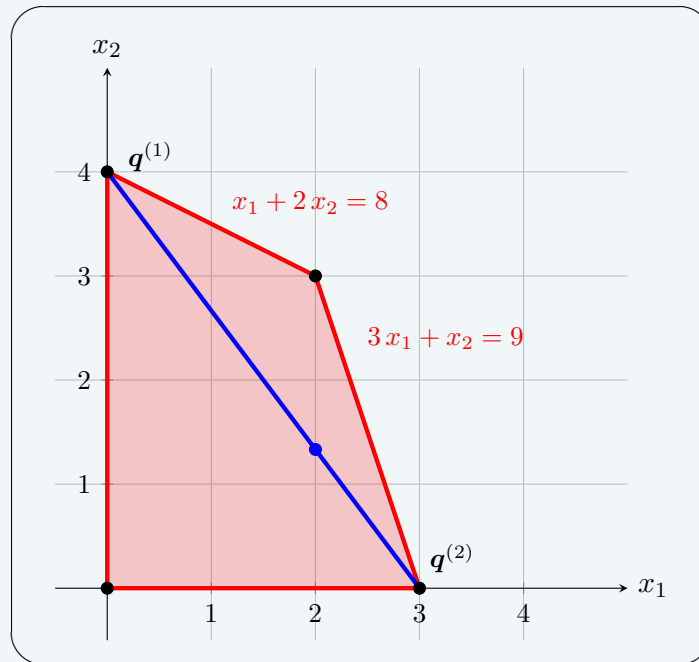
Therefore, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in F$, and the set is convex.

The same computation shows that each of the four sets defined by one inequality (a half-plane) is convex: F is their intersection, which is convex since the intersection of convex sets is convex (chapter «Convex sets»).

2. The vertices of F are the origin, the points where the lines $x_1 + 2x_2 = 8$ and $3x_1 + x_2 = 9$ meet the axes inside F , namely $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$, and the intersection point of the two lines:

$$\begin{cases} x_1 + 2x_2 = 8 \\ 3x_1 + x_2 = 9 \end{cases} \implies \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{8 \cdot 1 - 2 \cdot 9}{1 \cdot 1 - 2 \cdot 3} \\ \frac{1 \cdot 9 - 8 \cdot 3}{1 \cdot 1 - 2 \cdot 3} \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

The set F is the quadrilateral with vertices $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$ (red). The figure also shows the segment joining $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$ (blue) and the convex combination of the next question.



3. The convex combination of $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$ with $\lambda = \frac{1}{3}$ is the point:

$$\frac{1}{3} \begin{pmatrix} 0 \\ 4 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{4}{3} \end{pmatrix}.$$

It belongs to F , as guaranteed by the first question:

$$2 + 2 \cdot \frac{4}{3} = \frac{14}{3} \leq 8, \quad 3 \cdot 2 + \frac{4}{3} = \frac{22}{3} \leq 9, \quad 2 \geq 0, \quad \frac{4}{3} \geq 0.$$

Exercise 1.4 — Convex sets (E4)

Consider the following symmetric matrix

$$\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

and the following set:

$$F = \left\{ \mathbf{x} \in \mathbb{R}^2 : \underbrace{\sqrt{\sum_{i=1}^2 \sum_{j=1}^2 q_{ij} x_i x_j}}_{\sqrt{\mathbf{x}' \mathbf{Q} \mathbf{x}}} \leq 1 \right\}.$$

Questions

1. Show that $\mathbf{x}' \mathbf{Q} \mathbf{x} = x_1^2 + (x_1 - x_2)^2 + x_2^2$ for every $\mathbf{x} \in \mathbb{R}^2$.
2. Prove that F is a convex set.
3. Determine whether the points $\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$ belong to F .
4. Draw the set F .

Solution

1. For every $\mathbf{x} \in \mathbb{R}^2$ we have:

$$\begin{aligned} & (x_1 \ x_2) \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (2x_1 - x_2)x_1 + (-x_1 + 2x_2)x_2 \\ &= 2x_1^2 - 2x_1x_2 + 2x_2^2 \\ &= x_1^2 + (x_1^2 - 2x_1x_2 + x_2^2) + x_2^2 \\ &= x_1^2 + (x_1 - x_2)^2 + x_2^2. \end{aligned}$$

Therefore, for every $\mathbf{x} \in \mathbb{R}^2$, the number $\sqrt{\mathbf{x}' \mathbf{Q} \mathbf{x}}$ is the Euclidean norm of a vector of $\mathbb{R}^3$:

$$\sqrt{\mathbf{x}' \mathbf{Q} \mathbf{x}} = \|\mathbf{u}(\mathbf{x})\|_2, \quad \text{where} \quad \mathbf{u}(\mathbf{x}) = \begin{pmatrix} x_1 \\ x_1 - x_2 \\ x_2 \end{pmatrix} \in \mathbb{R}^3.$$

The vector $\mathbf{u}(\mathbf{x})$ depends linearly on $\mathbf{x}$: for every $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have $\mathbf{u}(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) = \lambda \mathbf{u}(\mathbf{p}) + (1 - \lambda) \mathbf{u}(\mathbf{w})$.

2. For every pair of points $\mathbf{p}, \mathbf{w} \in F$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 & \sqrt{(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})' \mathbf{Q} (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})} \\
 &= \|\lambda \mathbf{u}(\mathbf{p}) + (1 - \lambda) \mathbf{u}(\mathbf{w})\|_2 \\
 &\quad \text{(by the first question)} \\
 &\leq \|\lambda \mathbf{u}(\mathbf{p})\|_2 + \|(1 - \lambda) \mathbf{u}(\mathbf{w})\|_2 \\
 &\quad \text{(by the triangle inequality)} \\
 &= |\lambda| \|\mathbf{u}(\mathbf{p})\|_2 + |1 - \lambda| \|\mathbf{u}(\mathbf{w})\|_2 \\
 &\quad \text{(by the absolute homogeneity)} \\
 &= \lambda \|\mathbf{u}(\mathbf{p})\|_2 + (1 - \lambda) \|\mathbf{u}(\mathbf{w})\|_2 \\
 &\quad \text{(since } \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\
 &= \lambda \sqrt{\mathbf{p}' \mathbf{Q} \mathbf{p}} + (1 - \lambda) \sqrt{\mathbf{w}' \mathbf{Q} \mathbf{w}} \\
 &\quad \text{(by the first question)} \\
 &\leq \lambda \cdot 1 + (1 - \lambda) \cdot 1 \\
 &\quad \text{(since } \mathbf{p}, \mathbf{w} \in F, \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\
 &= 1
 \end{aligned}$$

Therefore, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in F$, and the set is convex.

3. We have:

$$\begin{aligned}
 \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} : & \quad \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2} - \frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1}{2} \leq 1 \quad \implies \quad \text{it is in } F, \\
 \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} : & \quad \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2} + \frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2 = \frac{3}{2} > 1 \quad \implies \quad \text{it is not in } F.
 \end{aligned}$$

(Since the square root is increasing, $\sqrt{\mathbf{x}' \mathbf{Q} \mathbf{x}} \leq 1$ if and only if $\mathbf{x}' \mathbf{Q} \mathbf{x} \leq 1$.)

4. The boundary of F is given by the points with

$$2x_1^2 - 2x_1x_2 + 2x_2^2 = 1.$$

We express x_2 as a function of x_1 . Writing the previous equation as

$$2x_2^2 - 2x_1x_2 + (2x_1^2 - 1) = 0,$$

with $a = 2$, $b = -2x_1$ and $c = 2x_1^2 - 1$ we have $x_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, hence:

$$x_2 = \frac{2x_1 \pm \sqrt{4x_1^2 - 8(2x_1^2 - 1)}}{4} = \frac{2x_1 \pm \sqrt{8 - 12x_1^2}}{4} = \frac{x_1 \pm \sqrt{2 - 3x_1^2}}{2}.$$

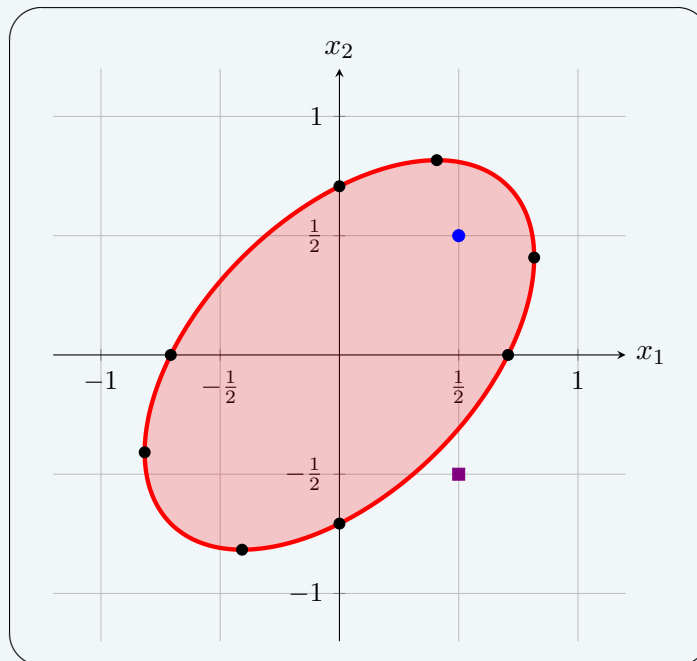
These values exist if and only if

$$2 - 3x_1^2 \geq 0 \iff x_1 \in \left[-\sqrt{\frac{2}{3}}, \sqrt{\frac{2}{3}}\right].$$

Some reference points of the boundary:

- if $x_2 = 0$, then $2x_1^2 = 1$, so $x_1 = \pm\frac{1}{\sqrt{2}}$; if $x_1 = 0$, then $2x_2^2 = 1$, so $x_2 = \pm\frac{1}{\sqrt{2}}$;
- if $x_1 = \pm\sqrt{\frac{2}{3}}$, then $2 - 3x_1^2 = 0$ and $x_2 = \frac{x_1}{2} = \pm\frac{1}{\sqrt{6}}$; in the same way, if $x_2 = \pm\sqrt{\frac{2}{3}}$, then $x_1 = \pm\frac{1}{\sqrt{6}}$.

The set F is the region inside (and on the boundary of) an ellipse whose axes are the lines $x_2 = x_1$ and $x_2 = -x_1$ (red). The figure shows the reference points (black), the point $(\frac{1}{2}, \frac{1}{2})$ that belongs to F (blue) and the point $(\frac{1}{2}, -\frac{1}{2})$ that does not belong to F (violet).



Exercise 1.5 — Convex sets (E5)

Consider the following sets:

$$A = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 1 \leq x_1^2 + x_2^2 \leq 4 \right\},$$

$$B_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 0 \leq x_1 \leq 4, 0 \leq x_2 \leq 1 \right\},$$

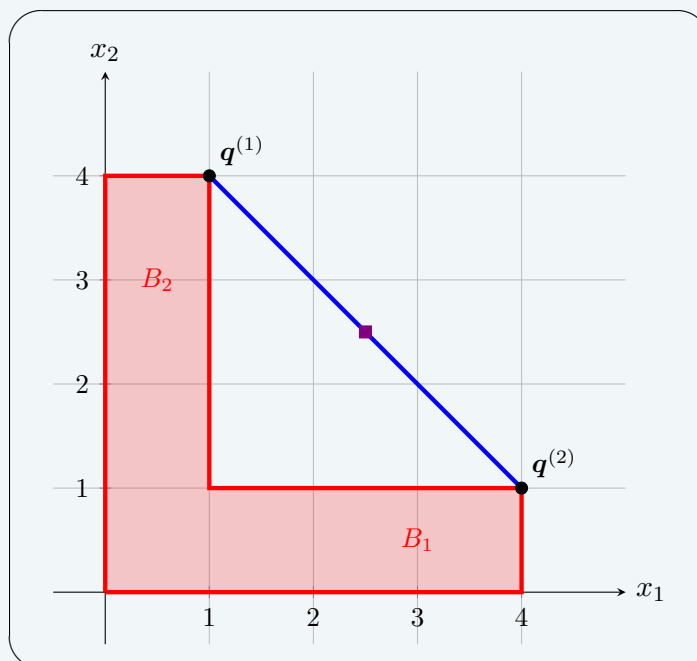
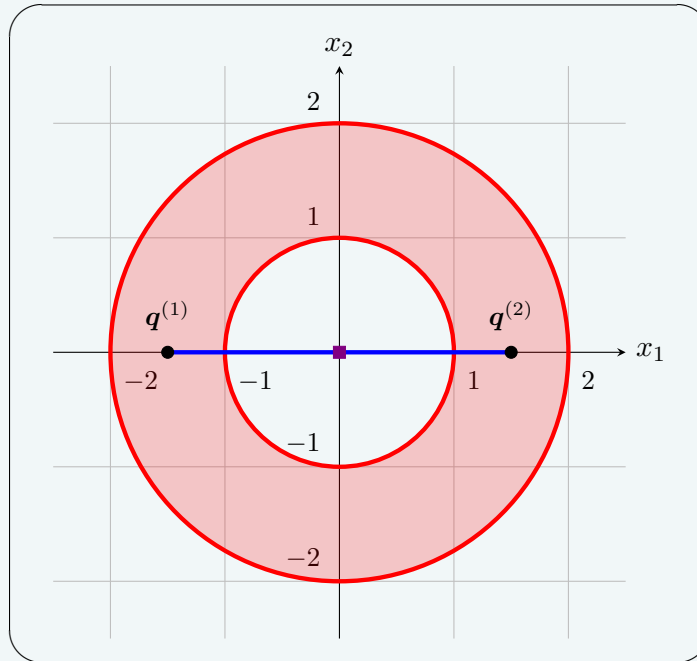
$$B_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 4 \right\} \quad \text{and} \quad B = B_1 \cup B_2.$$

Questions

1. Draw the sets A and B .
2. Show that A is not a convex set.
3. Prove that B_1 and B_2 are convex sets.
4. Show that B is not a convex set.

Solution

1. The set A is the ring between the circle of radius 1 and the circle of radius 2 centered at the origin (red, left). The set B is the L-shaped union of the rectangles B_1 and B_2 (red, right). The figures also show the segments used in the next questions (blue), whose midpoint (violet) does not belong to the set.



2. To show that a set is not convex, it is enough to find two points of the set and one scalar $\lambda \in [0, 1]$ whose convex combination does not belong to the set. Take

$$\mathbf{q}^{(1)} = \begin{pmatrix} -\frac{3}{2} \\ 0 \end{pmatrix}, \quad \mathbf{q}^{(2)} = \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix} \quad \text{and} \quad \lambda = \frac{1}{2}.$$

Both points belong to A , since $(\pm \frac{3}{2})^2 + 0^2 = \frac{9}{4} \in [1, 4]$. Their convex combination is

$$\frac{1}{2} \begin{pmatrix} -\frac{3}{2} \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \frac{3}{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \text{with} \quad 0^2 + 0^2 = 0 < 1.$$

The convex combination does not belong to A : the set A is not convex.

3. Let $k \in \{1, 2\}$. The set B_k is defined by the bounds $0 \leq x_1 \leq u_1$ and $0 \leq x_2 \leq u_2$, with $(u_1, u_2) = (4, 1)$ for B_1 and $(u_1, u_2) = (1, 4)$ for B_2 . For every pair of points $\mathbf{p}, \mathbf{w} \in B_k$ and every $\lambda \in [0, 1]$, for each coordinate $i \in \{1, 2\}$ we have:

$$\begin{aligned} \lambda p_i + (1 - \lambda)w_i &\in [\lambda \cdot 0 + (1 - \lambda) \cdot 0, \lambda u_i + (1 - \lambda) u_i] \\ &\quad (\text{since } p_i, w_i \in [0, u_i], \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\ &= [0, u_i] \end{aligned}$$

Therefore, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in B_k$, and the set B_k is convex.

4. Take

$$\mathbf{q}^{(1)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \in B_2, \quad \mathbf{q}^{(2)} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \in B_1 \quad \text{and} \quad \lambda = \frac{1}{2}.$$

Both points belong to B . Their convex combination is

$$\frac{1}{2} \begin{pmatrix} 1 \\ 4 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{2} \\ \frac{5}{2} \end{pmatrix},$$

which does not belong to B_1 (since $x_2 = \frac{5}{2} > 1$) nor to B_2 (since $x_1 = \frac{5}{2} > 1$), hence it does not belong to B : the set B is not convex.

The intersection of convex sets is always convex (chapter «Convex sets»), while the union of two convex sets, as $B = B_1 \cup B_2$, may not be convex.

Functions of several variables

Exercise 2.1 — Functions of several variables (E1)

Consider the following real-valued function:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^3 + 2x_1x_2^2 - 3x_2$$

Questions

1. Compute the value of the function at the points $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$.
2. Compute the gradient of the function and evaluate it at the two points.
3. Compute the Hessian of the function and evaluate it at the two points.

Solution

1. Substituting the coordinates of the points:

$$f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = 1^3 + 2 \cdot 1 \cdot 1^2 - 3 \cdot 1 = 1 + 2 - 3 = 0$$

$$f\left(\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right) = (-1)^3 + 2 \cdot (-1) \cdot 2^2 - 3 \cdot 2 = -1 - 8 - 6 = -15$$

2. The partial derivatives are $\frac{\partial f(\mathbf{x})}{\partial x_1} = 3x_1^2 + 2x_2^2$ and $\frac{\partial f(\mathbf{x})}{\partial x_2} = 4x_1x_2 - 3$, so the gradient is:

$$\nabla f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 3x_1^2 + 2x_2^2 \\ 4x_1x_2 - 3 \end{pmatrix}$$

At the two points:

$$\nabla f \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 3+2 \\ 4-3 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}, \quad \nabla f \left(\begin{pmatrix} -1 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} 3+8 \\ -8-3 \end{pmatrix} = \begin{pmatrix} 11 \\ -11 \end{pmatrix}$$

3. Differentiating the two partial derivatives once more:

$$\nabla^2 f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 6x_1 & 4x_2 \\ 4x_2 & 4x_1 \end{pmatrix}$$

which is a symmetric matrix depending on the point. At the two points:

$$\nabla^2 f \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 6 & 4 \\ 4 & 4 \end{pmatrix}, \quad \nabla^2 f \left(\begin{pmatrix} -1 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} -6 & 8 \\ 8 & -4 \end{pmatrix}$$

Exercise 2.2 — Functions of several variables (E2)

Consider the following real-valued function:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^4 - 2x_1^2x_2 + x_2^2$$

Questions

1. Show that $f(\mathbf{x}) = (x_1^2 - x_2)^2$.
2. Compute the gradient and the Hessian of the function.
3. Evaluate the gradient and the Hessian at the points $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.
4. Find all the points at which the gradient is the zero vector, and the value of the function at these points.

Solution

1. Expanding the square:

$$(x_1^2 - x_2)^2 = x_1^4 - 2x_1^2x_2 + x_2^2 = f(\mathbf{x})$$

2. The partial derivatives are $\frac{\partial f(\mathbf{x})}{\partial x_1} = 4x_1^3 - 4x_1x_2$ and $\frac{\partial f(\mathbf{x})}{\partial x_2} = -2x_1^2 + 2x_2$, so:

$$\nabla f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 4x_1^3 - 4x_1x_2 \\ -2x_1^2 + 2x_2 \end{pmatrix} = \begin{pmatrix} 4x_1(x_1^2 - x_2) \\ -2(x_1^2 - x_2) \end{pmatrix}$$

and

$$\nabla^2 f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 12x_1^2 - 4x_2 & -4x_1 \\ -4x_1 & 2 \end{pmatrix}$$

3. At the point $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ we have $x_1^2 - x_2 = 0$, and:

$$\nabla f \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \nabla^2 f \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 12 - 4 & -4 \\ -4 & 2 \end{pmatrix} = \begin{pmatrix} 8 & -4 \\ -4 & 2 \end{pmatrix}$$

- At the point $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ we have $x_1^2 - x_2 = 1$, and:

$$\nabla f \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4 \cdot 1 \cdot 1 \\ -2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \quad \nabla^2 f \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 12 & -4 \\ -4 & 2 \end{pmatrix}$$

4. The second component of the gradient is zero if and only if $x_2 = x_1^2$, and in this case also the first component $4x_1(x_1^2 - x_2)$ is zero. Therefore, the gradient is the zero vector at all the points of the parabola:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_2 = x_1^2 \right\}$$

(for example $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$), which is an infinite set of points. At all these points the value of the function is $f(\mathbf{x}) = (x_1^2 - x_2)^2 = 0$: this parabola is the level set $L(0)$ of the function.

Exercise 2.3 — Functions of several variables (E3)

Consider the following real-valued function of three variables:

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto x_1^2 x_3 + x_2 x_3^2 - 2x_1 x_2 + x_3$$

Questions

1. Compute the value of the function at the points $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.
2. Compute the gradient of the function and evaluate it at the two points.
3. Compute the Hessian of the function and evaluate it at the two points.

Solution

1. Substituting the coordinates of the points:

$$f\left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = 1^2 \cdot 2 + 0 \cdot 2^2 - 2 \cdot 1 \cdot 0 + 2 = 2 + 0 - 0 + 2 = 4$$

$$f\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right) = 0^2 \cdot (-1) + 1 \cdot (-1)^2 - 2 \cdot 0 \cdot 1 + (-1) = 0 + 1 - 0 - 1 = 0$$

2. The three partial derivatives are:

$$\frac{\partial f(\mathbf{x})}{\partial x_1} = 2x_1x_3 - 2x_2, \quad \frac{\partial f(\mathbf{x})}{\partial x_2} = x_3^2 - 2x_1, \quad \frac{\partial f(\mathbf{x})}{\partial x_3} = x_1^2 + 2x_2x_3 + 1$$

so the gradient is:

$$\nabla f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 2x_1x_3 - 2x_2 \\ x_3^2 - 2x_1 \\ x_1^2 + 2x_2x_3 + 1 \end{pmatrix}$$

At the two points:

$$\nabla f\left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 4 - 0 \\ 4 - 2 \\ 1 + 0 + 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix}, \quad \nabla f\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 0 - 2 \\ 1 - 0 \\ 0 - 2 + 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ -1 \end{pmatrix}$$

3. Differentiating each partial derivative with respect to x_1 , x_2 and x_3 :

$$\nabla^2 f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 2x_3 & -2 & 2x_1 \\ -2 & 0 & 2x_3 \\ 2x_1 & 2x_3 & 2x_2 \end{pmatrix}$$

which is a symmetric matrix depending on the point. At the two points:

$$\nabla^2 f\left(\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 4 & -2 & 2 \\ -2 & 0 & 4 \\ 2 & 4 & 0 \end{pmatrix}, \quad \nabla^2 f\left(\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} -2 & -2 & 0 \\ -2 & 0 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

Exercise 2.4 — Functions of several variables (E4)

Consider the following real-valued function:

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f: \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_2 - x_1^2$$

Questions

1. Compute the gradient and the Hessian of the function.
2. Describe the level sets $L(k)$ of the function for every $k \in \mathbb{R}$.
3. Evaluate the gradient at the point $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ of the level set $L(0)$ and at the point $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ of the level set $L(2)$.
4. Draw the level sets $L(0)$ and $L(2)$ and the two gradients of the previous question, and verify that each gradient is perpendicular to the level set through its point.

Solution

1. The partial derivatives are $\frac{\partial f(\mathbf{x})}{\partial x_1} = -2x_1$ and $\frac{\partial f(\mathbf{x})}{\partial x_2} = 1$, so:

$$\nabla f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix}, \quad \nabla^2 f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$$

The Hessian is constant. The second component of the gradient is always 1, so the gradient is never the zero vector.

2. For every $k \in \mathbb{R}$:

$$L(k) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_2 - x_1^2 = k \right\} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_2 = x_1^2 + k \right\}$$

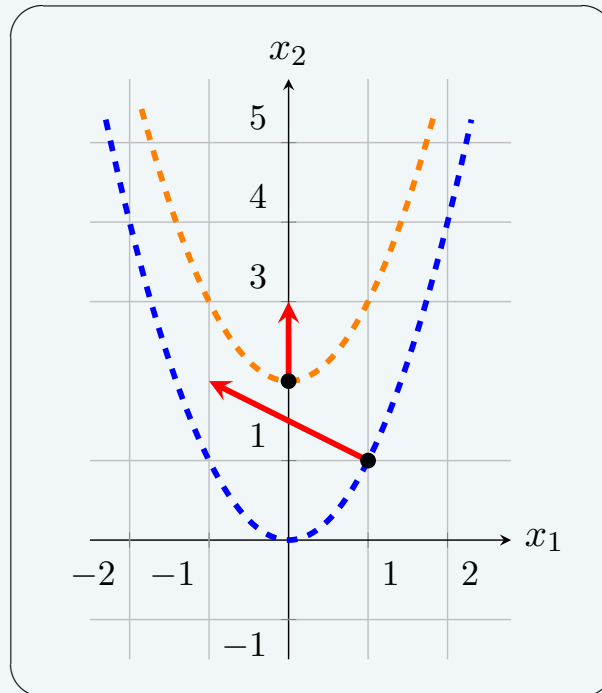
Every level set is a parabola with vertex $\begin{pmatrix} 0 \\ k \end{pmatrix}$, obtained by translating the parabola $x_2 = x_1^2$ vertically by k .

3. The point $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ belongs to $L(0)$ since $1 - 1^2 = 0$, and the point $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ belongs to $L(2)$ since $2 - 0^2 = 2$. The gradients are:

$$\nabla f \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \nabla f \left(\begin{pmatrix} 0 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

4. We draw the two level sets with the gradients, and then we verify the perpendicularity with the tangent directions.

Graphical solution: the figure shows the level sets $L(0)$ (in blue) and $L(2)$ (in orange) and the two gradients (in red), drawn from their points:



The tangent direction of a level set $x_2 = x_1^2 + k$ at a point with first coordinate x_1 is $\begin{pmatrix} 1 \\ 2x_1 \end{pmatrix}$ (the derivative of $x_1^2 + k$ is $2x_1$):

- at $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ the tangent direction is $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = -2 + 2 = 0$;
- at $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ the tangent direction is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$.

Therefore, both gradients are perpendicular to the level set through their point, and they point toward the level sets with larger values of k .

Convex and concave functions

In the exercises of this chapter convexity and concavity are proved only with the definitions of the chapter «Convex and concave functions», using the following two inequalities.

For every $a, b \in \mathbb{R}$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 (\lambda a + (1 - \lambda)b)^2 &= \lambda^2 a^2 + 2\lambda(1 - \lambda)ab + (1 - \lambda)^2 b^2 \\
 &= \lambda a^2 + (1 - \lambda)b^2 - \lambda(1 - \lambda)a^2 + 2\lambda(1 - \lambda)ab - \lambda(1 - \lambda)b^2 \\
 &= \lambda a^2 + (1 - \lambda)b^2 - \lambda(1 - \lambda)(a^2 - 2ab + b^2) \\
 &= \lambda a^2 + (1 - \lambda)b^2 - \underbrace{\lambda(1 - \lambda)(a - b)^2}_{\geq 0} \\
 &\leq \lambda a^2 + (1 - \lambda)b^2 \quad (\text{since } \lambda \geq 0, 1 - \lambda \geq 0 \text{ and } (a - b)^2 \geq 0)
 \end{aligned}$$

For every $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ and every $\alpha \in \mathbb{R}$, the Euclidean norm $\|\mathbf{u}\|_2 = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}$ satisfies:

$$\begin{aligned}
 \|\mathbf{u} + \mathbf{v}\|_2 &\leq \|\mathbf{u}\|_2 + \|\mathbf{v}\|_2 && \text{(triangle inequality)} \\
 \|\alpha \mathbf{u}\|_2 &= |\alpha| \|\mathbf{u}\|_2 && \text{(absolute homogeneity)}
 \end{aligned}$$

Exercise 3.1 — Convex and concave functions (E1)

Consider the following real-valued functions:

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f : x \mapsto x^2 - 4x + 5 \quad \text{and} \quad g : \mathbb{R} \rightarrow \mathbb{R}, \quad g : x \mapsto -x^2 + 4x - 5$$

and the following two points of $\mathbb{R}$ and scalar:

$$p = -1, \quad w = 4 \quad \text{and} \quad \lambda = \frac{2}{5}.$$

Questions

1. Prove that the function f is convex in its domain.
2. Compute both sides of the inequality of the definition of convex function for f with the given p , w and λ .
3. Draw the graph of f , the points $(p, f(p))$ and $(w, f(w))$ and the segment joining them.
4. Prove that the function g is concave in its domain.

Solution

1. For every $x \in \mathbb{R}$ we have $f(x) = (x - 2)^2 + 1$. For every pair of points $p, w \in \mathbb{R}$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 f(\lambda p + (1 - \lambda)w) &= (\lambda p + (1 - \lambda)w - 2)^2 + 1 \\
 &= (\lambda(p - 2) + (1 - \lambda)(w - 2))^2 + 1 \\
 &\leq \lambda(p - 2)^2 + (1 - \lambda)(w - 2)^2 + 1 \\
 &\quad (\text{since } (\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2) \\
 &= \lambda((p - 2)^2 + 1) + (1 - \lambda)((w - 2)^2 + 1) \\
 &= \lambda f(p) + (1 - \lambda)f(w)
 \end{aligned}$$

Therefore,

$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

By nature, the domain $\mathbb{R}$ is a convex set. Accordingly, the function is convex in its domain.

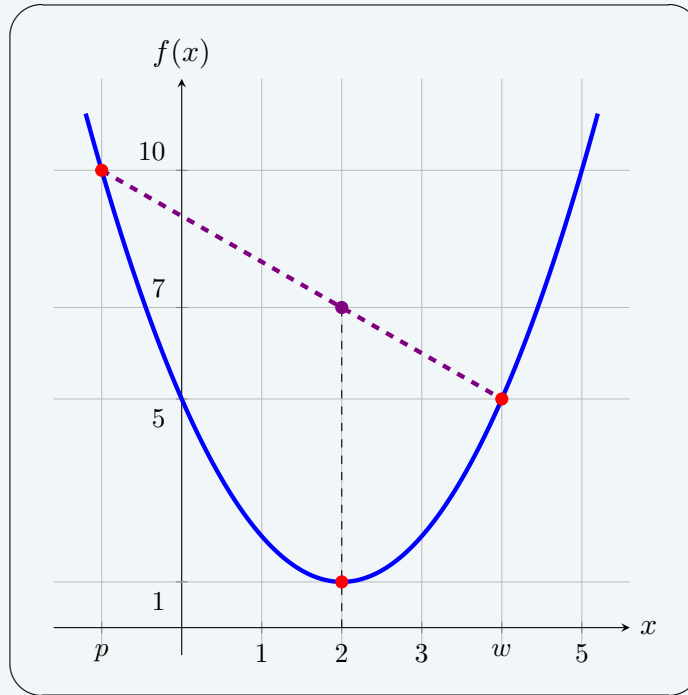
2. With $p = -1$, $w = 4$ and $\lambda = \frac{2}{5}$ we have

$$\lambda p + (1 - \lambda)w = \frac{2}{5} \cdot (-1) + \frac{3}{5} \cdot 4 = 2, \quad f(-1) = 10, \quad f(4) = 5,$$

and the two sides of the inequality are:

$$f(2) = 1 \leq \frac{2}{5} f(-1) + \frac{3}{5} f(4) = \frac{2}{5} \cdot 10 + \frac{3}{5} \cdot 5 = 7.$$

3. In the figure, the segment joining $(p, f(p))$ and $(w, f(w))$ (violet) lies above the graph of f (blue). The red point is $(2, f(2))$ and the violet point is $(2, \frac{2}{5}f(p) + \frac{3}{5}f(w))$.



4. For every $x \in \mathbb{R}$ we have $g(x) = -(x - 2)^2 - 1$. For every pair of points $p, w \in \mathbb{R}$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 g(\lambda p + (1 - \lambda)w) &= -(\lambda(p - 2) + (1 - \lambda)(w - 2))^2 - 1 \\
 &\geq -\lambda(p - 2)^2 - (1 - \lambda)(w - 2)^2 - 1 \\
 &\quad (\text{since } (\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2) \\
 &= \lambda(-(p - 2)^2 - 1) + (1 - \lambda)(-(w - 2)^2 - 1) \\
 &= \lambda g(p) + (1 - \lambda)g(w)
 \end{aligned}$$

Therefore,

$$g(\lambda p + (1 - \lambda)w) \geq \lambda g(p) + (1 - \lambda)g(w)$$

By nature, the domain $\mathbb{R}$ is a convex set. Accordingly, the function is concave in its domain.

The same conclusion follows from the first question, since $g = -f$ (chapter «Convex and concave functions»: f is convex if and only if $-f$ is concave).

Exercise 3.2 — Convex and concave functions (E2)

Consider the following real-valued function:

$$f : [-1, 4] \rightarrow \mathbb{R}, \quad f : x \mapsto \max \left\{ 1 - x, \frac{1}{2}x, 2x - 3 \right\}$$

and the following two points of $[-1, 4]$ and scalar:

$$p = -1, \quad w = 4 \quad \text{and} \quad \lambda = \frac{3}{5}.$$

Questions

1. Compute $f(-1)$, $f(\frac{2}{3})$, $f(2)$ and $f(4)$, and draw the graph of the function.
2. Prove that the function is convex in its domain.
3. Compute both sides of the inequality of the definition of convex function with the given p , w and λ .
4. Show that the function is not concave in its domain.

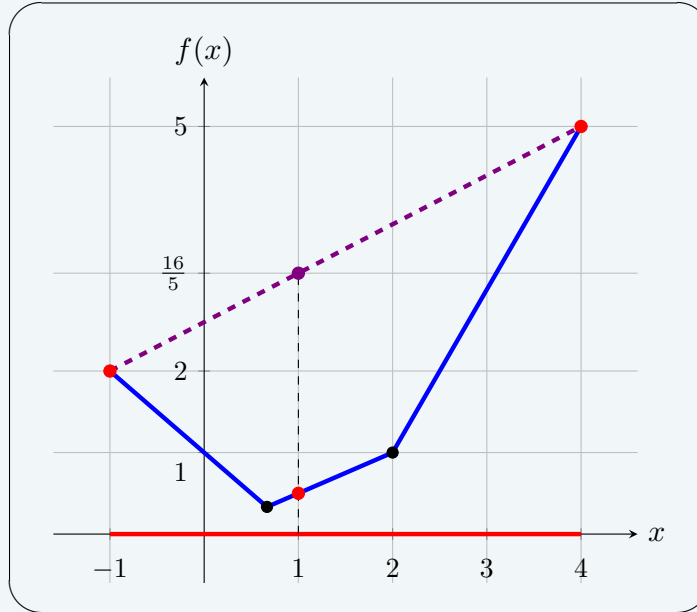
Solution

1. We have:

$$f(-1) = \max \left\{ 2, -\frac{1}{2}, -5 \right\} = 2, \quad f\left(\frac{2}{3}\right) = \max \left\{ \frac{1}{3}, \frac{1}{3}, -\frac{5}{3} \right\} = \frac{1}{3},$$

$$f(2) = \max \left\{ -1, 1, 1 \right\} = 1, \quad f(4) = \max \left\{ -3, 2, 5 \right\} = 5.$$

The lines $1 - x$ and $\frac{1}{2}x$ meet at $x = \frac{2}{3}$, and the lines $\frac{1}{2}x$ and $2x - 3$ meet at $x = 2$. The graph of f (blue) is made of three segments: $f(x) = 1 - x$ on $[-1, \frac{2}{3}]$, $f(x) = \frac{1}{2}x$ on $[\frac{2}{3}, 2]$ and $f(x) = 2x - 3$ on $[2, 4]$. The domain is red; the violet segment and the points are those of the third question.



2. For every $x \in [-1, 4]$, $f(x)$ is the largest of the three values $h_1(x) = 1 - x$, $h_2(x) = \frac{1}{2}x$ and $h_3(x) = 2x - 3$, and each of them has the form $h_k(x) = a_k x + b_k$. For every pair of points $p, w \in [-1, 4]$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 f(\lambda p + (1 - \lambda)w) &= h_k(\lambda p + (1 - \lambda)w) \\
 &\quad \text{(for an index } k \in \{1, 2, 3\} \text{ attaining the maximum)} \\
 &= a_k (\lambda p + (1 - \lambda)w) + b_k \\
 &= \lambda (a_k p + b_k) + (1 - \lambda) (a_k w + b_k) \\
 &\quad \text{(since } \lambda + (1 - \lambda) = 1 \text{)} \\
 &\leq \lambda f(p) + (1 - \lambda)f(w) \\
 &\quad \text{(since } a_k p + b_k \leq f(p), a_k w + b_k \leq f(w), \lambda, 1 - \lambda \geq 0 \text{)}
 \end{aligned}$$

Therefore,

$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

The domain $[-1, 4]$ is an interval, hence a convex set. Accordingly, the function is convex in its domain.

The same conclusion follows from the property «the maximum of convex functions is convex» (chapter «Convex and concave functions»), since h_1 , h_2 and h_3 are linear.

3. With $p = -1$, $w = 4$ and $\lambda = \frac{3}{5}$ we have

$$\lambda p + (1 - \lambda)w = \frac{3}{5} \cdot (-1) + \frac{2}{5} \cdot 4 = 1, \quad f(1) = \max \left\{ 0, \frac{1}{2}, -1 \right\} = \frac{1}{2},$$

and the two sides of the inequality are:

$$f(1) = \frac{1}{2} \leq \frac{3}{5} f(-1) + \frac{2}{5} f(4) = \frac{3}{5} \cdot 2 + \frac{2}{5} \cdot 5 = \frac{16}{5}.$$

4. The function is concave if $f(\lambda p + (1 - \lambda)w) \geq \lambda f(p) + (1 - \lambda)f(w)$ for every p, w in the domain and every $\lambda \in [0, 1]$. With the points and the scalar of the previous question we have

$$f(1) = \frac{1}{2} < \frac{16}{5} = \frac{3}{5} f(-1) + \frac{2}{5} f(4) :$$

the inequality of the definition of concave function does not hold, so the function is not concave in its domain.

Exercise 3.3 — Convex and concave functions (E3)

Consider the following real-valued functions:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \underbrace{x_1^2 + x_2^2 - 9}_{\left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\|_2^2 - 9}$$

and

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \left\| \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right\|_2 + 1.$$

Questions

1. Prove that the function f is convex in its domain.
2. Prove that the function g is convex in its domain.
3. Prove that the set $\{ \mathbf{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 9 \}$ is convex.
4. Draw the graph of the function f .

Solution

1. For every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) &= (\lambda p_1 + (1 - \lambda)w_1)^2 + (\lambda p_2 + (1 - \lambda)w_2)^2 - 9 \\ &\leq \lambda p_1^2 + (1 - \lambda)w_1^2 + \lambda p_2^2 + (1 - \lambda)w_2^2 - 9 \\ &\quad (\text{since } (\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2) \end{aligned}$$

$$\begin{aligned}
&= \lambda (p_1^2 + p_2^2 - 9) + (1 - \lambda) (w_1^2 + w_2^2 - 9) \\
&= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})
\end{aligned}$$

Therefore,

$$f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is convex in its domain.

2. Let $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$. For every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
g(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_2 + 1 \\
&= \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2 + 1 \\
&\leq \|\lambda(\mathbf{p} - \mathbf{c})\|_2 + \|(1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2 + 1 \\
&\quad \text{(by the triangle inequality)} \\
&= |\lambda| \|\mathbf{p} - \mathbf{c}\|_2 + |1 - \lambda| \|\mathbf{w} - \mathbf{c}\|_2 + 1 \\
&\quad \text{(by the absolute homogeneity)} \\
&= \lambda \|\mathbf{p} - \mathbf{c}\|_2 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_2 + 1 \\
&\quad \text{(since } \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\
&= \lambda (\|\mathbf{p} - \mathbf{c}\|_2 + 1) + (1 - \lambda) (\|\mathbf{w} - \mathbf{c}\|_2 + 1) \\
&= \lambda g(\mathbf{p}) + (1 - \lambda) g(\mathbf{w})
\end{aligned}$$

Therefore,

$$g(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda g(\mathbf{p}) + (1 - \lambda) g(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is convex in its domain.

3. The set is $\{\mathbf{x} \in \mathbb{R}^2 : f(\mathbf{x}) \leq 0\}$. For every pair of points $\mathbf{p}$ and $\mathbf{w}$ in the set and for every $\lambda \in [0, 1]$, we have:

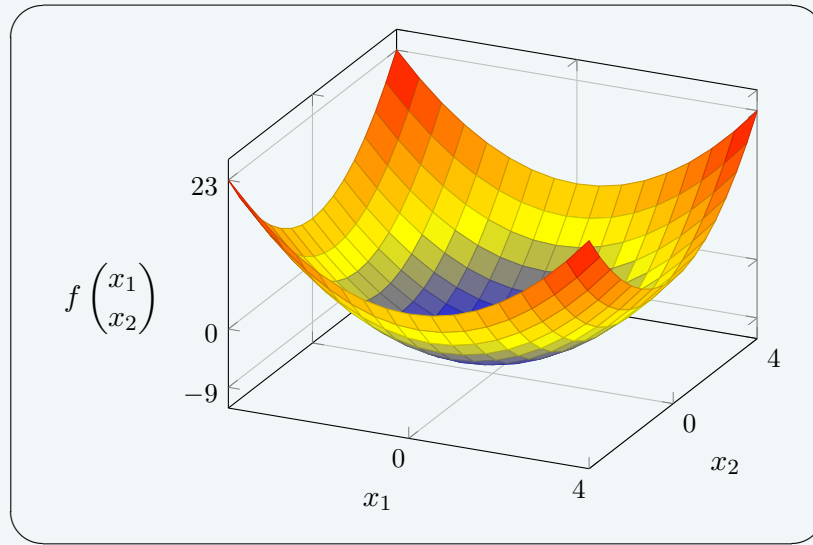
$$\begin{aligned}
f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &\leq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \\
&\quad \text{(by the first question)} \\
&\leq \lambda \cdot 0 + (1 - \lambda) \cdot 0 \\
&\quad \text{(since } f(\mathbf{p}) \leq 0, f(\mathbf{w}) \leq 0, \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\
&= 0
\end{aligned}$$

Therefore,

$$f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq 0$$

Accordingly, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}$ belongs to the set and the set is convex. It is the disk of radius 3 centered at the origin.

4. This is the graph of the function f in the domain $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in [-4, 4] \times [-4, 4]$; its lowest point is $(0, 0, f(0, 0)) = (0, 0, -9)$.



Exercise 3.4 — Convex and concave functions (E4)

Consider the following real-valued functions:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \underbrace{x_1 + x_2 - 1}_{(1 \quad 1) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 1}$$

and

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \min \{ x_1 + x_2 - 1, 2 - x_1 \}$$

and the following two points of $\mathbb{R}^2$ and scalar:

$$\mathbf{p} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad \text{and} \quad \lambda = \frac{1}{2}.$$

Questions

1. Prove that the function f is both convex and concave in its domain.
2. Draw the graph of the function f .
3. Prove that the function g is concave in its domain.
4. Compute both sides of the inequality of the definition of concave function for g with the given $\mathbf{p}$, $\mathbf{w}$ and λ , and show that g is not convex in its domain.

Solution

1. For every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= (\lambda p_1 + (1 - \lambda) w_1) + (\lambda p_2 + (1 - \lambda) w_2) - (\lambda + (1 - \lambda)) \\
 &= \lambda (p_1 + p_2 - 1) + (1 - \lambda) (w_1 + w_2 - 1) \\
 &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})
 \end{aligned}$$

Therefore, the inequalities of both definitions hold (with equality):

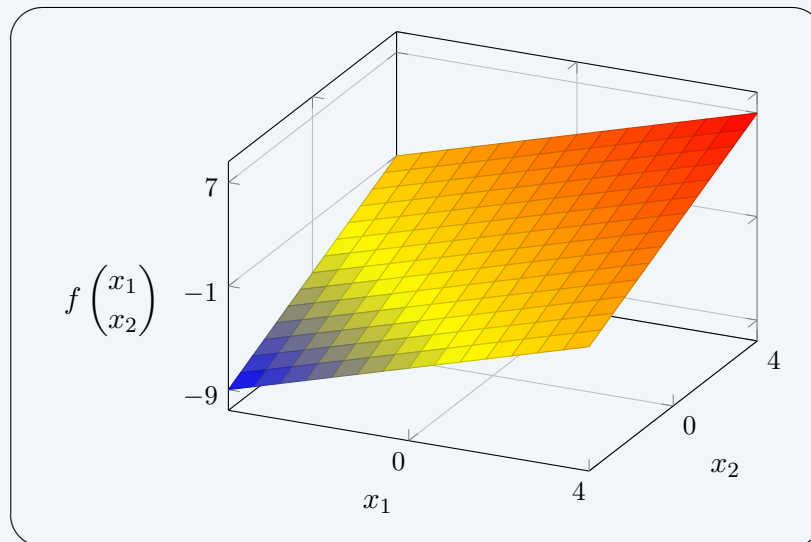
$$f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})$$

and

$$f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \geq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is both convex and concave in its domain.

2. This is the graph of the function f in the domain $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in [-4, 4] \times [-4, 4]$: a plane.



3. For every $\mathbf{x} \in \mathbb{R}^2$, $g(\mathbf{x})$ is the smallest of the two values $h_1(\mathbf{x}) = x_1 + x_2 - 1$ and $h_2(\mathbf{x}) = 2 - x_1$; as shown for f in the first question, each of them satisfies $h_k(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) = \lambda h_k(\mathbf{p}) + (1 - \lambda)h_k(\mathbf{w})$. For every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} g(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) &= h_k(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \\ &\quad (\text{for an index } k \in \{1, 2\} \text{ attaining the minimum}) \\ &= \lambda h_k(\mathbf{p}) + (1 - \lambda) h_k(\mathbf{w}) \\ &\quad (\text{as shown for } f \text{ in the first question}) \\ &\geq \lambda g(\mathbf{p}) + (1 - \lambda)g(\mathbf{w}) \\ &\quad (\text{since } h_k(\mathbf{p}) \geq g(\mathbf{p}), h_k(\mathbf{w}) \geq g(\mathbf{w}), \lambda, 1 - \lambda \geq 0) \end{aligned}$$

Therefore,

$$g(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \geq \lambda g(\mathbf{p}) + (1 - \lambda)g(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is concave in its domain.

The same conclusion follows from the property «the minimum of concave functions is concave» (chapter «Convex and concave functions»), since h_1 and h_2 are linear.

4. With $\mathbf{p} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$ and $\lambda = \frac{1}{2}$ we have

$$\lambda \mathbf{p} + (1 - \lambda)\mathbf{w} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad g(\mathbf{p}) = \min\{-1, 2\} = -1, \quad g(\mathbf{w}) = \min\{5, -2\} = -2,$$

and the two sides of the inequality are:

$$g\left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}\right) = \min\{2, 0\} = 0 \geq \frac{1}{2} g(\mathbf{p}) + \frac{1}{2} g(\mathbf{w}) = \frac{1}{2} \cdot (-1) + \frac{1}{2} \cdot (-2) = -\frac{3}{2}.$$

Since $0 > -\frac{3}{2}$, the inequality of the definition of convex function does not hold for these points and this scalar: the function g is not convex in its domain.

Exercise 3.5 — Convex and concave functions (E5)

Given $a, b \in \mathbb{R}$ with $a < b$, consider the following real-valued function:

$$f : [a, b] \rightarrow \mathbb{R}, \quad f : x \mapsto 2 + 9x - 6x^2 + x^3$$

whose domain is the interval $[a, b]$.

Questions

1. Prove that for every $p, w \in \mathbb{R}$ and every $\lambda \in [0, 1]$, with $m = \lambda p + (1 - \lambda)w$, we have

$$\lambda f(p) + (1 - \lambda)f(w) - f(m) = \lambda(1 - \lambda)(p - w)^2 (p + w + m - 6).$$

2. Prove that with $a = 2$ and $b = 4$ the function is convex.
3. Prove that with $a = -1$ and $b = 2$ the function is concave.
4. Show that with $a = -1$ and $b = 4$ the function is neither convex nor concave.
5. Draw the graph of the function with $a = -1$ and $b = 4$.

Solution

1. For every $p, w \in \mathbb{R}$ and every $\lambda \in [0, 1]$, let $d = p - w$, so that $p = w + d$ and $m = w + \lambda d$. We have:

$$\begin{aligned} \lambda f(p) + (1 - \lambda)f(w) - f(m) &= 2(\lambda + (1 - \lambda) - 1) + 9(\lambda p + (1 - \lambda)w - m) \\ &\quad - 6(\lambda p^2 + (1 - \lambda)w^2 - m^2) \\ &\quad + \lambda p^3 + (1 - \lambda)w^3 - m^3 \\ &= -6(\lambda p^2 + (1 - \lambda)w^2 - m^2) \\ &\quad + \lambda p^3 + (1 - \lambda)w^3 - m^3 \\ &\quad \text{(since } \lambda + (1 - \lambda) = 1 \\ &\quad \text{and } m = \lambda p + (1 - \lambda)w) \\ &= -6(\lambda w^2 + 2\lambda w d + \lambda d^2 + (1 - \lambda)w^2 \\ &\quad - w^2 - 2\lambda w d - \lambda^2 d^2) \\ &\quad + \lambda(w^3 + 3w^2 d + 3w d^2 + d^3) + (1 - \lambda)w^3 \\ &\quad - (w^3 + 3\lambda w^2 d + 3\lambda^2 w d^2 + \lambda^3 d^3) \\ &\quad \text{(expanding the powers, since } p = w + d \\ &\quad \text{and } m = w + \lambda d) \\ &= -6(\lambda - \lambda^2)d^2 + 3(\lambda - \lambda^2)w d^2 + (\lambda - \lambda^3)d^3 \\ &= \lambda(1 - \lambda)d^2(3w + (1 + \lambda)d - 6) \\ &= \lambda(1 - \lambda)d^2(p + w + m - 6) \\ &\quad \text{(since } 3w + (1 + \lambda)d = (w + d) + w + (w + \lambda d) \\ &\quad = p + w + m) \\ &= \lambda(1 - \lambda)(p - w)^2(p + w + m - 6) \end{aligned}$$

2. For every pair of points $p, w \in [2, 4]$ and every $\lambda \in [0, 1]$, with $m = \lambda p + (1 - \lambda)w \in$

$[2, 4]$, we have:

$$\begin{aligned} f(\lambda p + (1 - \lambda)w) &= \lambda f(p) + (1 - \lambda)f(w) - \lambda(1 - \lambda)(p - w)^2(p + w + m - 6) \\ &\quad \text{(by question 1)} \\ &\leq \lambda f(p) + (1 - \lambda)f(w) \\ &\quad \text{(since } \lambda(1 - \lambda)(p - w)^2 \geq 0, p, w, m \geq 2) \end{aligned}$$

Therefore,

$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

The domain $[2, 4]$ is an interval, hence a convex set. Accordingly, the function is convex in its domain.

3. For every pair of points $p, w \in [-1, 2]$ and every $\lambda \in [0, 1]$, with $m = \lambda p + (1 - \lambda)w \in [-1, 2]$, we have:

$$\begin{aligned} f(\lambda p + (1 - \lambda)w) &= \lambda f(p) + (1 - \lambda)f(w) - \lambda(1 - \lambda)(p - w)^2(p + w + m - 6) \\ &\quad \text{(by question 1)} \\ &\geq \lambda f(p) + (1 - \lambda)f(w) \\ &\quad \text{(since } \lambda(1 - \lambda)(p - w)^2 \geq 0, p, w, m \leq 2) \end{aligned}$$

Therefore,

$$f(\lambda p + (1 - \lambda)w) \geq \lambda f(p) + (1 - \lambda)f(w)$$

The domain $[-1, 2]$ is an interval, hence a convex set. Accordingly, the function is concave in its domain.

4. We have $f(0) = 2$, $f(1) = 6$, $f(2) = 4$, $f(3) = 2$ and $f(4) = 6$.

- With $p = 0$, $w = 2$ and $\lambda = \frac{1}{2}$ we have $\lambda p + (1 - \lambda)w = 1$ and

$$f(1) = 6 > \frac{1}{2}f(0) + \frac{1}{2}f(2) = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 4 = 3 :$$

the inequality of the definition of convex function does not hold, so the function is not convex in $[-1, 4]$.

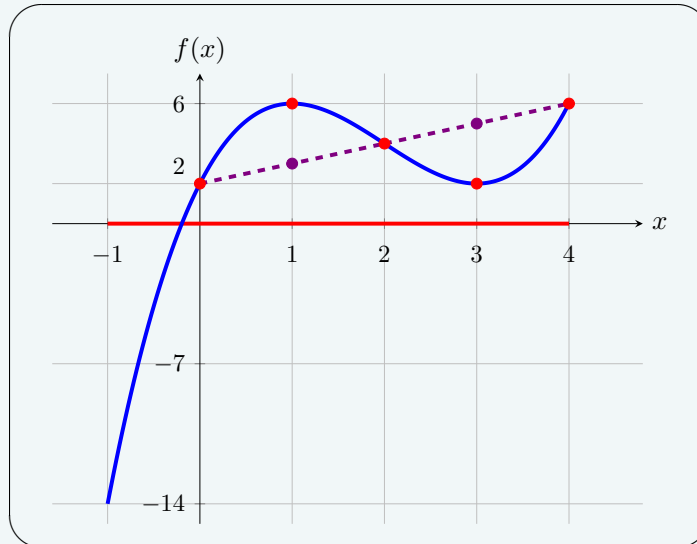
- With $p = 2$, $w = 4$ and $\lambda = \frac{1}{2}$ we have $\lambda p + (1 - \lambda)w = 3$ and

$$f(3) = 2 < \frac{1}{2}f(2) + \frac{1}{2}f(4) = \frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 6 = 5 :$$

the inequality of the definition of concave function does not hold, so the function is not concave in $[-1, 4]$.

The two differences agree with the first question: $\frac{1}{4} \cdot 4 \cdot (0 + 2 + 1 - 6) = -3 = 3 - 6$ and $\frac{1}{4} \cdot 4 \cdot (2 + 4 + 3 - 6) = 3 = 5 - 2$.

5. This is the graph of the function (blue) in the domain $[-1, 4]$ (red), with $f(-1) = -14$. The violet segments join the points of the previous question: the first one lies below the graph (on $[-1, 2]$ the function is concave), the second one lies above the graph (on $[2, 4]$ the function is convex).



Quadratic functions

In the exercises of this chapter, convexity is proved only with the definition of convex function of the chapter «Convex and concave functions»: the quadratic function is written as a sum of squares of linear functions, and the inequality

$$(\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2, \quad \forall a, b \in \mathbb{R}, \forall \lambda \in [0, 1]$$

of the same chapter is applied to each square.

Exercise 4.1 — Quadratic functions (E1)

Consider the following real-valued function:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + 2x_1x_2 + 5x_2^2$$

Questions

1. Write the function in the matrix form $f(\mathbf{x}) = \mathbf{x}'\mathbf{Q}\mathbf{x}$ with a symmetric matrix $\mathbf{Q}$.
2. Show that the non-symmetric matrix $\begin{pmatrix} 1 & 6 \\ -4 & 5 \end{pmatrix}$ defines the same function, and that replacing each off-diagonal pair with its average gives the matrix $\mathbf{Q}$ of the first question.
3. Compute the gradient and the Hessian of the function, and the gradient at the points $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$.
4. Show that $f(\mathbf{x}) = (x_1 + x_2)^2 + 4x_2^2$ and prove, from the definition, that the function is convex in its domain.
5. Draw the graph of the function.

Solution

1. The coefficient 2 of x_1x_2 is split equally between the two off-diagonal entries:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = (x_1 \ x_2) \underbrace{\begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix}}_{\mathbf{Q}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_1x_2 + x_2x_1 + 5x_2^2 = x_1^2 + 2x_1x_2 + 5x_2^2$$

2. With the non-symmetric matrix we have:

$$(x_1 \ x_2) \begin{pmatrix} 1 & 6 \\ -4 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 6x_1x_2 - 4x_2x_1 + 5x_2^2 = x_1^2 + 2x_1x_2 + 5x_2^2$$

which is the same function. Replacing the off-diagonal pair with its average:

$$\begin{pmatrix} 1 & \frac{6+(-4)}{2} \\ \frac{6+(-4)}{2} & 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} = \mathbf{Q}$$

3. The gradient is $2\mathbf{Q}\mathbf{x}$ and the Hessian is $2\mathbf{Q}$:

$$\nabla f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 + 2x_2 \\ 2x_1 + 10x_2 \end{pmatrix}, \quad \nabla^2 f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 2 & 2 \\ 2 & 10 \end{pmatrix}$$

At the two points:

$$\nabla f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ 12 \end{pmatrix}, \quad \nabla f\left(\begin{pmatrix} -1 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} -2 + 4 \\ -2 + 20 \end{pmatrix} = \begin{pmatrix} 2 \\ 18 \end{pmatrix}$$

4. Expanding the square:

$$(x_1 + x_2)^2 + 4x_2^2 = x_1^2 + 2x_1x_2 + x_2^2 + 4x_2^2 = x_1^2 + 2x_1x_2 + 5x_2^2 = f(\mathbf{x})$$

To prove that the function is convex in its domain, we observe that for every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

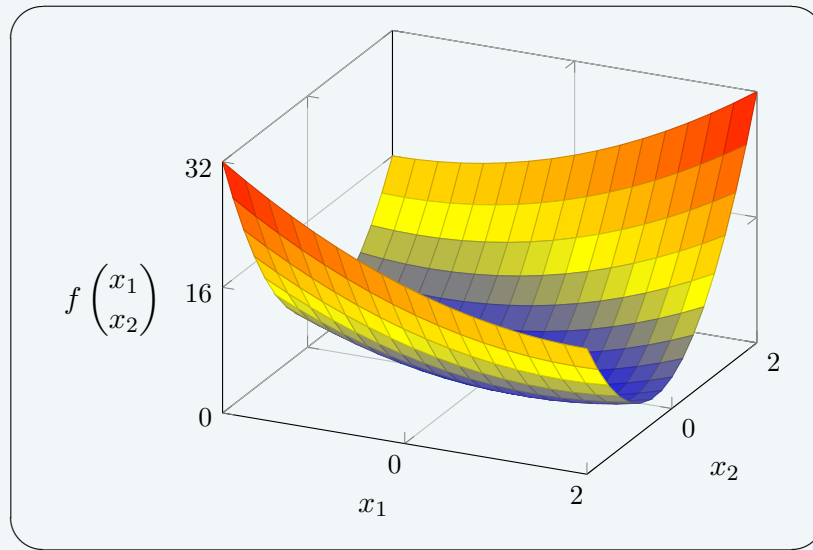
$$\begin{aligned} & f(\lambda\mathbf{p} + (1-\lambda)\mathbf{w}) \\ &= (\lambda p_1 + (1-\lambda)w_1 + \lambda p_2 + (1-\lambda)w_2)^2 + 4(\lambda p_2 + (1-\lambda)w_2)^2 \\ &= (\lambda(p_1 + p_2) + (1-\lambda)(w_1 + w_2))^2 + 4(\lambda p_2 + (1-\lambda)w_2)^2 \\ &\leq \lambda(p_1 + p_2)^2 + (1-\lambda)(w_1 + w_2)^2 + 4(\lambda p_2^2 + (1-\lambda)w_2^2) \\ &\quad (\text{since } (\lambda a + (1-\lambda)b)^2 \leq \lambda a^2 + (1-\lambda)b^2, \text{ twice}) \\ &= \lambda((p_1 + p_2)^2 + 4p_2^2) + (1-\lambda)((w_1 + w_2)^2 + 4w_2^2) \\ &= \lambda f(\mathbf{p}) + (1-\lambda)f(\mathbf{w}) \end{aligned}$$

Therefore,

$$f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is convex in its domain.

5. This is the graph of the function in the domain $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in [-2, 2] \times [-2, 2]$; its lowest point is $(0, 0, f(0, 0)) = (0, 0, 0)$ and its highest points are at $(2, 2)$ and $(-2, -2)$, where $f = 32$.



Exercise 4.2 — Quadratic functions (E2)

Consider the following real-valued function:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + 2x_2^2$$

Questions

1. Write the function in the matrix form $f(\mathbf{x}) = \mathbf{x}' \mathbf{Q} \mathbf{x}$ with a symmetric matrix $\mathbf{Q}$.
2. Compute the gradient of the function and its value at the four points $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} -1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.
3. Describe the level sets of the function for $k = 3$ and $k = 12$ (with their intersections with the axes) and draw them, together with the gradients of the previous question applied at their points.
4. Prove, from the definition, that the function is convex in its domain.

Solution

1. The function has no term x_1x_2 , so the off-diagonal entries are zero:

$$f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}}_Q \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + 2x_2^2$$

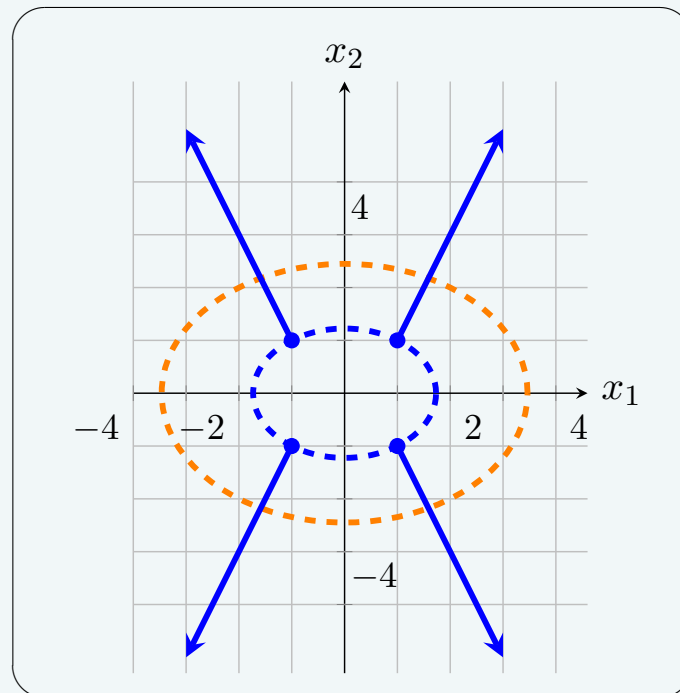
2. The gradient is $2Qx$:

$$\nabla f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 \\ 4x_2 \end{pmatrix}$$

At the four points:

$$\begin{aligned} \nabla f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) &= \begin{pmatrix} 2 \\ 4 \end{pmatrix}, & \nabla f\left(\begin{pmatrix} -1 \\ 1 \end{pmatrix}\right) &= \begin{pmatrix} -2 \\ 4 \end{pmatrix}, \\ \nabla f\left(\begin{pmatrix} -1 \\ -1 \end{pmatrix}\right) &= \begin{pmatrix} -2 \\ -4 \end{pmatrix}, & \nabla f\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) &= \begin{pmatrix} 2 \\ -4 \end{pmatrix} \end{aligned}$$

3. The level set for $k = 3$ is the ellipse $x_1^2 + 2x_2^2 = 3$, which meets the x_1 -axis at $x_1 = \pm\sqrt{3}$ and the x_2 -axis at $x_2 = \pm\sqrt{3/2}$; it contains the four points of the previous question, since $1 + 2 = 3$. The level set for $k = 12$ is the ellipse $x_1^2 + 2x_2^2 = 12$, which meets the x_1 -axis at $x_1 = \pm\sqrt{12} = \pm 2\sqrt{3}$ and the x_2 -axis at $x_2 = \pm\sqrt{6}$; it contains the four points $\begin{pmatrix} \pm 2 \\ \pm 2 \end{pmatrix}$, since $4 + 8 = 12$. In the figure, the level set $k = 3$ is blue, the level set $k = 12$ is orange and the gradients at the four points of the level set $k = 3$ are blue arrows: each gradient is perpendicular to the level set.



4. To prove that the function is convex in its domain, we observe that for every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= (\lambda p_1 + (1 - \lambda) w_1)^2 + 2(\lambda p_2 + (1 - \lambda) w_2)^2 \\ &\leq \lambda p_1^2 + (1 - \lambda) w_1^2 + 2(\lambda p_2^2 + (1 - \lambda) w_2^2) \\ &\quad (\text{since } (\lambda a + (1 - \lambda) b)^2 \leq \lambda a^2 + (1 - \lambda) b^2, \text{ twice}) \\ &= \lambda (p_1^2 + 2 p_2^2) + (1 - \lambda) (w_1^2 + 2 w_2^2) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Therefore,

$$f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is convex in its domain.

Exercise 4.3 — Quadratic functions (E3)

Consider the following real-valued function:

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_1 + 4$$

Questions

1. Write the function in the form $g(\mathbf{x}) = \mathbf{x}' \mathbf{Q} \mathbf{x} + \mathbf{c}' \mathbf{x} + d$ with a symmetric matrix $\mathbf{Q}$, a vector $\mathbf{c}$ and a scalar d .
2. Compute the gradient of the function and the point at which the gradient is the zero vector.
3. Show that $g(\mathbf{x}) = (x_1 - x_2 - 1)^2 + (x_2 - 1)^2 + 2$ and deduce that $g(\mathbf{x}) \geq 2$ for every $\mathbf{x} \in \mathbb{R}^2$, with equality only at the point of the previous question.
4. Describe the level sets of the function for $k = 3$ and $k = 6$ and draw them.
5. Prove, from the definition, that the function is convex in its domain.

Solution

1. The coefficient -2 of x_1x_2 is split equally between the two off-diagonal entries, and the linear part is $-2x_1 + 0x_2$:

$$\mathbf{Q} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad d = 4$$

Indeed:

$$(x_1 \ x_2) \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (-2 \ 0) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + 4 = x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_1 + 4$$

2. The gradient is $2\mathbf{Q}\mathbf{x} + \mathbf{c}$:

$$\nabla g \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 2x_1 - 2x_2 - 2 \\ -2x_1 + 4x_2 \end{pmatrix}$$

The gradient is the zero vector when:

$$\begin{cases} 2x_1 - 2x_2 - 2 = 0 \\ -2x_1 + 4x_2 = 0 \end{cases}$$

From the second equation $x_1 = 2x_2$; substituting in the first, $4x_2 - 2x_2 - 2 = 0$, that is $x_2 = 1$ and $x_1 = 2$. The point is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

3. Expanding the squares:

$$\begin{aligned} & (x_1 - x_2 - 1)^2 + (x_2 - 1)^2 + 2 \\ &= x_1^2 + x_2^2 + 1 - 2x_1x_2 - 2x_1 + 2x_2 + x_2^2 - 2x_2 + 1 + 2 \\ &= x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_1 + 4 = g(\mathbf{x}) \end{aligned}$$

Since the two squares are nonnegative, $g(\mathbf{x}) \geq 2$ for every $\mathbf{x} \in \mathbb{R}^2$. The equality holds only if both squares are zero, that is $x_2 = 1$ and $x_1 = x_2 + 1 = 2$: the point $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ of the previous question, where $g = 2$.

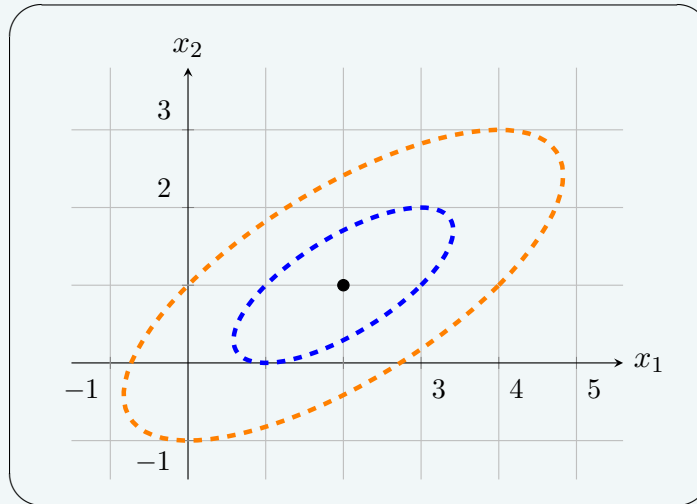
4. By the previous question, the level set for $k = 3$ is

$$\{\mathbf{x} \in \mathbb{R}^2 : (x_1 - x_2 - 1)^2 + (x_2 - 1)^2 = 1\}$$

an ellipse centered at $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ that contains, for instance, the points $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. The level set for $k = 6$ is

$$\{\mathbf{x} \in \mathbb{R}^2 : (x_1 - x_2 - 1)^2 + (x_2 - 1)^2 = 4\}$$

an ellipse with the same center that contains, for instance, the points $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$. In the figure, the level set $k = 3$ is blue, the level set $k = 6$ is orange and the center $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is black.



5. To prove that the function is convex in its domain, we observe that for every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 & g(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\
 &= (\lambda p_1 + (1 - \lambda)w_1 - \lambda p_2 - (1 - \lambda)w_2 - 1)^2 + (\lambda p_2 + (1 - \lambda)w_2 - 1)^2 + 2 \\
 &= (\lambda(p_1 - p_2 - 1) + (1 - \lambda)(w_1 - w_2 - 1))^2 + (\lambda(p_2 - 1) + (1 - \lambda)(w_2 - 1))^2 + 2 \\
 &\quad (\text{since } \lambda + (1 - \lambda) = 1) \\
 &\leq \lambda(p_1 - p_2 - 1)^2 + (1 - \lambda)(w_1 - w_2 - 1)^2 + \lambda(p_2 - 1)^2 + (1 - \lambda)(w_2 - 1)^2 + 2 \\
 &\quad (\text{since } (\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2, \text{ twice}) \\
 &= \lambda((p_1 - p_2 - 1)^2 + (p_2 - 1)^2 + 2) + (1 - \lambda)((w_1 - w_2 - 1)^2 + (w_2 - 1)^2 + 2) \\
 &\quad (\text{since } \lambda + (1 - \lambda) = 1) \\
 &= \lambda g(\mathbf{p}) + (1 - \lambda)g(\mathbf{w})
 \end{aligned}$$

Therefore,

$$g(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda g(\mathbf{p}) + (1 - \lambda)g(\mathbf{w})$$

By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is convex in its domain.

Exercise 4.4 — Quadratic functions (E4)

Consider the following real-valued functions:

$$f: \mathbb{R}^3 \rightarrow \mathbb{R}, \quad f: \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto (x_1 \ x_2 \ x_3) \begin{pmatrix} 1 & 4 & 0 \\ -2 & 2 & -6 \\ 0 & 4 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

and

$$h: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad h: \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + 4x_1x_2 + x_2^2$$

Questions

1. Find the symmetric matrix $\mathbf{Q}$ that defines the same function f , and write f as a polynomial in x_1, x_2, x_3 .
2. Compute the gradient and the Hessian of f , and the gradient at the point $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.
3. Show that $f(\mathbf{x}) = (x_1 + x_2)^2 + (x_2 - x_3)^2 + 2x_3^2$ and prove, from the definition, that f is convex in its domain.
4. Write h in the matrix form with a symmetric matrix and show that h is not convex in its domain, by finding two points and a value of λ for which the inequality of the definition of convex function does not hold.

Solution

1. Replacing each off-diagonal pair with its average:

$$\mathbf{Q} = \begin{pmatrix} 1 & \frac{4+(-2)}{2} & 0 \\ \frac{4+(-2)}{2} & 2 & \frac{-6+4}{2} \\ 0 & \frac{-6+4}{2} & 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 3 \end{pmatrix}$$

and the function is:

$$\begin{aligned} f(\mathbf{x}) &= x_1^2 + x_1x_2 + x_2x_1 + 2x_2^2 - x_2x_3 - x_3x_2 + 3x_3^2 \\ &= x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 - 2x_2x_3 \end{aligned}$$

2. The gradient is $2\mathbf{Q}\mathbf{x}$ and the Hessian is $2\mathbf{Q}$:

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 2x_1 + 2x_2 \\ 2x_1 + 4x_2 - 2x_3 \\ -2x_2 + 6x_3 \end{pmatrix}, \quad \nabla^2 f(\mathbf{x}) = \begin{pmatrix} 2 & 2 & 0 \\ 2 & 4 & -2 \\ 0 & -2 & 6 \end{pmatrix}$$

At the point:

$$\nabla f \left(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 2-2 \\ 2-4-2 \\ 2+6 \end{pmatrix} = \begin{pmatrix} 0 \\ -4 \\ 8 \end{pmatrix}$$

3. Expanding the squares:

$$\begin{aligned} & (x_1 + x_2)^2 + (x_2 - x_3)^2 + 2x_3^2 \\ &= x_1^2 + 2x_1x_2 + x_2^2 + x_2^2 - 2x_2x_3 + x_3^2 + 2x_3^2 \\ &= x_1^2 + 2x_2^2 + 3x_3^2 + 2x_1x_2 - 2x_2x_3 = f(\mathbf{x}) \end{aligned}$$

To prove that the function is convex in its domain, we observe that for every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^3$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} & f(\lambda\mathbf{p} + (1-\lambda)\mathbf{w}) \\ &= (\lambda(p_1 + p_2) + (1-\lambda)(w_1 + w_2))^2 + (\lambda(p_2 - p_3) + (1-\lambda)(w_2 - w_3))^2 \\ &\quad + 2(\lambda p_3 + (1-\lambda)w_3)^2 \\ &\leq \lambda(p_1 + p_2)^2 + (1-\lambda)(w_1 + w_2)^2 + \lambda(p_2 - p_3)^2 + (1-\lambda)(w_2 - w_3)^2 \\ &\quad + 2(\lambda p_3^2 + (1-\lambda)w_3^2) \\ &\quad (\text{since } (\lambda a + (1-\lambda)b)^2 \leq \lambda a^2 + (1-\lambda)b^2, \text{ three times}) \\ &= \lambda((p_1 + p_2)^2 + (p_2 - p_3)^2 + 2p_3^2) \\ &\quad + (1-\lambda)((w_1 + w_2)^2 + (w_2 - w_3)^2 + 2w_3^2) \\ &= \lambda f(\mathbf{p}) + (1-\lambda)f(\mathbf{w}) \end{aligned}$$

Therefore,

$$f(\lambda\mathbf{p} + (1-\lambda)\mathbf{w}) \leq \lambda f(\mathbf{p}) + (1-\lambda)f(\mathbf{w})$$

By nature, the domain $\mathbb{R}^3$ is a convex set. Accordingly, the function is convex in its domain.

4. The coefficient 4 of x_1x_2 is split equally between the two off-diagonal entries:

$$h \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

With $\mathbf{p} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $\lambda = \frac{1}{2}$, we have $\frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{w} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and:

$$h(\mathbf{p}) = 1 - 4 + 1 = -2, \quad h(\mathbf{w}) = 1 - 4 + 1 = -2, \quad h \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = 0$$

Therefore,

$$h\left(\frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{w}\right) = 0 > -2 = \frac{1}{2}h(\mathbf{p}) + \frac{1}{2}h(\mathbf{w})$$

The inequality of the definition does not hold, so the function h is not convex in its domain. In agreement with the chapter, the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ is not positive semidefinite, since $h(\mathbf{p}) = -2 < 0$.