



SAPIENZA
UNIVERSITÀ DI ROMA

Convex Optimization

Volume 1 — Preliminaries: convex sets, functions and convex functions

Convex combinations and convex sets, functions of several variables (gradient, Hessian, level sets), norms, convex and concave functions, quadratic functions and their convexity. With worked examples.

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
1.1	Convex combination	2
1.2	Convex sets	6
1.3	Examples of convex sets	7
1.4	Properties of convex sets	12
2	Functions of several variables	14
2.1	Functions of several variables	14
2.2	Gradient, Hessian and level sets	15
2.3	Examples of functions that are not quadratic	18
3	Convex and concave functions	20
3.1	Convex functions	20
3.2	Concave functions	21
3.3	Examples of convex and concave functions	22
3.4	Convex functions defined by norms	28
3.5	Properties of convex and concave functions	33
4	Quadratic functions	39
4.1	Quadratic functions	39
4.2	Gradient, Hessian and level sets	41
4.3	Convexity of quadratic functions	48
4.4	Quadratic functions with a linear part	51

Convex sets

1.1 Convex combination

Definition 1.1: convex combination

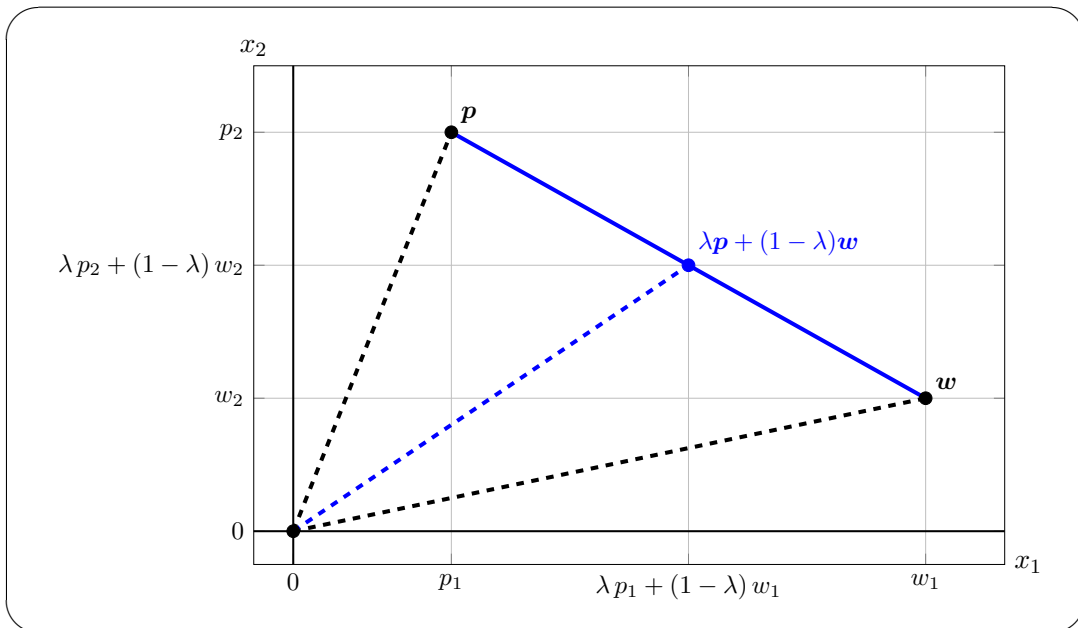
Given two points

$$\underbrace{\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}}_{\mathbf{p}} \in \mathbb{R}^n \quad \text{and} \quad \underbrace{\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}}_{\mathbf{w}} \in \mathbb{R}^n$$

their **convex combination** with a scalar $\lambda \in [0, 1]$ is the point

$$\underbrace{\begin{pmatrix} \lambda p_1 + (1 - \lambda)w_1 \\ \lambda p_2 + (1 - \lambda)w_2 \\ \vdots \\ \lambda p_n + (1 - \lambda)w_n \end{pmatrix}}_{\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}}$$

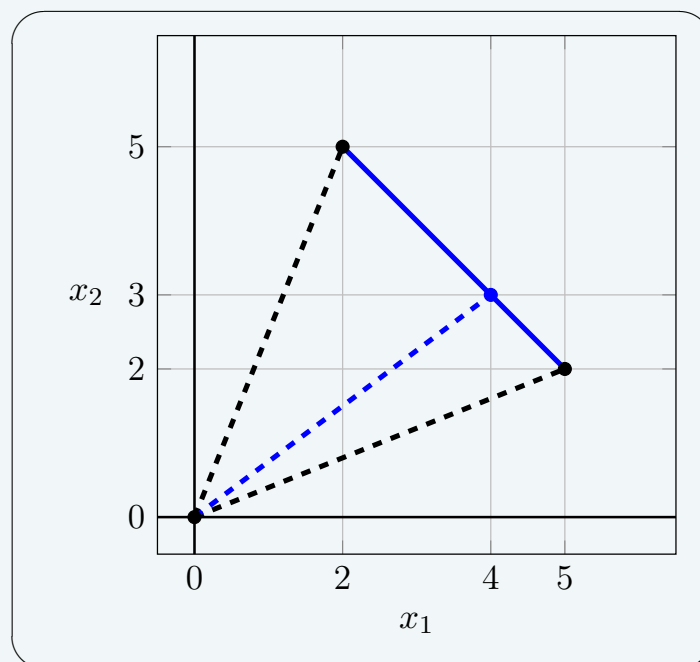
- The convex combinations for all values $\lambda \in [0, 1]$ are the points on the segment connecting the two points $\mathbf{p}$ and $\mathbf{w}$. In $\mathbb{R}^2$, the blue segment of the following figure:



Example 1.1: convex combination

Let us consider the points $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ and $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in $\mathbb{R}^2$. Their convex combination with $\lambda = \frac{1}{3}$ is the point:

$$\begin{pmatrix} \frac{1}{3} \cdot 2 + (1 - \frac{1}{3})5 \\ \frac{1}{3} \cdot 5 + (1 - \frac{1}{3})2 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$



Definition 1.2: convex combination of m points

Given m points

$$\underbrace{\begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ \vdots \\ p_n^{(1)} \end{pmatrix}}_{\mathbf{p}^{(1)}} \in \mathbb{R}^n \quad \underbrace{\begin{pmatrix} p_1^{(2)} \\ p_2^{(2)} \\ \vdots \\ p_n^{(2)} \end{pmatrix}}_{\mathbf{p}^{(2)}} \in \mathbb{R}^n \quad \dots \quad \underbrace{\begin{pmatrix} p_1^{(m)} \\ p_2^{(m)} \\ \vdots \\ p_n^{(m)} \end{pmatrix}}_{\mathbf{p}^{(m)}} \in \mathbb{R}^n$$

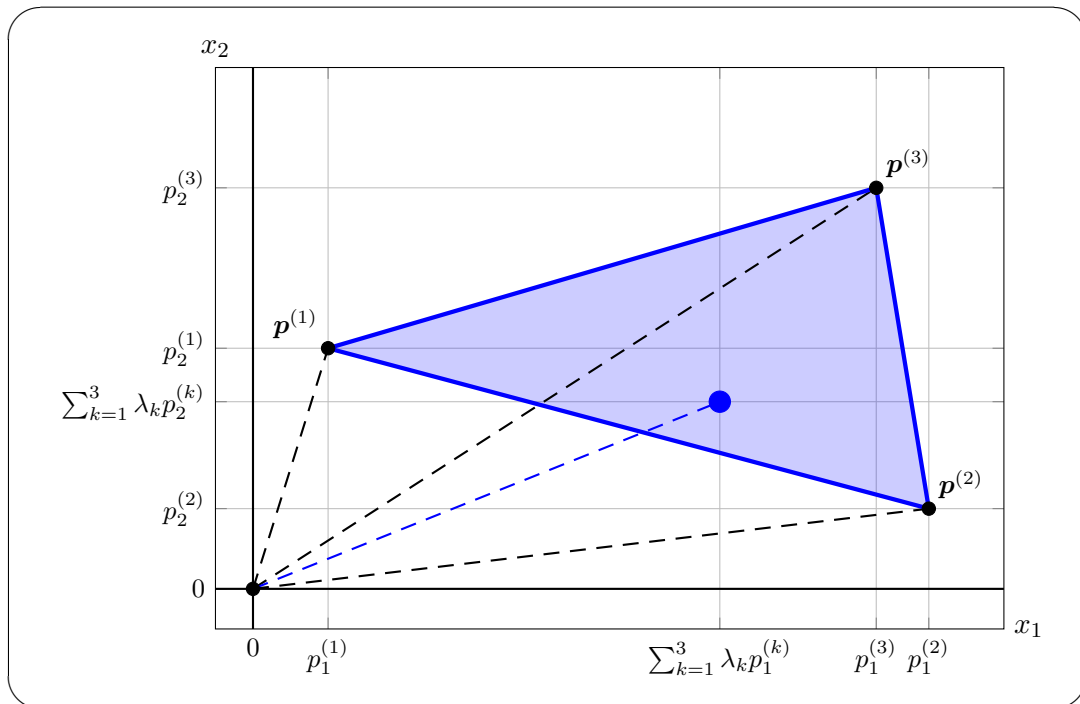
their **convex combination** with scalars $\lambda_1, \lambda_2, \dots, \lambda_m \geq 0$ satisfying

$$\sum_{k=1}^m \lambda_k = 1$$

is the point

$$\underbrace{\begin{pmatrix} \lambda_1 p_1^{(1)} + \lambda_2 p_1^{(2)} + \dots + \lambda_m p_1^{(m)} \\ \lambda_1 p_2^{(1)} + \lambda_2 p_2^{(2)} + \dots + \lambda_m p_2^{(m)} \\ \vdots \\ \lambda_1 p_n^{(1)} + \lambda_2 p_n^{(2)} + \dots + \lambda_m p_n^{(m)} \end{pmatrix}}_{\sum_{k=1}^m \lambda_k \mathbf{p}^{(k)}}$$

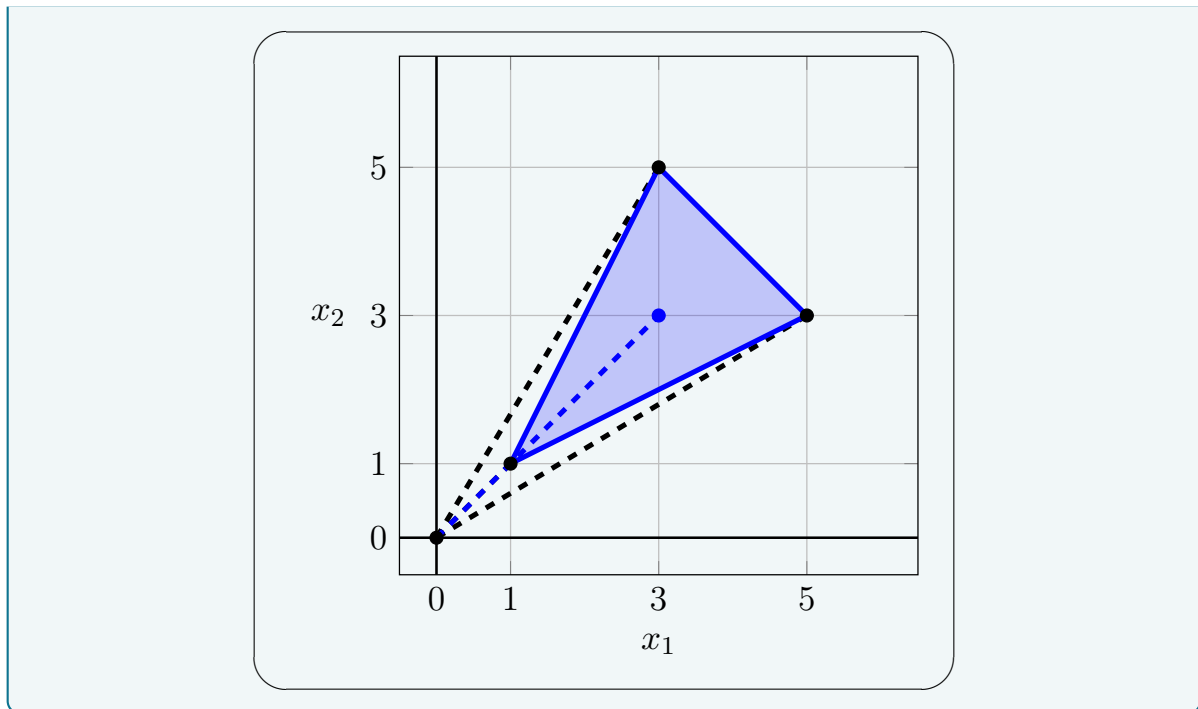
- The convex combinations for all valid scalars $\lambda_1, \lambda_2, \dots, \lambda_m \geq 0$ are the points in the **convex hull** of the m points.
- In $\mathbb{R}^2$ with $m = 3$ points, the convex hull is the set of all the points inside (and on the boundary of) the triangle connecting the three points: the blue filled region of the following figure, where the blue point is the convex combination with $\lambda_1 = \frac{1}{3}$, $\lambda_2 = \frac{1}{2}$ and $\lambda_3 = \frac{1}{6}$:



Example 1.2: convex combination of 3 points

Let us consider the three points $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ in $\mathbb{R}^2$. Their convex combination with $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$ is the point:

$$\begin{pmatrix} \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 5 + \frac{1}{3} \cdot 3 \\ \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 3 + \frac{1}{3} \cdot 5 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$



1.2 Convex sets

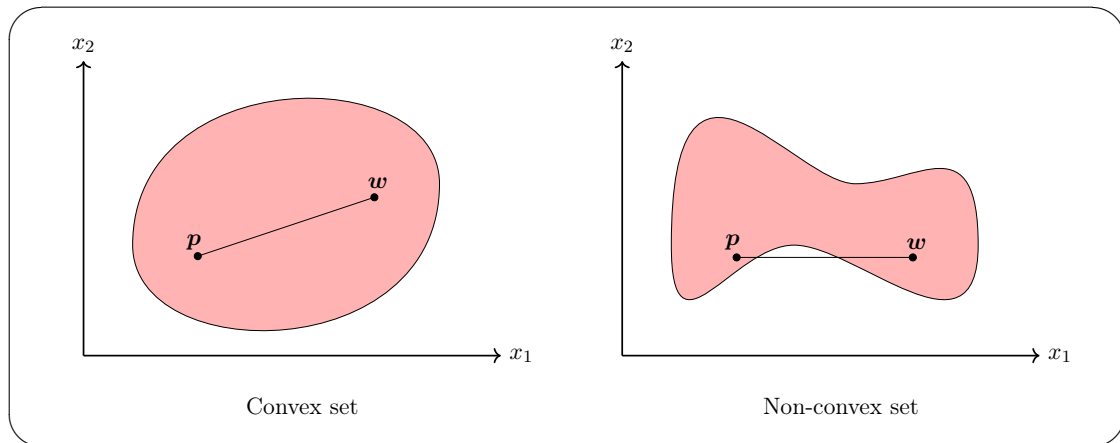
Definition 1.3: convex set

A set $A \subseteq \mathbb{R}^n$ is **convex** if

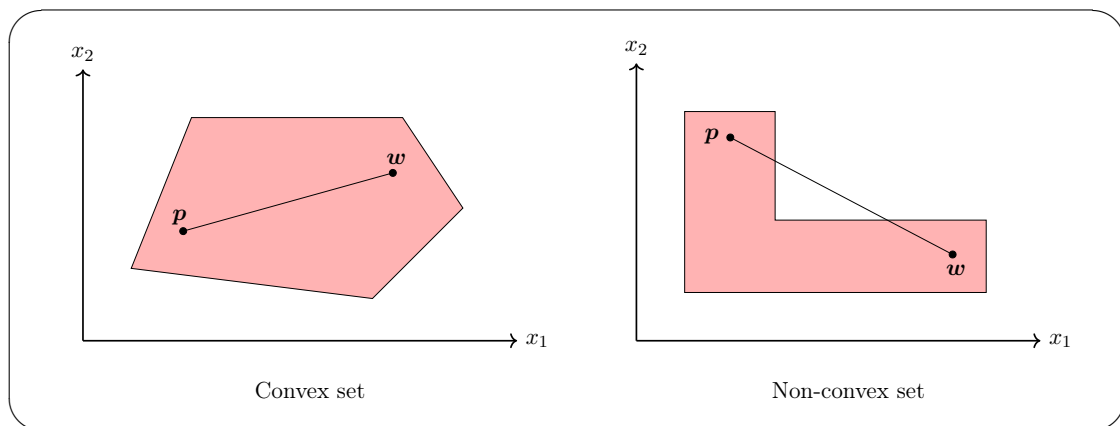
$$\forall \mathbf{p}, \mathbf{w} \in A \text{ and } \forall \lambda \in [0, 1] \text{ we have } \underbrace{\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}}_{\text{convex combination}} \in A$$

A subset of the n -dimensional space is convex if for any of its two points, any convex combination of the two points belongs to the subset. In other words, if all points on the segment that joins any two points of the subset belong to the subset.

- In $\mathbb{R}^2$ (the 2-dimensional space), an example of a convex set and an example of a non-convex set are given by the following picture, first with curved boundaries:



- The same with straight-edged sets: on the left a convex polygon, on the right a non-convex one.



1.3 Examples of convex sets

Observation 1.4

For every vector $\mathbf{c} \in \mathbb{R}^n$ and every scalar $r \in \mathbb{R}_{>0}$, the set

$$\{ \mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\|_1 \leq r \}$$

is convex.

It is the set of points in $\mathbb{R}^n$ whose sum of absolute deviations from $\mathbf{c}$ is at most r . In $\mathbb{R}^2$ this is a diamond (a square rotated by 45°), and in $\mathbb{R}^3$ it is a regular octahedron.

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

$$\begin{aligned} \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_1 &= \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_1 \\ &\leq \|\lambda(\mathbf{p} - \mathbf{c})\|_1 + \|(1 - \lambda)(\mathbf{w} - \mathbf{c})\|_1 \end{aligned}$$

$$\begin{aligned}
& \text{(by the triangle inequality)} \\
&= |\lambda| \|\mathbf{p} - \mathbf{c}\|_1 + |1 - \lambda| \|\mathbf{w} - \mathbf{c}\|_1 \\
& \text{(by the absolute homogeneity)} \\
&= \lambda \|\mathbf{p} - \mathbf{c}\|_1 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_1 \\
& \text{(since } \lambda \geq 0, 1 - \lambda \geq 0 \text{)} \\
&\leq \lambda r + (1 - \lambda)r \\
& \text{(since } \|\mathbf{p} - \mathbf{c}\|_1 \leq r \text{ and } \|\mathbf{w} - \mathbf{c}\|_1 \leq r \text{)} \\
&= r
\end{aligned}$$

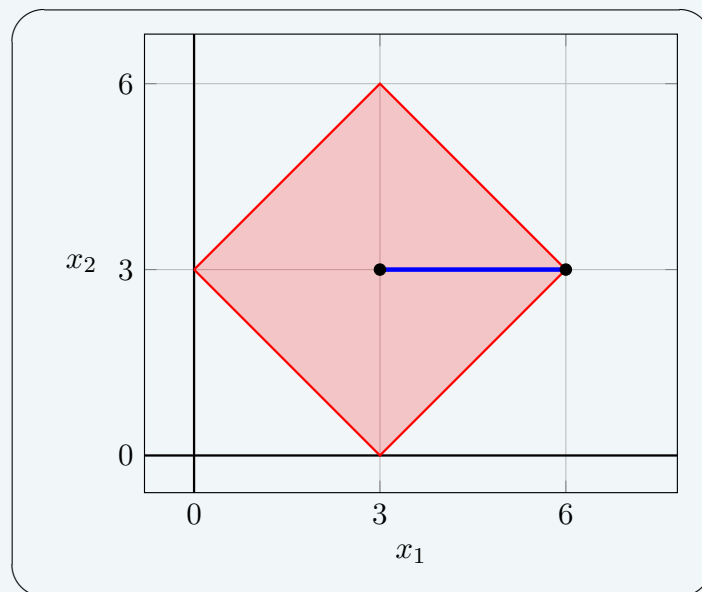
Therefore, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}$ belongs to the set, and the set is convex. $\square$

Example 1.3: diamond in $\mathbb{R}^2$

For example, in $\mathbb{R}^2$, with $r = 3$ and $\mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, we have the following convex set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : |x_1 - 3| + |x_2 - 3| \leq 3 \right\}$$

The points of this set are all those lying inside (and on) the diamond (rotated square) centered at $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$, with vertices at $\begin{pmatrix} 0 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 6 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 6 \end{pmatrix}$, shown as the red shaded region in the figure below.



Observation 1.5

For every vector $\mathbf{c} \in \mathbb{R}^n$ and every scalar $r \in \mathbb{R}_{>0}$, the set

$$\{ \mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\|_2 \leq r \}$$

is convex.

It is the set of points in $\mathbb{R}^n$ contained in the closed ball of radius r centered at $\mathbf{c}$. In $\mathbb{R}^2$ this is a disk, and in $\mathbb{R}^3$ it is a solid ball.

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

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

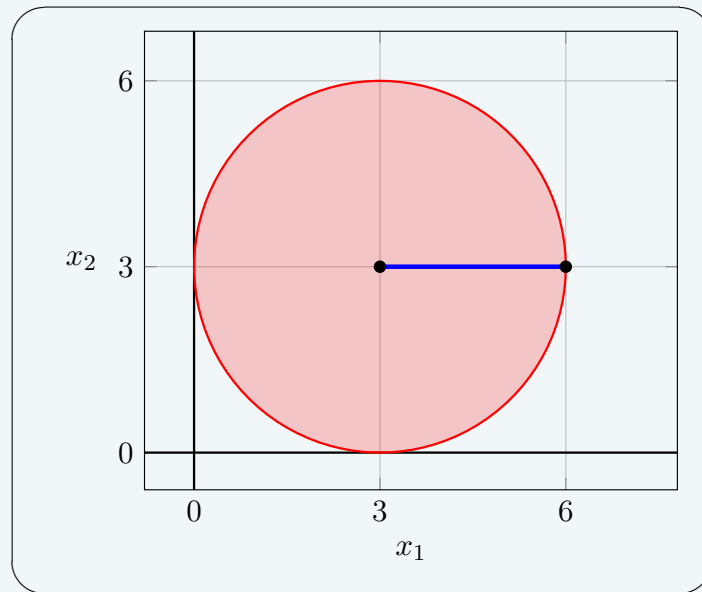
Therefore, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}$ belongs to the set, and the set is convex. □

Example 1.4: closed ball in $\mathbb{R}^2$

For example, in $\mathbb{R}^2$, with $r = 3$ and $\mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, we have the following convex set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : \sqrt{(x_1 - 3)^2 + (x_2 - 3)^2} \leq 3 \right\}$$

The points of this set are all those lying inside (and on the boundary of) the disk of radius 3 centered at $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$, shown as the red shaded region in the figure below.


Observation 1.6

For every vector $\mathbf{c} \in \mathbb{R}^n$ and every scalar $r \in \mathbb{R}_{>0}$, the set

$$\{ \mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\|_\infty \leq r \}$$

is convex.

It is the set of points in $\mathbb{R}^n$ whose maximum absolute deviation from $\mathbf{c}$ is at most r . In $\mathbb{R}^2$ this is an axis-aligned square of side $2r$ centered at $\mathbf{c}$, and in $\mathbb{R}^3$ it is a cube.

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

$$\|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_\infty = \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_\infty$$

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

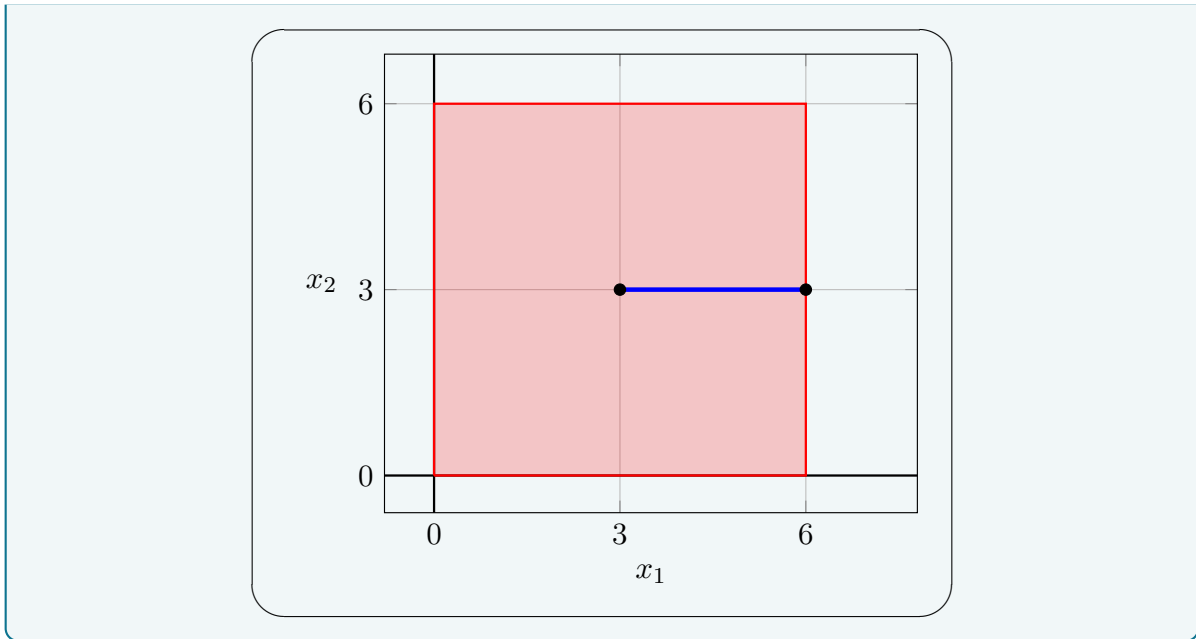
Therefore, $\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}$ belongs to the set, and the set is convex. $\square$

Example 1.5: square in $\mathbb{R}^2$

For example, in $\mathbb{R}^2$, with $r = 3$ and $\mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, we have the following convex set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : \max \{ |x_1 - 3|, |x_2 - 3| \} \leq 3 \right\}$$

The points of this set are all those lying inside (and on) the axis-aligned square of side 6 centered at $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$, with vertices at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 6 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 6 \\ 6 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 6 \end{pmatrix}$, shown as the red shaded region in the figure below.



1.4 Properties of convex sets

Observation 1.7

For every m convex sets $A_i \subseteq \mathbb{R}^n$ with $i \in \{1, 2, \dots, m\}$, the set

$$\bigcap_{i=1}^m A_i$$

is convex.

It is the set of points in $\mathbb{R}^n$ that belong to all the sets $A_1, A_2, \dots, A_m$ at the same time.

Proof. For every pair of points $\mathbf{p}, \mathbf{w} \in \bigcap_{i=1}^m A_i$ and every $\lambda \in [0, 1]$, we have:

$$\mathbf{p}, \mathbf{w} \in \bigcap_{i=1}^m A_i \implies \mathbf{p} \in A_i \text{ and } \mathbf{w} \in A_i, \quad \forall i \in \{1, 2, \dots, m\}$$

(by definition of intersection)

$$\implies \lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in A_i, \quad \forall i \in \{1, 2, \dots, m\}$$

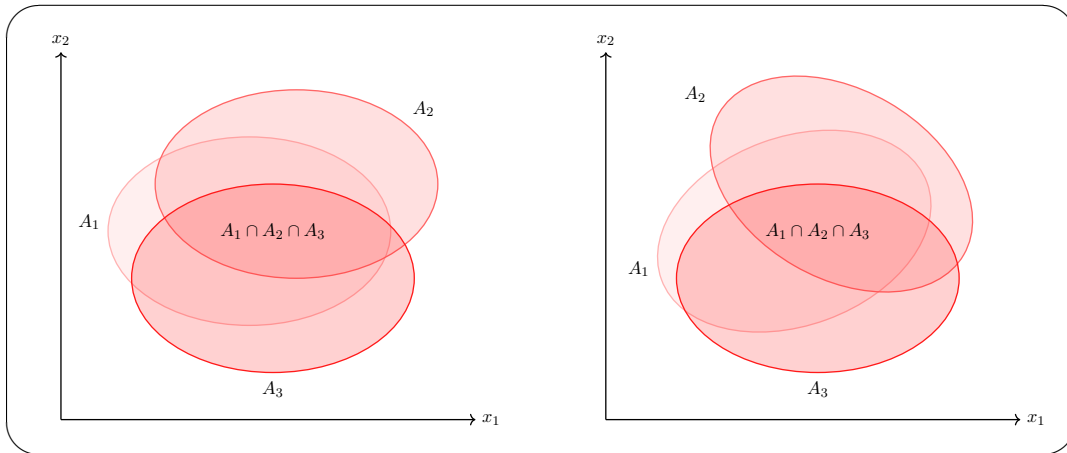
(since each A_i is convex)

$$\implies \lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in \bigcap_{i=1}^m A_i$$

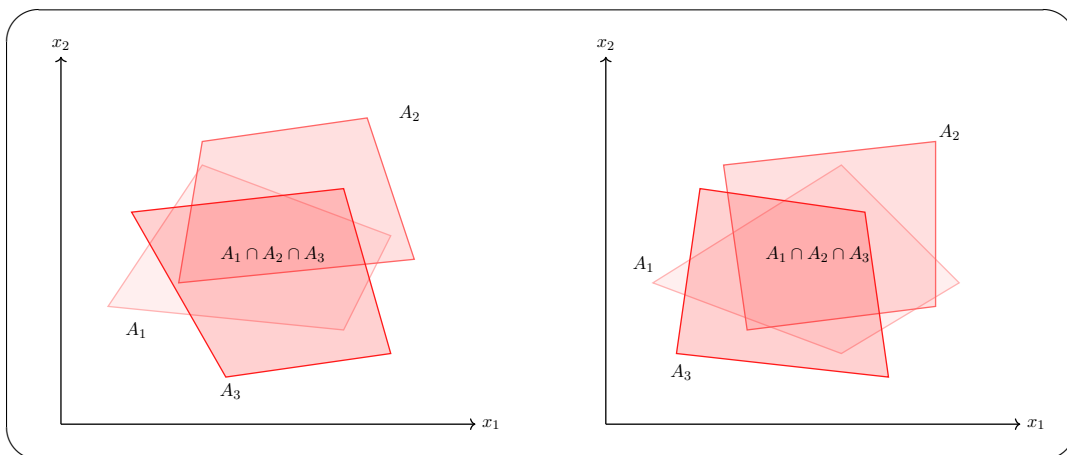
(by definition of intersection)

Therefore, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in \bigcap_{i=1}^m A_i$, and the set is convex. $\square$

- In $\mathbb{R}^2$ (the 2-dimensional space), examples of intersection of three convex sets A_1, A_2 and A_3 are given by the following pictures, first with three ellipses:



- The same with three convex quadrilaterals (straight edges, like the polyhedra of linear programming, which are all convex):



Functions of several variables

2.1 Functions of several variables

Definition 2.1: function of several real variables

The function:

$$f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}, \quad f : \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} \mapsto f(\mathbf{x})$$

is called a (scalar) **function of n real variables** and $D \subseteq \mathbb{R}^n$ is its **domain**.

- A function, commonly denoted f , of n real variables takes as input a vector of n real numbers, commonly represented by the (column) variable vector $\mathbf{x}$, and produces a real number, the value of the function, commonly denoted $f(\mathbf{x})$.
- Vectors are written in bold and are column vectors; the transpose of a vector $\mathbf{x}$ is the row vector $\mathbf{x}' = (x_1 \ x_2 \ \dots \ x_n)$.

Example 2.1: a function of three variables

Consider the function:

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

At the point $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ the value of the function is:

$$f \left(\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} \right) = 2 \cdot (-1) + 3^2 = -2 + 9 = 7$$

2.2 Gradient, Hessian and level sets

Definition 2.2: gradient

Given $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, the **gradient** $\nabla f(\mathbf{x})$ of f at $\mathbf{x}$ is the (column) vector of first-order partial derivatives:

$$\nabla f(\mathbf{x}) := \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

The gradient $\nabla f(\mathbf{x})$ (if it is nonzero) gives the direction of steepest ascent (maximum directional derivative among unit directions), while $-\nabla f(\mathbf{x})$ gives the direction of steepest descent.

Definition 2.3: Hessian

Given a function $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, the **Hessian** $\nabla^2 f(\mathbf{x})$ of f at $\mathbf{x}$ is the matrix of second-order partial derivatives:

$$\nabla^2 f(\mathbf{x}) := \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}$$

- The entry in row i and column j of the Hessian is obtained by differentiating $\frac{\partial f(\mathbf{x})}{\partial x_j}$ with respect to x_i .
- If the second-order partial derivatives of f are continuous (as for all the functions of these notes), the order of differentiation does not matter, $\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$, and the Hessian is a **symmetric** matrix.

The Hessian $\nabla^2 f(\mathbf{x})$ describes the local curvature of f around $\mathbf{x}$, that is, how the gradient changes when moving away from $\mathbf{x}$.

Definition 2.4: level set

Given a function $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ and a constant $k \in \mathbb{R}$, the **level set** $L(k)$ of f at level k is:

$$L(k) := \{ \mathbf{x} \in D : f(\mathbf{x}) = k \}$$

A level set $L(k)$ is the set of points where the function has a constant value k . The gradient $\nabla f(\mathbf{x})$ (if nonzero) is perpendicular to the level set passing through $\mathbf{x}$ and points toward increasing values of f .

Example 2.2: gradient, Hessian and level sets of a quadratic function

Consider the quadratic function:

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

- The **gradient** of the function is:

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

For example:

- at the point $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$, we have

$$\nabla f \left(\begin{pmatrix} 3 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} 2(3 - 2) \\ 2(2 - 2) \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

- at the point $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$, we have

$$\nabla f \left(\begin{pmatrix} 2 \\ 3 \end{pmatrix} \right) = \begin{pmatrix} 2(2 - 2) \\ 2(3 - 2) \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

- The **Hessian** of the function is:

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

Note that the Hessian is constant (it does not depend on the point).

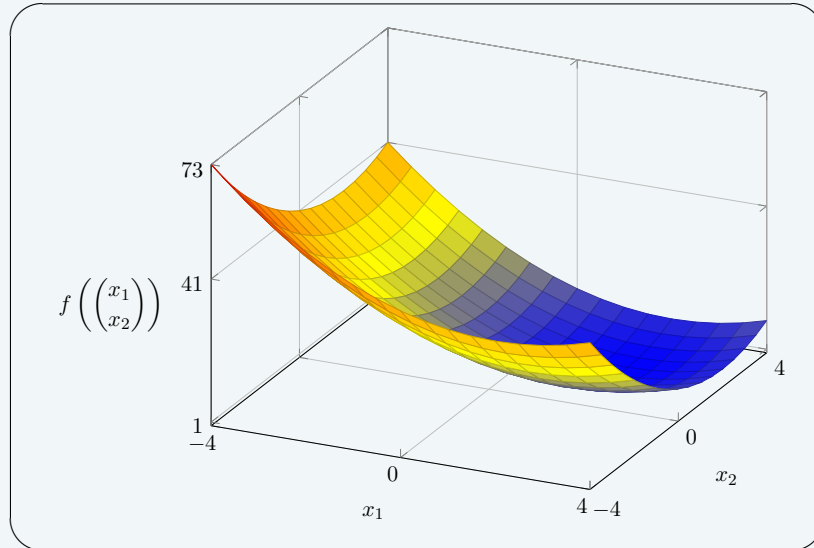
- The **level sets** of the function are:

$$L(k) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : (x_1 - 2)^2 + (x_2 - 2)^2 + 1 = k \right\}, \quad \forall k \in \mathbb{R}$$

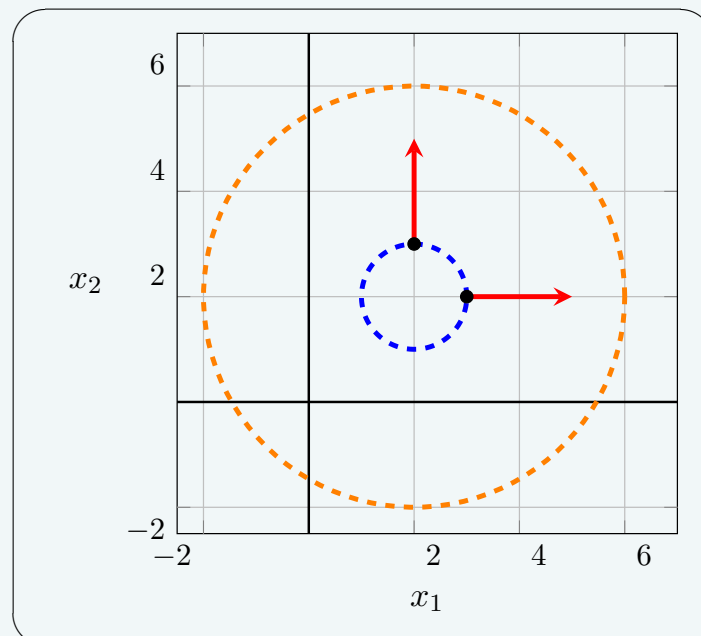
which are circles centered at $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ with radius $\sqrt{k - 1}$ for $k > 1$; the level set $L(1)$ is the single point $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ and the level sets with $k < 1$ are empty.

Example 2.3: graph, level sets and gradients of a quadratic function

For the function of the previous example, this is the graph in the domain $[-4, 4] \times [-4, 4]$; its lowest point is at $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$, where $f = 1$, and its highest point is at $\begin{pmatrix} -4 \\ -4 \end{pmatrix}$, where $f = 73$:



The following figure shows the level sets $L(2)$ (circle of radius 1, in blue) and $L(17)$ (circle of radius 4, in orange), and the two gradients of the previous example (in red), drawn from the points $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ of the level set $L(2)$:



The gradient vectors are perpendicular to the level sets and point in the direction of maximum increase of the function (away from the center).

2.3 Examples of functions that are not quadratic

- For a quadratic function the Hessian is constant (chapter «Quadratic functions»). In general, the Hessian depends on the point, as in the following two examples.

Example 2.4: gradient and Hessian of a cubic function

Consider the function:

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

- The **gradient** and the **Hessian** of the function are:

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

- At the point $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, we have:

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

- At the point $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$, we have:

$$\nabla f \left(\begin{pmatrix} -1 \\ 3 \end{pmatrix} \right) = \begin{pmatrix} 2 \cdot (-1) \cdot 3 \\ (-1)^2 \end{pmatrix} = \begin{pmatrix} -6 \\ 1 \end{pmatrix}, \quad \nabla^2 f \left(\begin{pmatrix} -1 \\ 3 \end{pmatrix} \right) = \begin{pmatrix} 6 & -2 \\ -2 & 0 \end{pmatrix}$$

The Hessian changes from point to point and, as expected, it is symmetric at every point.

Example 2.5: gradient, Hessian and a level set of an exponential function

Consider the function:

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

- The **gradient** and the **Hessian** of the function are:

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

- The level set $L(2)$ is the curve $e^{x_1} + x_2^2 = 2$, that is, $x_1 = \ln(2 - x_2^2)$ for $|x_2| < \sqrt{2}$. It contains the points $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} \ln 2 \\ 0 \end{pmatrix}$, since $e^0 + 1^2 = 2$ and $e^{\ln 2} + 0^2 = 2$. At these

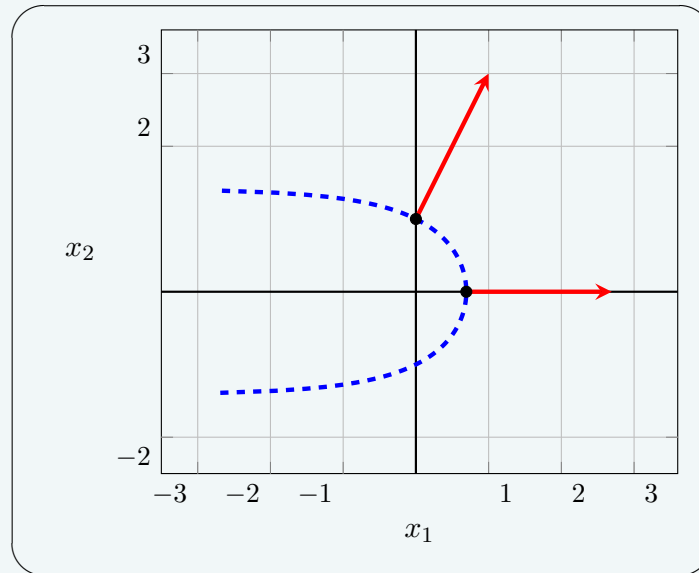
points:

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

and

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

The following figure shows the level set $L(2)$ (in blue) and the two gradient vectors (in red):



The tangent direction of the level set at $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ (the derivative of $x_1 = \ln(2 - x_2^2)$ with respect to x_2 is $\frac{-2x_2}{2-x_2^2}$, equal to -2 at $x_2 = 1$), and indeed:

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

so the gradient is perpendicular to the level set also when the level set is not a circle.

The gradient $\nabla f(\mathbf{x})$ is the column vector of first-order partial derivatives: it points in the direction of steepest ascent and is perpendicular to the level set through $\mathbf{x}$. The Hessian $\nabla^2 f(\mathbf{x})$ is the (symmetric) matrix of second-order partial derivatives and describes the local curvature of f .

Convex and concave functions

3.1 Convex functions

Definition 3.1: convex function

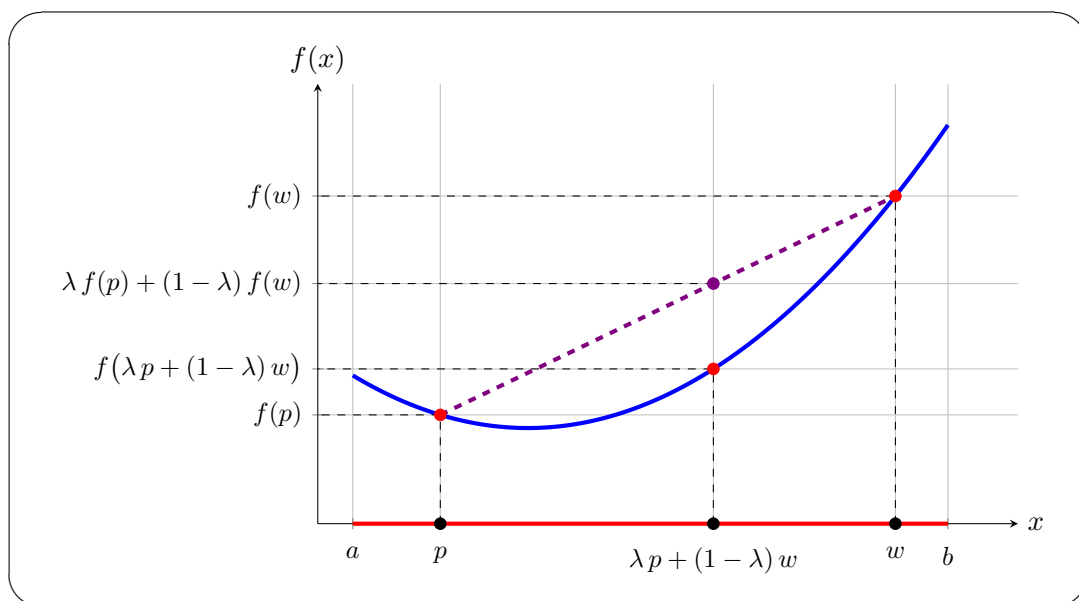
A function $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is **convex** in its domain D if:

$$\forall \mathbf{p}, \mathbf{w} \in D \text{ and } \forall \lambda \in [0, 1] \text{ we have } f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \quad (3.1)$$

and the domain D of the function f is a convex set.

A function is convex in its domain if, for any two points in its convex domain, the value at any convex combination of the two points is smaller than or equal to the convex combination of the values at those two points.

- If $D = [a, b] \subseteq \mathbb{R}$ (an interval which is by nature a convex set), the graph of a convex function is given by the following picture:



3.2 Concave functions

Definition 3.2: concave function

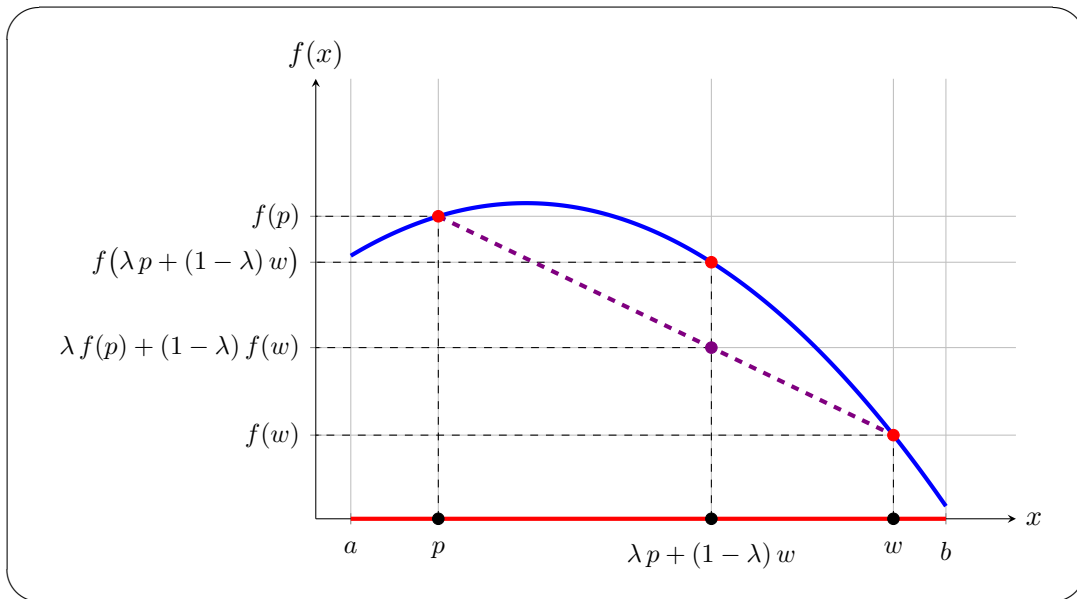
A function $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is **concave** in its domain D if:

$$\forall \mathbf{p}, \mathbf{w} \in D \text{ and } \forall \lambda \in [0, 1] \text{ we have } f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \geq \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \quad (3.2)$$

and the domain D of the function f is a convex set.

A function is concave in its domain if, for any two points in its convex domain, the value at any convex combination of the two points is greater than or equal to the convex combination of the values at those two points.

- If $D = [a, b] \subseteq \mathbb{R}$ (an interval which is by nature a convex set), the graph of a concave function is given by the following picture:



3.3 Examples of convex and concave functions

- In this chapter we also consider nonlinear functions. The convexity of a function is always proved by checking the definition: for arbitrary points $\mathbf{p}, \mathbf{w}$ of the domain and an arbitrary $\lambda \in [0, 1]$, we verify the inequality of the definition. For the quadratic functions of the first examples, we use the following inequality.

Observation 3.3

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

$$(\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2. \quad (3.3)$$

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

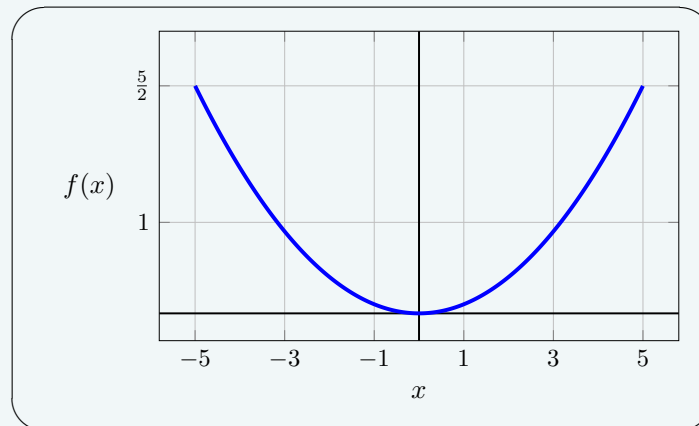
$$\begin{aligned} (\lambda a + (1 - \lambda)b)^2 &= \lambda^2 a^2 + 2\lambda(1 - \lambda)ab + (1 - \lambda)^2 b^2 \\ &= \lambda a^2 + (1 - \lambda)b^2 - \lambda(1 - \lambda)a^2 + 2\lambda(1 - \lambda)ab - \lambda(1 - \lambda)b^2 \\ &= \lambda a^2 + (1 - \lambda)b^2 - \lambda(1 - \lambda)(a^2 - 2ab +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

□

Example 3.1: a convex quadratic function

Consider the following function:

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f : x \mapsto \frac{x^2}{10}$$



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

$$\begin{aligned}
 f(\lambda p + (1 - \lambda)w) &= \frac{1}{10}(\lambda p + (1 - \lambda)w)^2 \\
 &\leq \frac{1}{10}(\lambda p^2 + (1 - \lambda)w^2) \quad (\text{by inequality (3.3)}) \\
 &= \lambda \frac{p^2}{10} + (1 - \lambda) \frac{w^2}{10} \\
 &= \lambda f(p) + (1 - \lambda)f(w)
 \end{aligned}$$

Therefore,

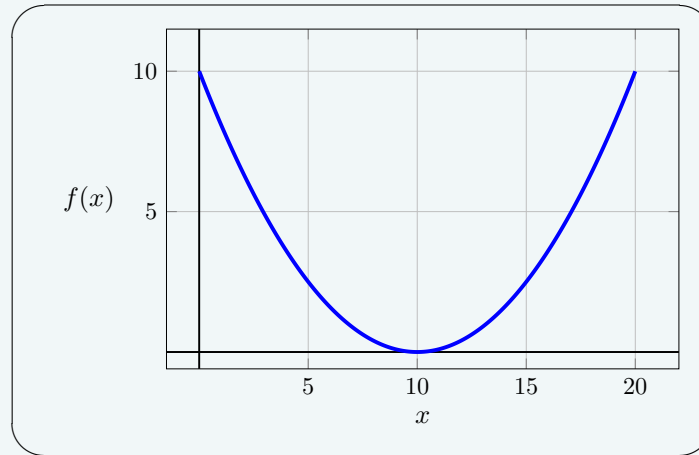
$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

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

Example 3.2: a translated convex quadratic function

Consider the following function:

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f : x \mapsto \frac{1}{10}(x - 10)^2$$



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

$$\begin{aligned}
 f(\lambda p + (1 - \lambda)w) &= \frac{1}{10}(\lambda p + (1 - \lambda)w - 10)^2 \\
 &= \frac{1}{10}(\lambda(p - 10) + (1 - \lambda)(w - 10))^2 \quad (\text{since } \lambda + (1 - \lambda) = 1) \\
 &\leq \frac{1}{10}(\lambda(p - 10)^2 + (1 - \lambda)(w - 10)^2) \quad (\text{by inequality (3.3)}) \\
 &= \lambda \frac{(p - 10)^2}{10} + (1 - \lambda) \frac{(w - 10)^2}{10}
 \end{aligned}$$

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

Therefore,

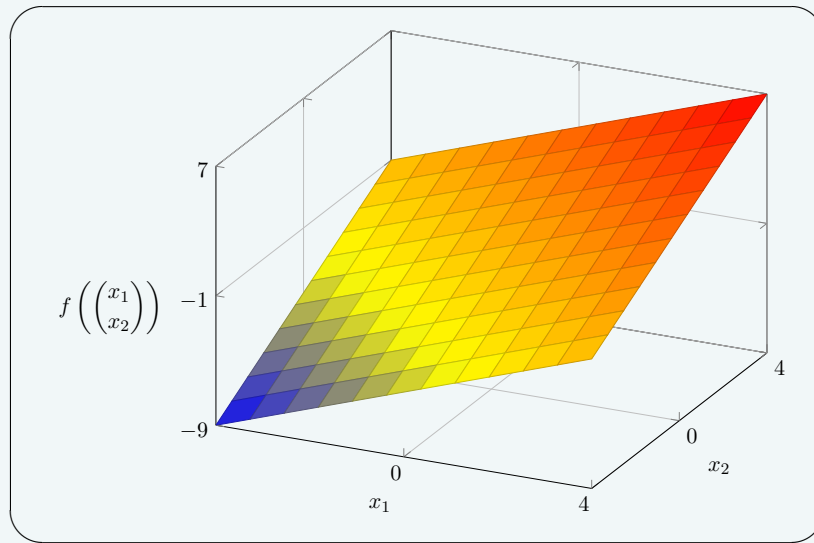
$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

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

Example 3.3: a linear function of two variables

Consider the following function:

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



For every pair of points $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

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

Therefore,

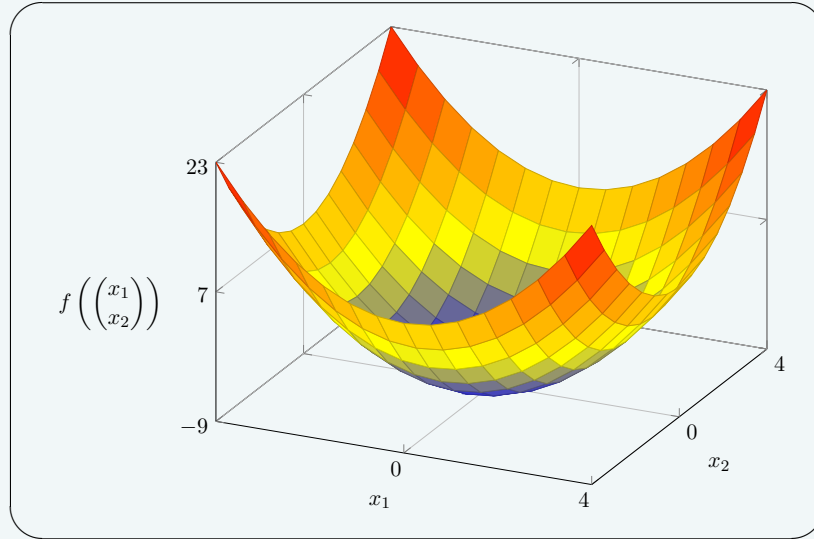
$$f(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w})$$

(with equality). By nature, the domain $\mathbb{R}^2$ is a convex set. Accordingly, the function is convex in its domain. Since the equality holds, also the inequality with $\geq$ holds: the function is also concave in its domain.

Example 3.4: a convex quadratic function of two variables

Consider the following function:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_2^2 - 9$$



For every pair of points $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= (\lambda p_1 + (1 - \lambda) w_1)^2 + (\lambda p_2 + (1 - \lambda) w_2)^2 - 9 \\ &\leq \lambda p_1^2 + (1 - \lambda) w_1^2 + \lambda p_2^2 + (1 - \lambda) w_2^2 - 9 \\ &\quad \text{(by inequality (3.3), twice)} \\ &= \lambda(p_1^2 + p_2^2 - 9) + (1 - \lambda)(w_1^2 + w_2^2 - 9) \\ &\quad \text{(since } \lambda + (1 - \lambda) = 1\text{)} \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

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.

- The following two examples are functions of one variable whose graph is made of two segments.

Example 3.5: the absolute value is a convex function

Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$. For every pair of points $p, w \in \mathbb{R}$ and every $\lambda \in [0, 1]$, we have:

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

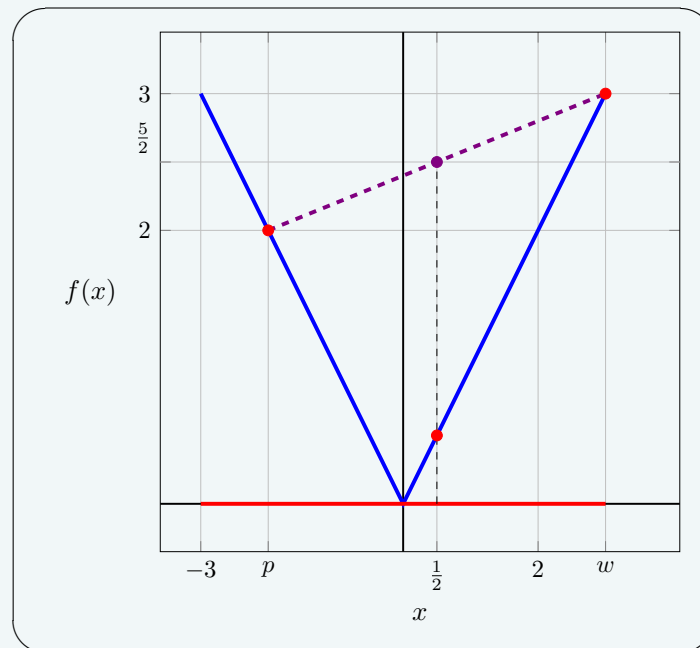
Therefore,

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

By nature, the domain $\mathbb{R}$ is a convex set. Accordingly, the function is convex in its domain. For example, with $p = -2$, $w = 3$ and $\lambda = \frac{1}{2}$ we obtain $\lambda p + (1 - \lambda)w = \frac{1}{2}$ and

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

In the figure, the segment joining $(p, f(p))$ and $(w, f(w))$ (violet) lies above the graph of f (blue).



Example 3.6: the minimum of two linear functions is a concave function

Consider the function $f : [-2, 4] \rightarrow \mathbb{R}$, $f(x) = \min\{x + 2, 4 - \frac{1}{2}x\}$. For every pair of points $p, w \in [-2, 4]$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} f(\lambda p + (1 - \lambda)w) &= \min \left\{ \lambda p + (1 - \lambda)w + 2, 4 - \frac{1}{2}(\lambda p + (1 - \lambda)w) \right\} \\ &= \min \left\{ \lambda(p + 2) + (1 - \lambda)(w + 2), \lambda(4 - \frac{1}{2}p) + (1 - \lambda)(4 - \frac{1}{2}w) \right\} \\ &\quad (\text{since } \lambda + (1 - \lambda) = 1) \\ &\geq \min \left\{ \lambda f(p) + (1 - \lambda)f(w), \lambda f(p) + (1 - \lambda)f(w) \right\} \\ &\quad (\text{since } f(p) \leq p + 2, f(p) \leq 4 - \frac{1}{2}p, f(w) \leq w + 2, \\ &\quad \quad f(w) \leq 4 - \frac{1}{2}w, \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\ &= \lambda f(p) + (1 - \lambda)f(w) \end{aligned}$$

Therefore,

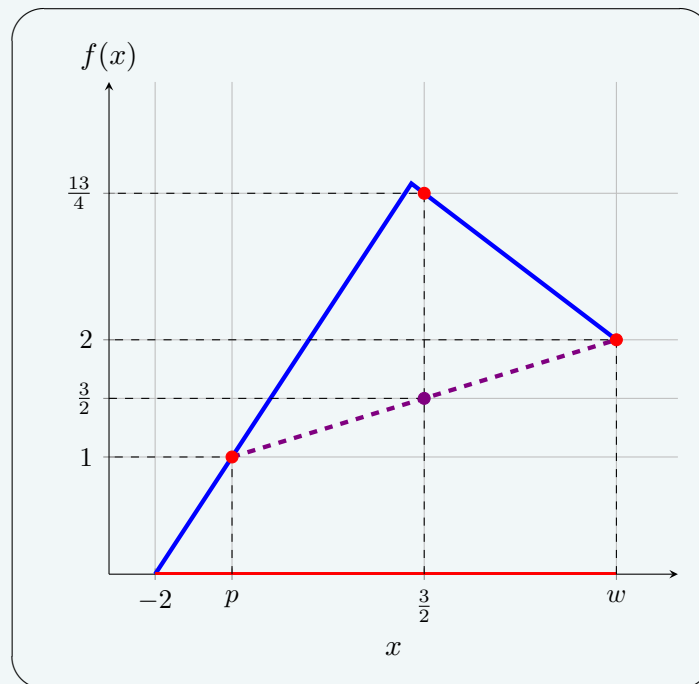
$$f(\lambda p + (1 - \lambda)w) \geq \lambda f(p) + (1 - \lambda)f(w)$$

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

The two lines meet at $x = \frac{4}{3}$, where $f(\frac{4}{3}) = \frac{10}{3}$. For example, with $p = -1$, $w = 4$ and $\lambda = \frac{1}{2}$ we obtain $\lambda p + (1 - \lambda)w = \frac{3}{2}$ and

$$f\left(\frac{3}{2}\right) = \min \left\{ \frac{7}{2}, \frac{13}{4} \right\} = \frac{13}{4} \geq \frac{1}{2} f(-1) + \frac{1}{2} f(4) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 = \frac{3}{2}.$$

In the figure, the segment joining $(p, f(p))$ and $(w, f(w))$ (violet) lies below the graph of f (blue).



3.4 Convex functions defined by norms

Definition 3.4: norm

A **norm** on $\mathbb{R}^n$ is a function $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ such that, for every $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ and every $\alpha \in \mathbb{R}$:

- $\|\mathbf{p}\| \geq 0$, and $\|\mathbf{p}\| = 0$ if and only if $\mathbf{p} = \mathbf{0}$ (positivity);
- $\|\alpha \mathbf{p}\| = |\alpha| \|\mathbf{p}\|$ (absolute homogeneity);
- $\|\mathbf{p} + \mathbf{w}\| \leq \|\mathbf{p}\| + \|\mathbf{w}\|$ (triangle inequality).

Definition 3.5: the norms 1, 2, q and ∞

For every $\mathbf{x} \in \mathbb{R}^n$:

$$\|\mathbf{x}\|_1 = \sum_{j=1}^n |x_j| \quad (\text{norm 1, or Manhattan norm})$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^n x_j^2} \quad (\text{norm 2, or Euclidean norm})$$

$$\|\mathbf{x}\|_q = \left(\sum_{j=1}^n |x_j|^q \right)^{\frac{1}{q}}, \quad q \geq 1 \quad (\text{norm } q)$$

$$\|\mathbf{x}\|_\infty = \max_{j=1, \dots, n} |x_j| \quad (\text{norm } \infty, \text{ or maximum norm})$$

The norms 1 and 2 are the norm q with $q = 1$ and $q = 2$, and $\|\mathbf{x}\|_\infty = \lim_{q \rightarrow \infty} \|\mathbf{x}\|_q$.

Example 3.7: the four norms of a vector

For $\mathbf{x} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$ we have:

$$\|\mathbf{x}\|_1 = |3| + |-4| = 7, \quad \|\mathbf{x}\|_2 = \sqrt{3^2 + (-4)^2} = 5,$$

$$\|\mathbf{x}\|_3 = (|3|^3 + |-4|^3)^{\frac{1}{3}} = \sqrt[3]{91}, \quad \|\mathbf{x}\|_\infty = \max\{|3|, |-4|\} = 4.$$

Observation 3.6: convexity of the Euclidean norm

The function

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f : \mathbf{x} \mapsto \|\mathbf{x}\|_2$$

is convex in its domain.

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

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

Therefore,

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

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

Observation 3.7: convexity of every norm

For every norm $\|\cdot\|$ on $\mathbb{R}^n$ (in particular $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_q$ with $q \geq 1$, and $\|\cdot\|_\infty$), the function

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f : \mathbf{x} \mapsto \|\mathbf{x}\|$$

is convex in its domain.

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

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

Therefore,

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

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

Corollary 3.8

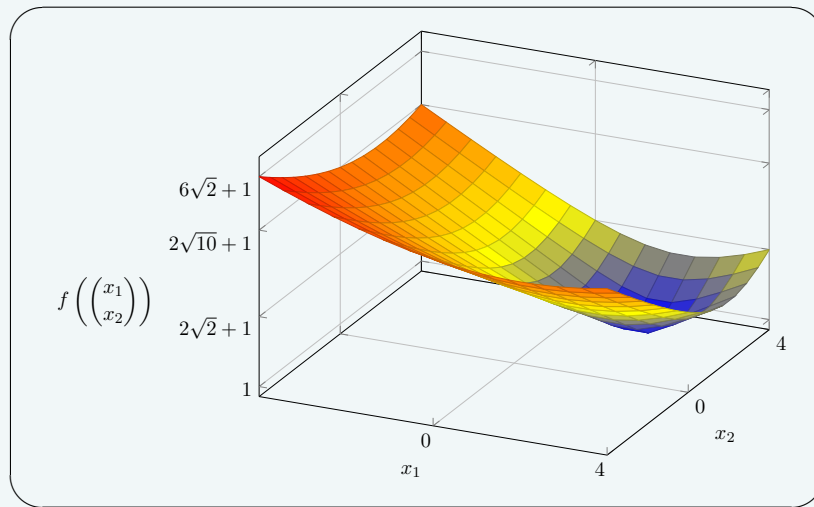
For every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have

$$\|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}\|_2 \leq \lambda \|\mathbf{p}\|_2 + (1 - \lambda) \|\mathbf{w}\|_2 \quad (3.4)$$

Example 3.8: the distance from a point

Let us consider the vector $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ and the function:

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



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

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_2 + 1 \\ &= \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2 + 1 \quad (\text{since } \lambda + (1 - \lambda) = 1) \\ &\leq \lambda \|\mathbf{p} - \mathbf{c}\|_2 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_2 + 1 \quad (\text{by inequality (3.4)}) \\ &= \lambda(\|\mathbf{p} - \mathbf{c}\|_2 + 1) + (1 - \lambda)(\|\mathbf{w} - \mathbf{c}\|_2 + 1) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

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.

Observation 3.9: convexity of the squared Euclidean norm

The function

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f : \mathbf{x} \mapsto \|\mathbf{x}\|_2^2$$

is convex in its domain.

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

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}\|_2^2 \\ &= (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})' (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\ &= \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + 2\lambda(1 - \lambda) (\mathbf{p}' \mathbf{w}) \\ &\leq \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + 2\lambda(1 - \lambda) \|\mathbf{p}\|_2 \|\mathbf{w}\|_2 \\ &\quad \text{(by the Cauchy-Schwarz inequality)} \\ &\leq \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + \lambda(1 - \lambda) \|\mathbf{p}\|_2^2 + \lambda(1 - \lambda) \|\mathbf{w}\|_2^2 \\ &\quad \text{(since } 2ab \leq a^2 + b^2 \text{ for all } a, b \in \mathbb{R}, \\ &\quad \text{with } a = \|\mathbf{p}\|_2 \text{ and } b = \|\mathbf{w}\|_2) \\ &= (\lambda^2 + \lambda(1 - \lambda)) \|\mathbf{p}\|_2^2 + ((1 - \lambda)^2 + \lambda(1 - \lambda)) \|\mathbf{w}\|_2^2 \\ &= \lambda \|\mathbf{p}\|_2^2 + (1 - \lambda) \|\mathbf{w}\|_2^2 \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Therefore,

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

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

- The next proof is more direct and does not require the use of the Cauchy-Schwarz inequality and the inequality $2ab \leq a^2 + b^2$ for every $a, b \in \mathbb{R}$.

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

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})' (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\ &= \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + 2\lambda(1 - \lambda) (\mathbf{p}' \mathbf{w}) \\ &= \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 \\ &\quad + \lambda(1 - \lambda) (\|\mathbf{p}\|_2^2 + \|\mathbf{w}\|_2^2 - \|\mathbf{p} - \mathbf{w}\|_2^2) \\ &\quad \text{(since } \|\mathbf{p} - \mathbf{w}\|_2^2 = \|\mathbf{p}\|_2^2 + \|\mathbf{w}\|_2^2 - 2\mathbf{p}' \mathbf{w}) \end{aligned}$$

$$\begin{aligned}
&= (\lambda^2 + \lambda(1 - \lambda))\|\mathbf{p}\|_2^2 + ((1 - \lambda)^2 + \lambda(1 - \lambda))\|\mathbf{w}\|_2^2 \\
&\quad - \lambda(1 - \lambda)\|\mathbf{p} - \mathbf{w}\|_2^2 \\
&= \lambda\|\mathbf{p}\|_2^2 + (1 - \lambda)\|\mathbf{w}\|_2^2 - \lambda(1 - \lambda)\|\mathbf{p} - \mathbf{w}\|_2^2 \\
&\leq \lambda\|\mathbf{p}\|_2^2 + (1 - \lambda)\|\mathbf{w}\|_2^2 \quad (\text{since } \lambda(1 - \lambda)\|\mathbf{p} - \mathbf{w}\|_2^2 \geq 0) \\
&= \lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w})
\end{aligned}$$

Therefore,

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

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

Corollary 3.10

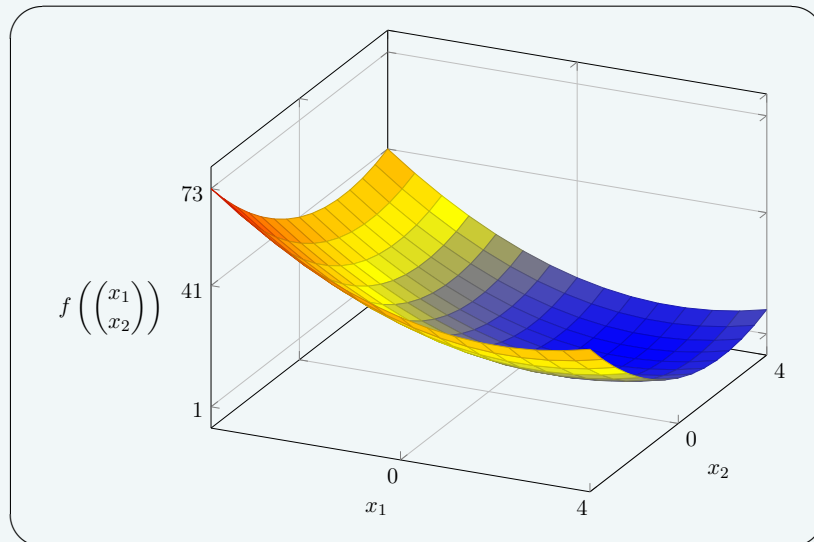
For every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have

$$\|\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}\|_2^2 \leq \lambda\|\mathbf{p}\|_2^2 + (1 - \lambda)\|\mathbf{w}\|_2^2 \quad (3.5)$$

Example 3.9: the squared distance from a point

Let us consider the vector $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ and the function:

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



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

$$\begin{aligned}
 f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_2^2 + 1 \\
 &= \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2^2 + 1 \quad (\text{since } \lambda + (1 - \lambda) = 1) \\
 &\leq \lambda \|\mathbf{p} - \mathbf{c}\|_2^2 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_2^2 + 1 \quad (\text{by inequality (3.5)}) \\
 &= \lambda(\|\mathbf{p} - \mathbf{c}\|_2^2 + 1) + (1 - \lambda)(\|\mathbf{w} - \mathbf{c}\|_2^2 + 1) \\
 &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})
 \end{aligned}$$

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.

3.5 Properties of convex and concave functions

Observation 3.11

A function $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ is convex in D if and only if the function $-f$ is concave in D .

Proof. Let f be convex in D . For every pair of points $\mathbf{p}, \mathbf{w} \in D$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 (-f)(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= -f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\
 &\geq -(\lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})) \\
 &\quad (\text{since } f \text{ is convex, and } -1 \text{ reverses the inequality}) \\
 &= \lambda (-f)(\mathbf{p}) + (1 - \lambda) (-f)(\mathbf{w})
 \end{aligned}$$

Therefore,

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

The domain D is a convex set, since f is convex in D . Accordingly, the function $-f$ is concave in its domain. The converse is proved in the same way, starting from the inequality of the concave function $-f$ and multiplying it by -1 . $\square$

Observation 3.12

Let $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ be convex in D and let $\alpha, \beta \geq 0$. Then the function $\alpha f + \beta g$ is convex in D . The same holds for concave functions.

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

$$\begin{aligned}
 (\alpha f + \beta g)(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \alpha f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) + \beta g(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\
 &\leq \alpha(\lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})) + \beta(\lambda g(\mathbf{p}) + (1 - \lambda) g(\mathbf{w})) \\
 &\quad (\text{since } f \text{ and } g \text{ are convex, and } \alpha \geq 0, \beta \geq 0) \\
 &= \lambda(\alpha f(\mathbf{p}) + \beta g(\mathbf{p})) + (1 - \lambda)(\alpha f(\mathbf{w}) + \beta g(\mathbf{w})) \\
 &= \lambda(\alpha f + \beta g)(\mathbf{p}) + (1 - \lambda)(\alpha f + \beta g)(\mathbf{w})
 \end{aligned}$$

Therefore,

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

The domain D is a convex set, since f is convex in D . Accordingly, the function $\alpha f + \beta g$ is convex in its domain. For concave functions the same argument holds with $\geq$. $\square$

- The point-wise maximum of convex functions is itself a convex function: this allows us to combine multiple convex functions into a single one while preserving convexity.

Observation 3.13

For every m functions $f_i : D_i \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convex in their domains, with $i \in \{1, 2, \dots, m\}$, the function

$$f : \bigcap_{i=1}^m D_i \subseteq \mathbb{R}^n \rightarrow \mathbb{R}, \quad \mathbf{x} \mapsto \max \{ f_i(\mathbf{x}) : i \in \{1, 2, \dots, m\} \}$$

is convex in its domain. Similarly, the minimum of m concave functions is concave in its domain.

Proof. For every pair of points $\mathbf{p}, \mathbf{w} \in \bigcap_{i=1}^m D_i$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned}
 f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \max \{ f_i(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) : i \in \{1, 2, \dots, m\} \} \\
 &= f_k(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\
 &\quad (\text{for an index } k \in \{1, 2, \dots, m\} \text{ attaining the maximum}) \\
 &\leq \lambda f_k(\mathbf{p}) + (1 - \lambda) f_k(\mathbf{w}) \quad (\text{since } f_k \text{ is convex}) \\
 &\leq \lambda \max \{ f_i(\mathbf{p}) : i \in \{1, 2, \dots, m\} \} \\
 &\quad + (1 - \lambda) \max \{ f_i(\mathbf{w}) : i \in \{1, 2, \dots, m\} \} \\
 &\quad (\text{since } \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\
 &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})
 \end{aligned}$$

Therefore,

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

The domain $\bigcap_{i=1}^m D_i$ is a convex set, since it is the intersection of convex sets. Accordingly, the function f is convex in its domain. The statement for the minimum of concave functions follows by applying this one to $-f_1, \dots, -f_m$. $\square$

- The absolute value $|x| = \max\{x, -x\}$ of Example 3.5 is the maximum of the two functions x and $-x$, which are convex (and concave), since the definition holds with equality: $-(\lambda p + (1 - \lambda)w) = \lambda(-p) + (1 - \lambda)(-w)$. The function of Example 3.6 is the minimum of the two functions $x + 2$ and $4 - \frac{1}{2}x$, which are concave for the same reason: $(\lambda p + (1 - \lambda)w) + 2 = \lambda(p + 2) + (1 - \lambda)(w + 2)$ and $4 - \frac{1}{2}(\lambda p + (1 - \lambda)w) = \lambda(4 - \frac{1}{2}p) + (1 - \lambda)(4 - \frac{1}{2}w)$. By the previous observation, we find again that the two functions are convex and concave, respectively.

Example 3.10: maximum of convex functions

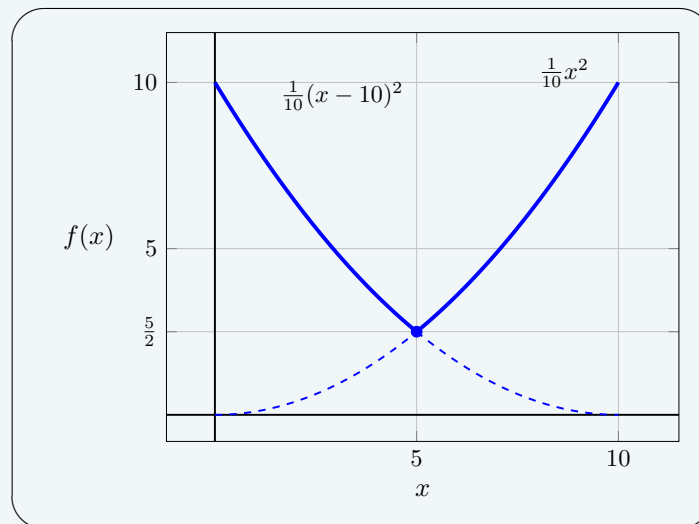
Consider the following two convex functions (Examples 3.2 and 3.1):

$$f_1 : \mathbb{R} \rightarrow \mathbb{R}, \quad f_1 : x \mapsto \frac{1}{10}(x - 10)^2 \quad \text{and} \quad f_2 : \mathbb{R} \rightarrow \mathbb{R}, \quad f_2 : x \mapsto \frac{1}{10}x^2$$

According to the previous observation, the following function is convex in its domain:

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f : x \mapsto \max \left\{ \frac{1}{10}(x - 10)^2, \frac{1}{10}x^2 \right\}$$

In the figure, the graph of f is the solid blue line; the dashed parts of the two parabolas lie below it.



Observation 3.14

For every m functions $f_i : D_i \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convex in their domains, with $i \in \{1, 2, \dots, m\}$, the set

$$\left\{ \mathbf{x} \in \bigcap_{i=1}^m D_i : f_i(\mathbf{x}) \leq 0, \forall i \in \{1, 2, \dots, m\} \right\}$$

is convex.

Proof. For every pair of points $\mathbf{p}$ and $\mathbf{w}$ in the set and every $\lambda \in [0, 1]$, the point $\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}$ belongs to $\bigcap_{i=1}^m D_i$, since $\mathbf{p}, \mathbf{w} \in D_i$ and each domain D_i is convex, and for every $i \in \{1, 2, \dots, m\}$ we have:

$$\begin{aligned} f_i(\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}) &\leq \lambda f_i(\mathbf{p}) + (1 - \lambda)f_i(\mathbf{w}) \quad (\text{since } f_i \text{ is convex}) \\ &\leq \lambda 0 + (1 - \lambda) 0 \\ &\quad (\text{since } f_i(\mathbf{p}) \leq 0, f_i(\mathbf{w}) \leq 0, \lambda \geq 0 \text{ and } 1 - \lambda \geq 0) \\ &= 0 \end{aligned}$$

Therefore,

$$f_i(\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}) \leq 0, \quad \forall i \in \{1, 2, \dots, m\}$$

Accordingly, $\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}$ belongs to the set and the set is convex. $\square$

Example 3.11: convex set from convex functions

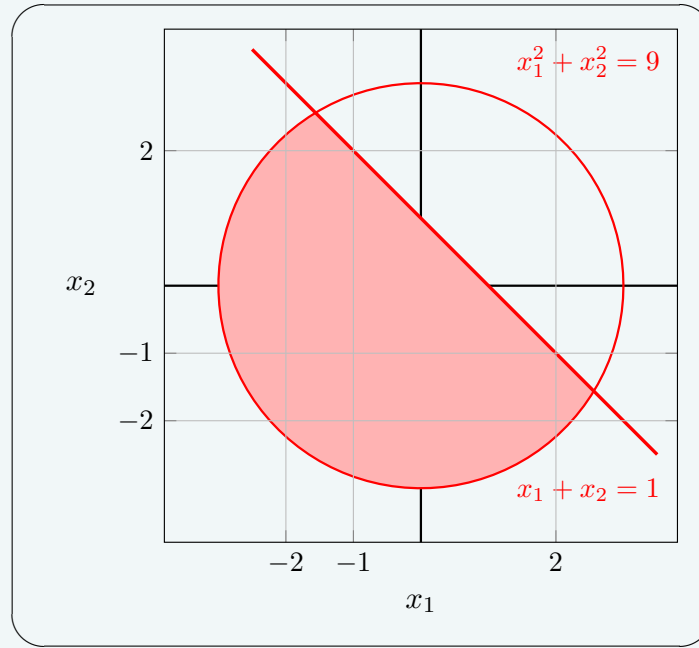
Consider the following two convex functions (Examples 3.4 and 3.3):

$$f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_1 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + x_2^2 - 9 \quad \text{and} \quad f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_2 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 + x_2 - 1$$

According to the previous observation, the following is a convex set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : x_1^2 + x_2^2 - 9 \leq 0 \quad \text{and} \quad x_1 + x_2 - 1 \leq 0 \right\}$$

It is the disk of radius 3 centered at the origin (red circle) cut by the half-plane $x_1 + x_2 \leq 1$ (red line), shaded in red.



Example 3.12: convex set from convex functions

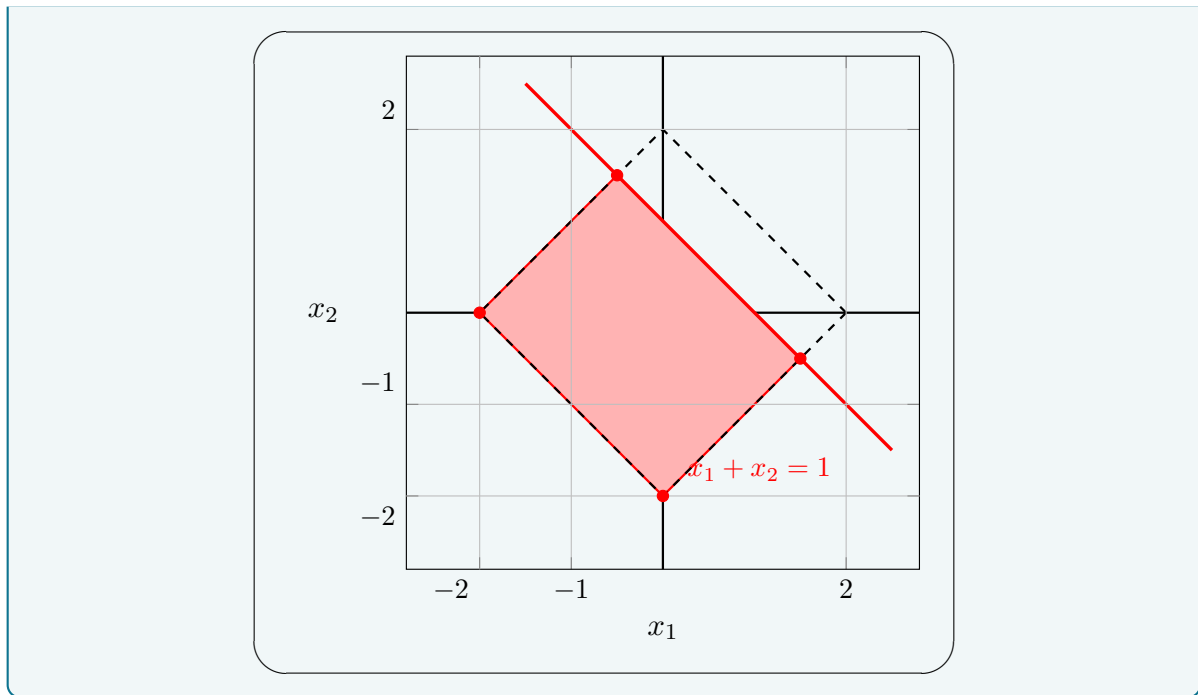
Consider the following two convex functions:

$$f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_1 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto |x_1| + |x_2| - 2 \quad \text{and} \quad f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_2 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1 + x_2 - 1.$$

The function f_1 is convex because $|x_1| = \max\{x_1, -x_1\}$ and $|x_2| = \max\{x_2, -x_2\}$ are maxima of convex functions (as $|x| = \max\{x, -x\}$ above), and the sum of convex functions is convex; f_2 is convex by Example 3.3. Also in this case, according to the same observation, the following is a convex set:

$$\left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 : |x_1| + |x_2| - 2 \leq 0 \text{ and } x_1 + x_2 - 1 \leq 0 \right\}.$$

It is the square with vertices $(\pm 2, 0)$ and $(0, \pm 2)$ (dashed) cut by the half-plane $x_1 + x_2 \leq 1$ (red line): the quadrilateral with vertices $(-2, 0)$, $(0, -2)$, $(\frac{3}{2}, -\frac{1}{2})$ and $(-\frac{1}{2}, \frac{3}{2})$ shaded in red.



The points that satisfy a finite number of constraints $f_i(\mathbf{x}) \leq 0$ with convex functions f_i form a convex set. For instance, every half-space $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'\mathbf{x} \leq b\}$, with $\mathbf{a} \in \mathbb{R}^n$ and $b \in \mathbb{R}$, is a convex set of this kind, since the linear function $\mathbf{a}'\mathbf{x} - b$ is convex; so is every intersection of finitely many half-spaces, as the polyhedra of linear programming.

Quadratic functions

4.1 Quadratic functions

Definition 4.1: quadratic function

Given a square symmetric matrix

$$\underbrace{\begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{pmatrix}}_Q \in \mathbb{R}^{n \times n}$$

the associated (scalar) function

$$f : \mathbb{R}^n \rightarrow \mathbb{R}, \quad f : \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_x \mapsto \underbrace{\sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j}_{x' Q x} \quad (4.1)$$

is called a **quadratic function**.

- If the square matrix of a quadratic function is not symmetric, it can be made symmetric by replacing each off-diagonal pair with its average:

$$q_{ij} \leftarrow \frac{q_{ij} + q_{ji}}{2}, \quad q_{ji} \leftarrow \frac{q_{ij} + q_{ji}}{2}, \quad \forall i, j \in \{1, 2, \dots, n\}, \quad i < j$$

This is valid because for every $i, j \in \{1, 2, \dots, n\}$ with $i < j$, we have:

$$\begin{aligned} q_{ij} x_i x_j + q_{ji} x_j x_i &= q_{ij} x_i x_j + q_{ji} x_i x_j \\ &= (q_{ij} + q_{ji}) x_i x_j \\ &= \frac{1}{2}(q_{ij} + q_{ji}) x_i x_j + \frac{1}{2}(q_{ij} + q_{ji}) x_i x_j \end{aligned}$$

$$= \frac{q_{ij} + q_{ji}}{2} x_i x_j + \frac{q_{ij} + q_{ji}}{2} x_j x_i$$

Therefore, the quadratic functions defined by $\mathbf{Q}$ and $\frac{1}{2}(\mathbf{Q} + \mathbf{Q}')$ are equivalent, since they yield the same value for every $\mathbf{x} \in \mathbb{R}^n$.

Without loss of generality, we can assume that the matrix $\mathbf{Q}$ of a quadratic function is symmetric.

Example 4.1: equivalent matrices for quadratic functions

Consider the following non-symmetric 3×3 matrix:

$$\begin{pmatrix} 2 & -4 & 0 \\ 2 & 2 & 6 \\ 0 & -8 & 2 \end{pmatrix}$$

An equivalent quadratic function is obtained by using the following symmetric matrix:

$$\begin{pmatrix} 2 & \frac{2+(-4)}{2} & 0 \\ \frac{2+(-4)}{2} & 2 & \frac{-8+6}{2} \\ 0 & \frac{-8+6}{2} & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

Indeed, both matrices define the same function:

$$\begin{aligned} f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) &= 2x_1^2 - 4x_1x_2 + 2x_2x_1 + 2x_2^2 + 6x_2x_3 - 8x_3x_2 + 2x_3^2 \\ &= 2x_1^2 - x_1x_2 - x_2x_1 + 2x_2^2 - x_2x_3 - x_3x_2 + 2x_3^2 \\ &= 2x_1^2 + 2x_2^2 + 2x_3^2 - 2x_1x_2 - 2x_2x_3 \end{aligned}$$

4.2 Gradient, Hessian and level sets

Observation 4.2: gradient of a quadratic function

The **gradient** (vector of partial derivatives) of a quadratic function is:

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \frac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n (q_{1j} + q_{j1})x_j \\ \sum_{j=1}^n (q_{2j} + q_{j2})x_j \\ \vdots \\ \sum_{j=1}^n (q_{nj} + q_{jn})x_j \end{pmatrix} = \underbrace{\begin{pmatrix} 2 \sum_{j=1}^n q_{1j}x_j \\ 2 \sum_{j=1}^n q_{2j}x_j \\ \vdots \\ 2 \sum_{j=1}^n q_{nj}x_j \end{pmatrix}}_{2 \mathbf{Q} \mathbf{x}}$$

Proof. For every $k \in \{1, 2, \dots, n\}$, only the terms of the double sum with $i = k$ or $j = k$ contain the variable x_k : the terms $q_{kj} x_k x_j$ contribute $q_{kj} x_j$ to the derivative and the terms $q_{ik} x_i x_k$ contribute $q_{ik} x_i$ (the term $q_{kk} x_k^2$, with $i = j = k$, contributes $2 q_{kk} x_k$, once in each sum). Therefore:

$$\frac{\partial f(\mathbf{x})}{\partial x_k} = \sum_{j=1}^n q_{kj} x_j + \sum_{i=1}^n q_{ik} x_i = \sum_{j=1}^n (q_{kj} + q_{jk}) x_j = 2 \sum_{j=1}^n q_{kj} x_j$$

where the last equality holds since the matrix $\mathbf{Q}$ is symmetric ($q_{jk} = q_{kj}$). $\square$

Observation 4.3: Hessian of a quadratic function

The **Hessian** (matrix of second-order partial derivatives) of a quadratic function is:

$$\begin{aligned} \nabla^2 f(\mathbf{x}) &= \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix} \\ &= 2 \underbrace{\begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{12} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1n} & q_{2n} & \cdots & q_{nn} \end{pmatrix}}_{2 \mathbf{Q}} \end{aligned}$$

Proof. By the previous observation, for every $k \in \{1, 2, \dots, n\}$ we have $\frac{\partial f(\mathbf{x})}{\partial x_k} = 2 \sum_{j=1}^n q_{kj} x_j$, whose derivative with respect to x_h is $2 q_{kh}$, for every $h \in \{1, 2, \dots, n\}$. $\square$

Definition 4.4: level sets of a quadratic function

The **level sets** of a quadratic function (the sets of points in which the function is **constant**) are:

$$\left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x}' \mathbf{Q} \mathbf{x} = k \right\}, \quad \forall k \in \mathbb{R}$$

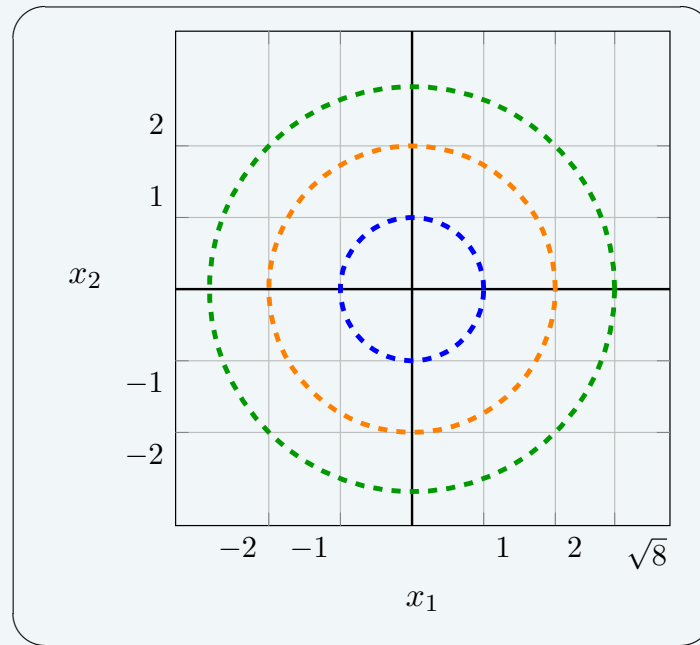
- Depending on the matrix $\mathbf{Q}$, the nonempty level sets can be **ellipsoids** (ellipses in $\mathbb{R}^2$), **hyperboloids** (hyperbolas in $\mathbb{R}^2$) or, when some directions do not change the value of the function, **cylinders** (pairs of parallel lines in $\mathbb{R}^2$). For $k = 0$ there are also degenerate cases: the single point $\mathbf{0}$ in place of an ellipsoid, two lines crossing at the origin in $\mathbb{R}^2$ (a cone in $\mathbb{R}^n$) in place of a hyperboloid, and a single line through the origin in $\mathbb{R}^2$ (a linear subspace in $\mathbb{R}^n$) in place of a cylinder. The graph of the quadratic functions of the following examples is a **paraboloid**.

For a quadratic function with symmetric matrix $\mathbf{Q}$, the gradient is $\nabla f(\mathbf{x}) = 2 \mathbf{Q} \mathbf{x}$ and the Hessian is the constant matrix $\nabla^2 f(\mathbf{x}) = 2 \mathbf{Q}$. At every point where the gradient is nonzero, the gradient is perpendicular to the level set through that point.

Example 4.2: level sets of quadratic functions

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, the associated quadratic function is:

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto (x_1 \ x_2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1^2 + x_2^2$$



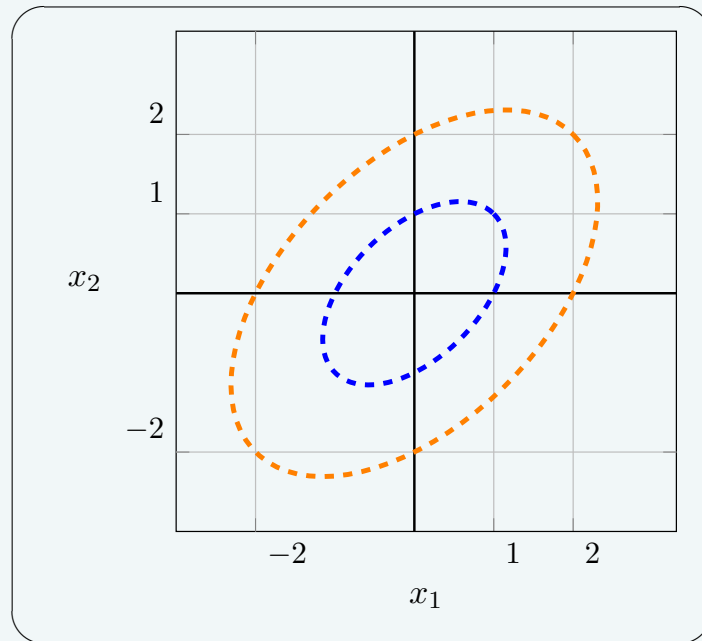
The figure shows:

- Blue circle: level set $x_1^2 + x_2^2 = 1$ (radius 1)
- Orange circle: level set $x_1^2 + x_2^2 = 4$ (radius 2)
- Green circle: level set $x_1^2 + x_2^2 = 8$ (radius $\sqrt{8}$)

Example 4.3: level sets of quadratic functions

With $\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, the associated quadratic function is:

$$\begin{aligned}
 f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\mapsto \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\
 &= 2x_1^2 - x_1x_2 - x_2x_1 + 2x_2^2 \\
 &= 2x_1^2 - 2x_1x_2 + 2x_2^2
 \end{aligned}$$



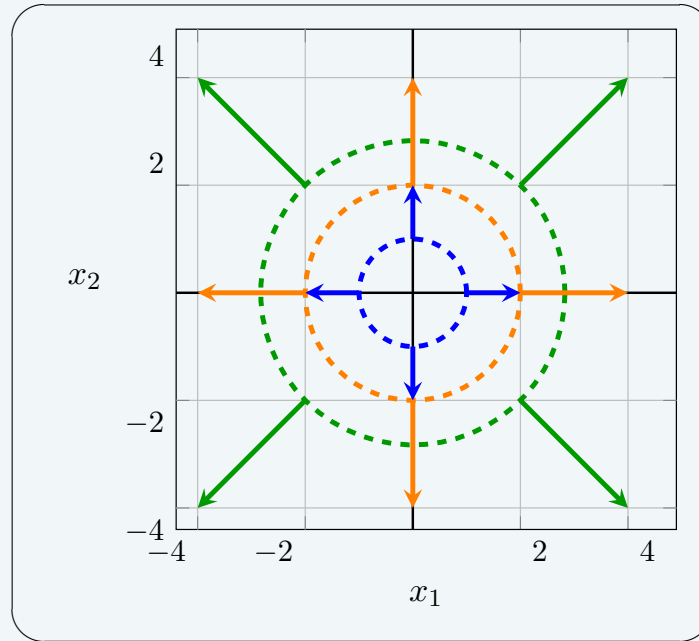
The figure shows:

- Blue ellipse: level set $2x_1^2 - 2x_1x_2 + 2x_2^2 = 2$
- Orange ellipse: level set $2x_1^2 - 2x_1x_2 + 2x_2^2 = 8$

Example 4.4: gradients of quadratic functions

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, the gradient of the quadratic function is:

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



The figure shows:

- Blue vectors (gradients with norm 2 on level set $x_1^2 + x_2^2 = 1$):

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

- Orange vectors (gradients with norm 4 on level set $x_1^2 + x_2^2 = 4$):

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

- Green vectors (gradients with norm $4\sqrt{2}$ on level set $x_1^2 + x_2^2 = 8$):

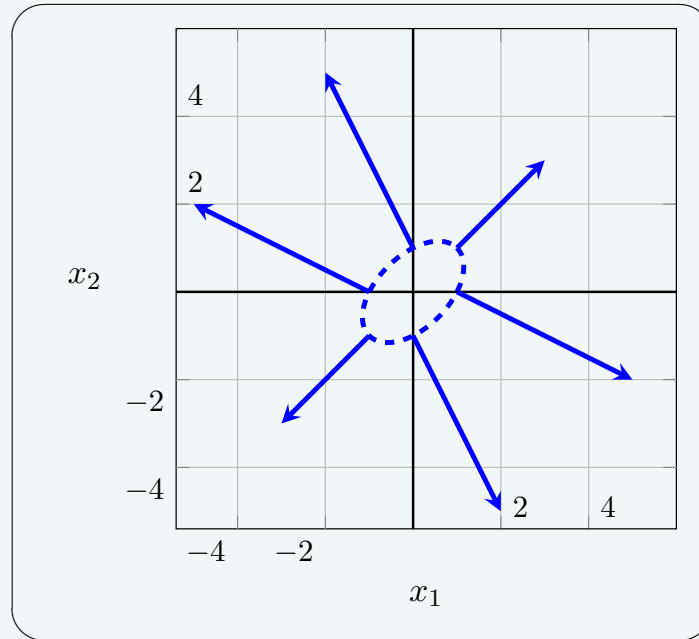
$$\begin{aligned}\nabla f\left(\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) &= \begin{pmatrix} 4 \\ 4 \end{pmatrix}, & \nabla f\left(\begin{pmatrix} -2 \\ -2 \end{pmatrix}\right) &= \begin{pmatrix} -4 \\ -4 \end{pmatrix}, \\ \nabla f\left(\begin{pmatrix} -2 \\ 2 \end{pmatrix}\right) &= \begin{pmatrix}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

Each gradient is perpendicular to the level set through its point.

Example 4.5: gradients of quadratic functions

With $Q = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, the gradient of the quadratic function is:

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



The figure shows:

- Blue ellipse: level set $2x_1^2 - 2x_1x_2 + 2x_2^2 = 2$
- Blue vectors (gradients on the level set with norm $2\sqrt{2}$ at diagonal points and norm $2\sqrt{5}$ at axis points):

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

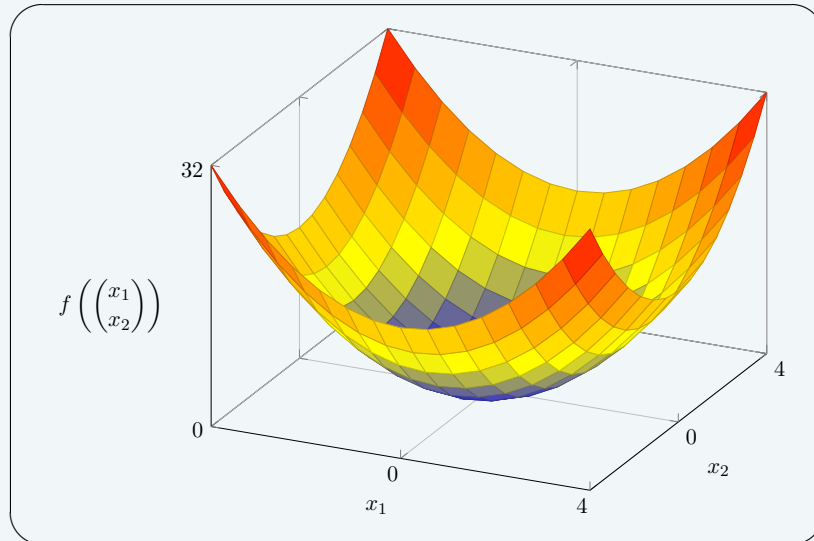
$$\nabla f \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \quad \nabla f \left(\begin{pmatrix} -1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} -4 \\ -2 \end{pmatrix},$$

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

Also in this case, each gradient is perpendicular to the level set through its point.

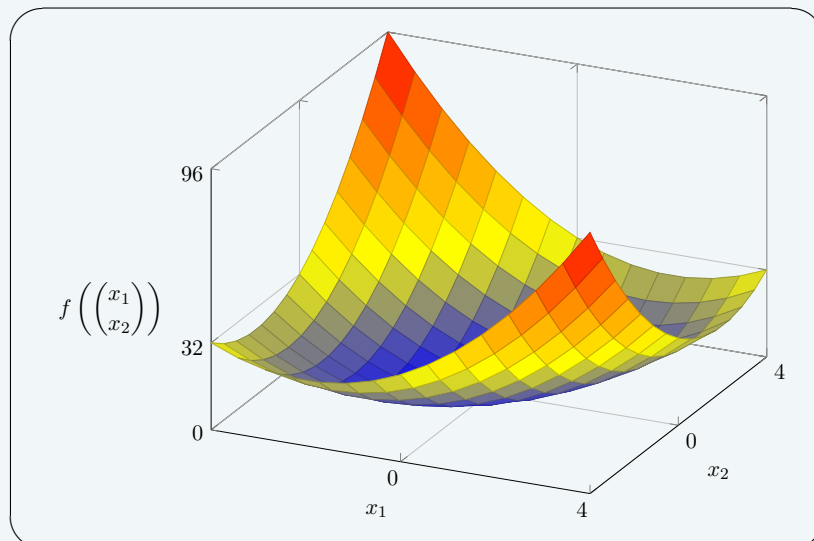
Example 4.6: graphs of quadratic functions

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, we have $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + x_2^2$. Its graph is:



Example 4.7: graphs of quadratic functions

With $\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, we have $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = 2x_1^2 - 2x_1x_2 + 2x_2^2$. Its graph is:



4.3 Convexity of quadratic functions

Definition 4.5: positive semidefinite matrix

A square symmetric matrix $Q \in \mathbb{R}^{n \times n}$ is **positive semidefinite** if:

$$x'Qx \geq 0, \quad \forall x \in \mathbb{R}^n$$

Proposition 4.6

For every square matrix Q , if Q is positive semidefinite then the associated quadratic function is convex in its domain.

Proof. If the matrix Q is positive semidefinite, for every pair of points $p, w \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} & f(\lambda p + (1 - \lambda) w) \\ &= (\lambda p + (1 - \lambda) w)' Q (\lambda p + (1 - \lambda) w) \\ &= \lambda^2 p'Qp + (1 - \lambda)^2 w'Qw \\ &\quad + \lambda(1 - \lambda) p'Qw + \lambda(1 - \lambda) w'Qp \\ &= \lambda p'Qp + (1 - \lambda) w'Qw \\ &\quad - \lambda(1 - \lambda) (p'Qw + w'Qp - p'Qw - w'Qp) \\ &\quad (\text{since } \lambda^2 = \lambda - \lambda(1 - \lambda) \text{ and } (1 - \lambda)^2 = (1 - \lambda) - \lambda(1 - \lambda)) \\ &= \lambda p'Qp + (1 - \lambda) w'Qw - \lambda(1 - \lambda) (p - w)'Q(p - w) \\ &\leq \lambda p'Qp + (1 - \lambda) w'Qw \\ &\quad (\text{since } Q \text{ is positive semidefinite and } \lambda(1 - \lambda) \geq 0) \\ &= \lambda f(p) + (1 - \lambda) f(w) \end{aligned}$$

Therefore,

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

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

- We have shown that:

$$Q \text{ positive semidefinite} \implies \text{convex quadratic function}$$

so “ Q positive semidefinite” is a sufficient condition for “convex quadratic function” and “convex quadratic function” is a necessary condition for “ Q positive semidefinite”.

- The implication works also in the other direction, as stated by the following proposition.

Proposition 4.7

For every quadratic function, if the function is convex in its domain then the associated square matrix $\mathbf{Q}$ is positive semidefinite.

Proof. If the quadratic function is convex, for every pair of column vectors $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ and every scalar $\lambda \in (0, 1)$, we have:

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

where the left-hand side is:

$$\begin{aligned} & (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})' \mathbf{Q} (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\ &= \lambda^2 \mathbf{p}' \mathbf{Q} \mathbf{p} + (1 - \lambda)^2 \mathbf{w}' \mathbf{Q} \mathbf{w} + \lambda (1 - \lambda) \mathbf{p}' \mathbf{Q} \mathbf{w} + \lambda (1 - \lambda) \mathbf{w}' \mathbf{Q} \mathbf{p} \end{aligned}$$

This implies, rearranging and simplifying, that:

$$\begin{aligned} & \lambda (1 - \lambda) \mathbf{p}' \mathbf{Q} \mathbf{p} + \lambda (1 - \lambda) \mathbf{w}' \mathbf{Q} \mathbf{w} \\ & - \lambda (1 - \lambda) \mathbf{p}' \mathbf{Q} \mathbf{w} - \lambda (1 - \lambda) \mathbf{w}' \mathbf{Q} \mathbf{p} \geq 0 \end{aligned}$$

which, since $\lambda (1 - \lambda) > 0$, implies:

$$\underbrace{\mathbf{p}' \mathbf{Q} \mathbf{p} + \mathbf{w}' \mathbf{Q} \mathbf{w} - \mathbf{p}' \mathbf{Q} \mathbf{w} - \mathbf{w}' \mathbf{Q} \mathbf{p}}_{=(\mathbf{p}-\mathbf{w})' \mathbf{Q} (\mathbf{p}-\mathbf{w})} \geq 0$$

Every vector $\mathbf{x} \in \mathbb{R}^n$ can be written as $\mathbf{x} = \mathbf{p} - \mathbf{w}$ (for instance, with $\mathbf{p} = \mathbf{x}$ and $\mathbf{w} = \mathbf{0}$). Therefore, $\mathbf{x}' \mathbf{Q} \mathbf{x} \geq 0$ for every $\mathbf{x} \in \mathbb{R}^n$, and the square matrix $\mathbf{Q}$ is positive semidefinite. $\square$

- We have shown that:

$$\text{convex quadratic function} \implies \mathbf{Q} \text{ positive semidefinite}$$

so “convex quadratic function” is a sufficient condition for “ $\mathbf{Q}$ positive semidefinite” and “ $\mathbf{Q}$ positive semidefinite” is a necessary condition for “convex quadratic function”.

- So we finally have:

$$\mathbf{Q} \text{ positive semidefinite} \iff \text{convex quadratic function}$$

so “ $\mathbf{Q}$ positive semidefinite” is a necessary and sufficient condition for “convex quadratic function” and vice versa.

Corollary 4.8

For every square matrix $\mathbf{Q}$, the associated quadratic function is convex in its domain if and only if $\mathbf{Q}$ is positive semidefinite.

- With $\mathbf{Q}$ equal to the identity matrix, the quadratic function is $f(\mathbf{x}) = \mathbf{x}'\mathbf{x} = \|\mathbf{x}\|_2^2$, the squared Euclidean norm, whose convexity is proved in the chapter «Convex and concave functions» directly from the definition. The identity matrix is positive semidefinite, since $\mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 + \dots + x_n^2 \geq 0$, in agreement with the corollary.
- When $\mathbf{x}'\mathbf{Q}\mathbf{x}$ can be written as a sum of squares of linear functions, as in the following example, the matrix $\mathbf{Q}$ is positive semidefinite, and the convexity can also be proved directly from the definition, with the inequality $(\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2$ of the chapter «Convex and concave functions».

Example 4.8: a convex quadratic function

With $\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, for every $\mathbf{x} \in \mathbb{R}^2$ we have:

$$\mathbf{x}'\mathbf{Q}\mathbf{x} = 2x_1^2 - 2x_1x_2 + 2x_2^2 = x_1^2 + x_2^2 + (x_1 - x_2)^2 \geq 0$$

Therefore, the matrix $\mathbf{Q}$ is positive semidefinite and, by the corollary, the quadratic function of the previous examples is convex in its domain.

The same conclusion follows from the definition: for every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} & f(\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}) \\ &= (\lambda p_1 + (1 - \lambda)w_1)^2 + (\lambda p_2 + (1 - \lambda)w_2)^2 \\ & \quad + (\lambda(p_1 - p_2) + (1 - \lambda)(w_1 - w_2))^2 \\ &\leq \lambda p_1^2 + (1 - \lambda)w_1^2 + \lambda p_2^2 + (1 - \lambda)w_2^2 \\ & \quad + \lambda(p_1 - p_2)^2 + (1 - \lambda)(w_1 - w_2)^2 \\ & \quad \text{(by the inequality of the chapter «Convex and concave functions»,} \\ & \quad \text{three times)} \\ &= \lambda(p_1^2 + p_2^2 + (p_1 - p_2)^2) \\ & \quad + (1 - \lambda)(w_1^2 + w_2^2 + (w_1 - w_2)^2) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w}) \end{aligned}$$

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.

Example 4.9: a quadratic function that is not convex

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, the quadratic function is

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

and its nonempty level sets with $k \neq 0$ are hyperbolas. The matrix $\mathbf{Q}$ is not positive semidefinite, since:

$$\begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1 < 0$$

By the corollary, the function is not convex in its domain. Indeed, with

$$\mathbf{p} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \lambda = \frac{1}{2}$$

the inequality of the definition of convex function does not hold:

$$f\left(\frac{1}{2}\mathbf{p} + \frac{1}{2}\mathbf{w}\right) = f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = 0 > -1 = \frac{1}{2}f(\mathbf{p}) + \frac{1}{2}f(\mathbf{w})$$

4.4 Quadratic functions with a linear part

- In optimization problems we often find functions that are the sum of a quadratic function, a linear function and a constant: given a square symmetric matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$, a vector $\mathbf{c} \in \mathbb{R}^n$ and a scalar $d \in \mathbb{R}$,

$$g: \mathbb{R}^n \rightarrow \mathbb{R}, \quad g: \mathbf{x} \mapsto \mathbf{x}'\mathbf{Q}\mathbf{x} + \mathbf{c}'\mathbf{x} + d = \sum_{i=1}^n \sum_{j=1}^n q_{ij}x_i x_j + \sum_{j=1}^n c_j x_j + d$$

- Since the derivative of $\mathbf{c}'\mathbf{x} + d$ with respect to x_k is c_k , the gradient and the Hessian are:

$$\nabla g(\mathbf{x}) = 2\mathbf{Q}\mathbf{x} + \mathbf{c}, \quad \nabla^2 g(\mathbf{x}) = 2\mathbf{Q}$$

- The function $h(\mathbf{x}) = \mathbf{c}'\mathbf{x} + d$ is convex (and concave) in its domain, since the definition holds with equality: for every pair of points $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ and every $\lambda \in [0, 1]$, we have:

$$\begin{aligned} h(\lambda\mathbf{p} + (1-\lambda)\mathbf{w}) &= \mathbf{c}'(\lambda\mathbf{p} + (1-\lambda)\mathbf{w}) + d \\ &= \lambda\mathbf{c}'\mathbf{p} + (1-\lambda)\mathbf{c}'\mathbf{w} + \lambda d + (1-\lambda)d \\ &\quad (\text{since } \lambda + (1-\lambda) = 1) \\ &= \lambda h(\mathbf{p}) + (1-\lambda)h(\mathbf{w}) \end{aligned}$$

Therefore, both inequalities of the definitions of convex and concave function hold with equality. By nature, the domain $\mathbb{R}^n$ is a convex set. Accordingly, the function h is convex

and concave in its domain. Since the sum of convex functions is convex (chapter «Convex and concave functions»), if the matrix $\mathbf{Q}$ is positive semidefinite, the function g is convex in its domain. Conversely, if g is convex, the quadratic function $\mathbf{x}'\mathbf{Q}\mathbf{x} = g(\mathbf{x}) + (-\mathbf{c}'\mathbf{x} - d)$ is convex, as the sum of two convex functions, and the matrix $\mathbf{Q}$ is positive semidefinite by the corollary.

A quadratic function $\mathbf{x}'\mathbf{Q}\mathbf{x}$ is convex in its domain if and only if the matrix $\mathbf{Q}$ is positive semidefinite. Adding a linear function and a constant does not change its convexity.