



SAPIENZA
UNIVERSITÀ DI ROMA

Ottimizzazione convessa

Volume 1 — Preliminari: insiemi convessi, funzioni e funzioni convesse

Combinazioni convesse e insiemi convessi, funzioni di più variabili (gradiente, Hessiano, insiemi di livello), norme, funzioni convesse e concave, funzioni quadratiche e loro convessità. Con esempi svolti.

Fabio Furini

fabiofurini.github.io

Dipartimento di Ingegneria informatica, automatica e gestionale — Sapienza Università di Roma

Versione online e materiale scaricabile:
fabiofurini.github.io/ottimizzazione-convessa/
github.com/fabiofurini/ottimizzazione-convessa

Anno accademico 2026–2027

Indice

1	Insiemi convessi	2
1.1	Combinazione convessa	2
1.2	Insiemi convessi	6
1.3	Esempi di insiemi convessi	7
1.4	Proprietà degli insiemi convessi	12
2	Funzioni di più variabili	14
2.1	Funzioni di più variabili	14
2.2	Gradiente, matrice Hessiana e insiemi di livello	15
2.3	Esempi di funzioni non quadratiche	18
3	Funzioni convesse e concave	21
3.1	Funzioni convesse	21
3.2	Funzioni concave	22
3.3	Esempi di funzioni convesse e concave	23
3.4	Funzioni convesse definite da norme	29
3.5	Proprietà delle funzioni convesse e concave	34
4	Funzioni quadratiche	40
4.1	Funzioni quadratiche	40
4.2	Gradiente, matrice Hessiana e insiemi di livello	42
4.3	Convessità delle funzioni quadratiche	49
4.4	Funzioni quadratiche con una parte lineare	52

Insiemi convessi

1.1 Combinazione convessa

Definizione 1.1: combinazione convessa

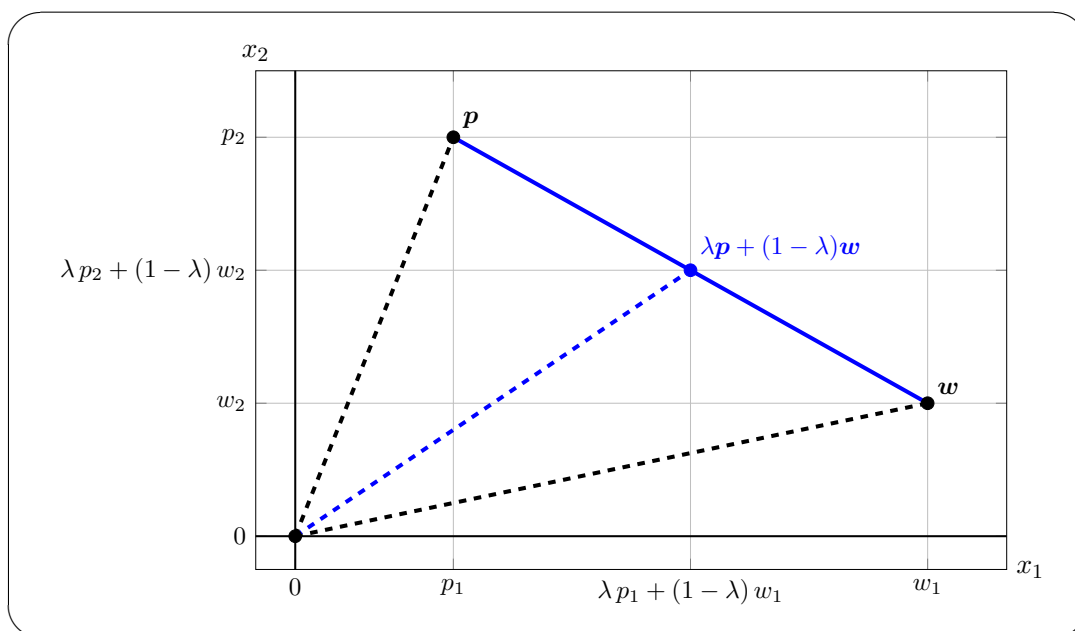
Dati due punti

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

la loro **combinazione convessa** con uno scalare $\lambda \in [0, 1]$ è il punto

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

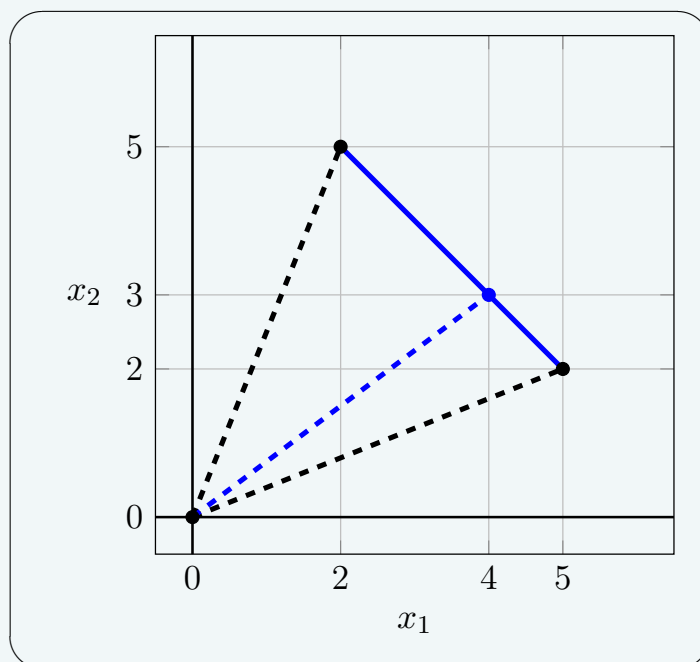
- Le combinazioni convesse per tutti i valori $\lambda \in [0, 1]$ sono i punti del segmento che congiunge i due punti $\mathbf{p}$ e $\mathbf{w}$. In $\mathbb{R}^2$, il segmento blu della figura seguente:



Esempio 1.1: combinazione convessa

Consideriamo i punti $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ e $\begin{pmatrix} 5 \\ 2 \end{pmatrix}$ in $\mathbb{R}^2$. La loro combinazione convessa con $\lambda = \frac{1}{3}$ è il punto:

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



Definizione 1.2: combinazione convessa di m punti

Dati m punti

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

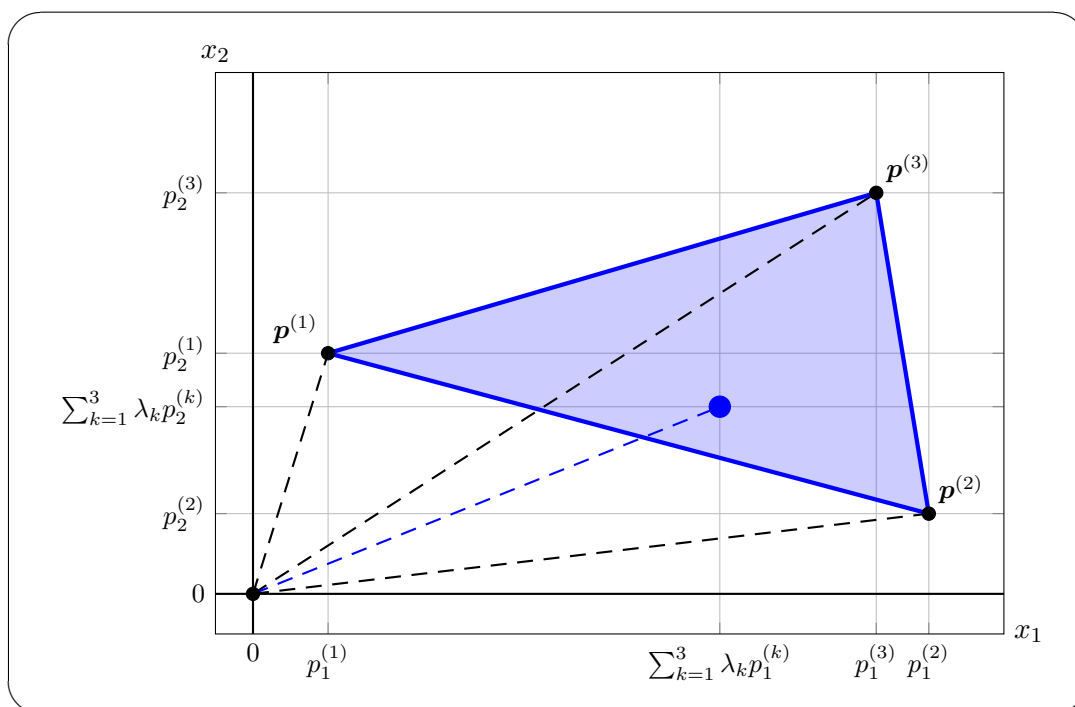
la loro **combinazione convessa** con scalari $\lambda_1, \lambda_2, \dots, \lambda_m \geq 0$ tali che

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

è il punto

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

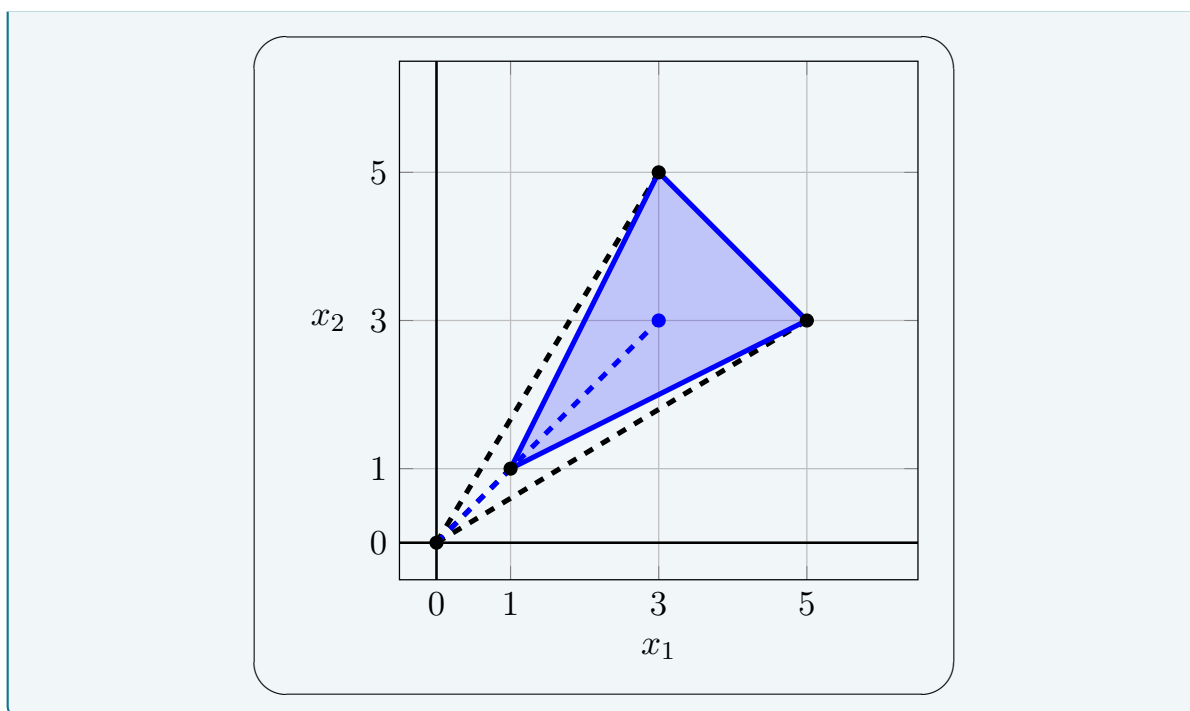
- Le combinazioni convesse che si ottengono con tutti gli scalari ammessi $\lambda_1, \lambda_2, \dots, \lambda_m \geq 0$ sono i punti dell'**inviluppo convesso** degli m punti.
- In $\mathbb{R}^2$ con $m = 3$ punti, l'inviluppo convesso è formato da tutti i punti interni (e sul bordo) del triangolo che ha per vertici i tre punti: la regione blu della figura seguente, dove il punto blu è la combinazione convessa con $\lambda_1 = \frac{1}{3}$, $\lambda_2 = \frac{1}{2}$ e $\lambda_3 = \frac{1}{6}$:



Esempio 1.2: combinazione convessa di 3 punti

Consideriamo i tre punti $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 5 \\ 3 \end{pmatrix}$ e $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ in $\mathbb{R}^2$. La loro combinazione convessa con $\lambda_1 = \lambda_2 = \lambda_3 = \frac{1}{3}$ è il punto:

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



1.2 Insiemi convessi

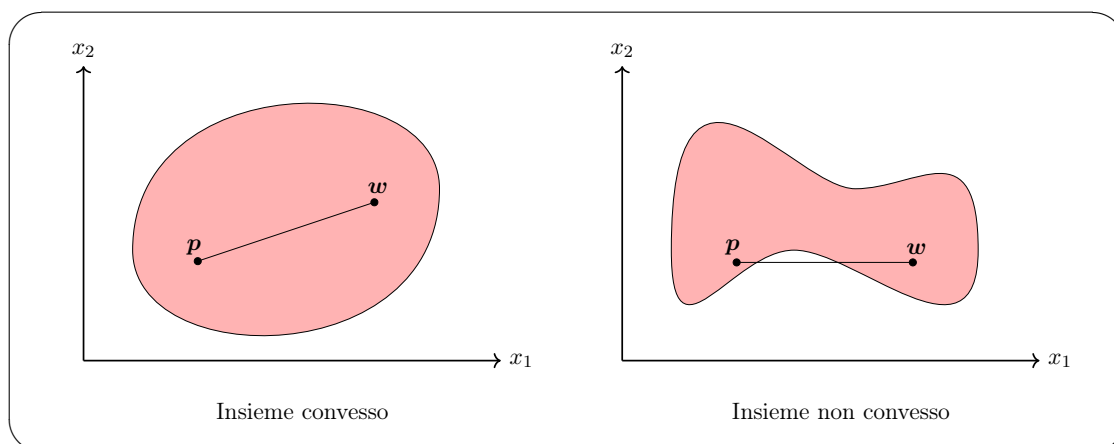
Definizione 1.3: insieme convesso

Un insieme $A \subseteq \mathbb{R}^n$ è **convesso** se

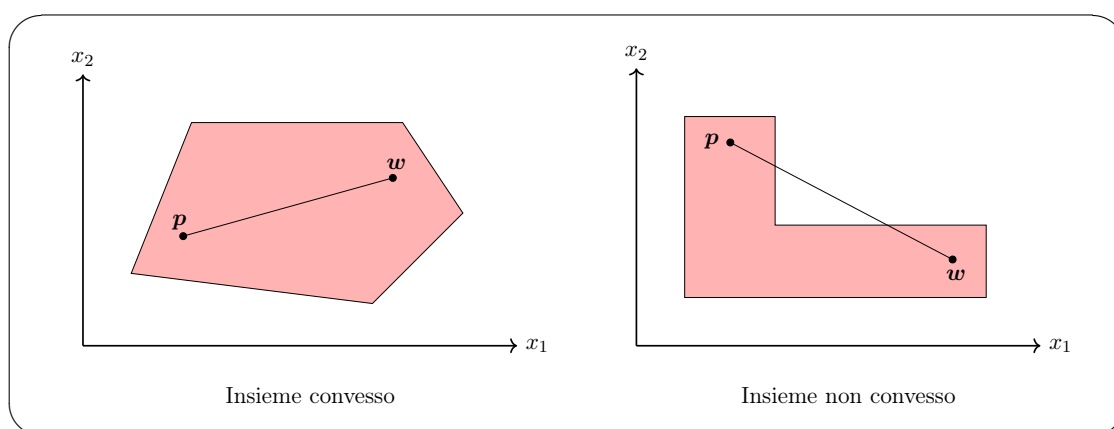
$$\forall \mathbf{p}, \mathbf{w} \in A \text{ e } \forall \lambda \in [0, 1] \text{ abbiamo } \underbrace{\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}}_{\text{combinazione convessa}} \in A$$

Un sottoinsieme dello spazio n -dimensionale è convesso se, per ogni coppia di suoi punti, ogni combinazione convessa dei due punti appartiene al sottoinsieme. In altre parole, se tutti i punti del segmento che congiunge due punti qualsiasi del sottoinsieme appartengono al sottoinsieme.

- In $\mathbb{R}^2$ (lo spazio a 2 dimensioni), un esempio di insieme convesso e un esempio di insieme non convesso sono dati dalla figura seguente, prima con bordi curvi:



- Lo stesso con insiemi a lati rettilinei: a sinistra un poligono convesso, a destra uno non convesso.



1.3 Esempi di insiemi convessi

Osservazione 1.4

Per ogni vettore $\mathbf{c} \in \mathbb{R}^n$ e ogni scalare $r \in \mathbb{R}_{>0}$, l'insieme

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

è convesso.

È l'insieme dei punti di $\mathbb{R}^n$ la cui somma degli scarti assoluti da $\mathbf{c}$ è al più r . In $\mathbb{R}^2$ è un rombo (un quadrato ruotato di 45°), e in $\mathbb{R}^3$ è un ottaedro regolare.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w}$ dell'insieme e per ogni $\lambda \in [0, 1]$, abbiamo:

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

$$\begin{aligned}
 & \text{(per la disuguaglianza triangolare)} \\
 &= |\lambda| \|\mathbf{p} - \mathbf{c}\|_1 + |1 - \lambda| \|\mathbf{w} - \mathbf{c}\|_1 \\
 & \text{(per l'omogeneità assoluta)} \\
 &= \lambda \|\mathbf{p} - \mathbf{c}\|_1 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_1 \\
 & \text{(poiché } \lambda \geq 0, 1 - \lambda \geq 0 \text{)} \\
 &\leq \lambda r + (1 - \lambda)r \\
 & \text{(poiché } \|\mathbf{p} - \mathbf{c}\|_1 \leq r \text{ e } \|\mathbf{w} - \mathbf{c}\|_1 \leq r \text{)} \\
 &= r
 \end{aligned}$$

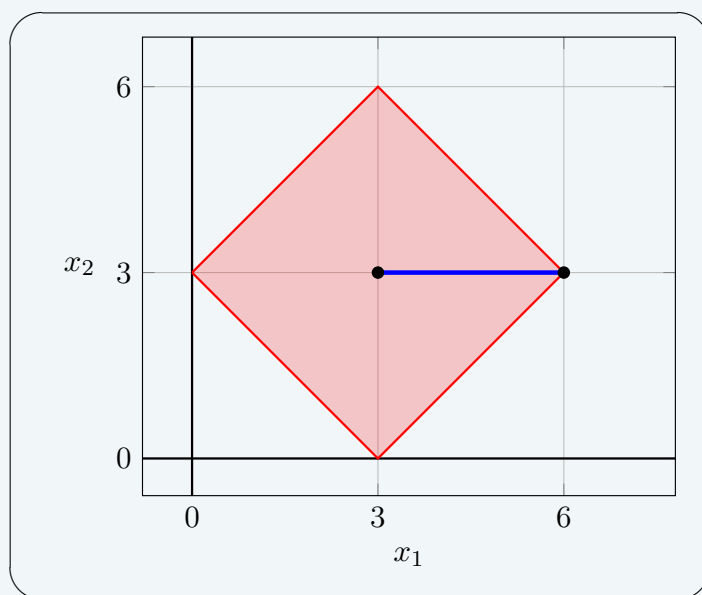
Quindi, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}$ appartiene all'insieme e l'insieme è convesso. $\square$

Esempio 1.3: rombo in $\mathbb{R}^2$

Per esempio, in $\mathbb{R}^2$, con $r = 3$ e $\mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, abbiamo il seguente insieme convesso:

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

I punti di questo insieme sono tutti quelli interni al rombo (quadrato ruotato) o sul suo bordo, con centro in $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$, con vertici in $\begin{pmatrix} 0 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 6 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 6 \end{pmatrix}$, rappresentati dalla regione ombreggiata in rosso nella figura seguente.



Osservazione 1.5

Per ogni vettore $\mathbf{c} \in \mathbb{R}^n$ e ogni scalare $r \in \mathbb{R}_{>0}$, l'insieme

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

è convesso.

È l'insieme dei punti di $\mathbb{R}^n$ contenuti nella palla chiusa di raggio r e centro $\mathbf{c}$. In $\mathbb{R}^2$ è un disco, e in $\mathbb{R}^3$ è una sfera piena.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w}$ dell'insieme e per ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_2 &= \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2 \\ &\leq \|\lambda(\mathbf{p} - \mathbf{c})\|_2 + \|(1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2 \\ &\quad \text{(per la disuguaglianza triangolare)} \\ &= |\lambda| \|\mathbf{p} - \mathbf{c}\|_2 + |1 - \lambda| \|\mathbf{w} - \mathbf{c}\|_2 \\ &\quad \text{(per l'omogeneità assoluta)} \\ &= \lambda \|\mathbf{p} - \mathbf{c}\|_2 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_2 \\ &\quad \text{(poiché } \lambda \geq 0, 1 - \lambda \geq 0) \\ &\leq \lambda r + (1 - \lambda)r \\ &\quad \text{(poiché } \|\mathbf{p} - \mathbf{c}\|_2 \leq r \text{ e } \|\mathbf{w} - \mathbf{c}\|_2 \leq r) \\ &= r \end{aligned}$$

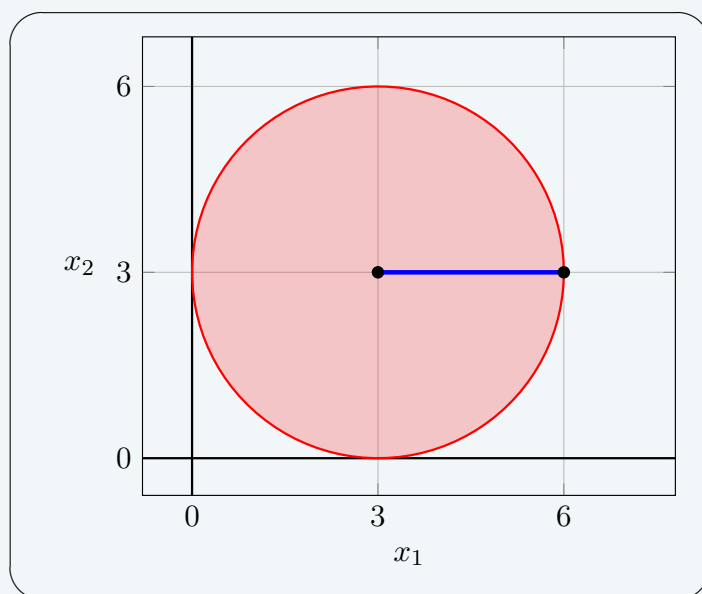
Quindi, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}$ appartiene all'insieme e l'insieme è convesso. □

Esempio 1.4: palla chiusa in $\mathbb{R}^2$

Per esempio, in $\mathbb{R}^2$, con $r = 3$ e $\mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, abbiamo il seguente insieme convesso:

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

I punti di questo insieme sono tutti quelli interni al disco di raggio 3 o sul suo bordo, con centro in $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$, rappresentati dalla regione ombreggiata in rosso nella figura seguente.

**Osservazione 1.6**

Per ogni vettore $\mathbf{c} \in \mathbb{R}^n$ e ogni scalare $r \in \mathbb{R}_{>0}$, l'insieme

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

è convesso.

È l'insieme dei punti di $\mathbb{R}^n$ il cui massimo scarto assoluto da $\mathbf{c}$ è al più r . In $\mathbb{R}^2$ è un quadrato con i lati paralleli agli assi, di lato $2r$ e centro $\mathbf{c}$, e in $\mathbb{R}^3$ è un cubo.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w}$ dell'insieme e per ogni $\lambda \in [0, 1]$, abbiamo:

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

$$\begin{aligned}
 &\leq \|\lambda(\mathbf{p} - \mathbf{c})\|_\infty + \|(1 - \lambda)(\mathbf{w} - \mathbf{c})\|_\infty \\
 &\quad (\text{per la disuguaglianza triangolare}) \\
 &= |\lambda| \|\mathbf{p} - \mathbf{c}\|_\infty + |1 - \lambda| \|\mathbf{w} - \mathbf{c}\|_\infty \\
 &\quad (\text{per l'omogeneità assoluta}) \\
 &= \lambda \|\mathbf{p} - \mathbf{c}\|_\infty + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_\infty \\
 &\quad (\text{poiché } \lambda \geq 0, 1 - \lambda \geq 0) \\
 &\leq \lambda r + (1 - \lambda)r \\
 &\quad (\text{poiché } \|\mathbf{p} - \mathbf{c}\|_\infty \leq r \text{ e } \|\mathbf{w} - \mathbf{c}\|_\infty \leq r) \\
 &= r
 \end{aligned}$$

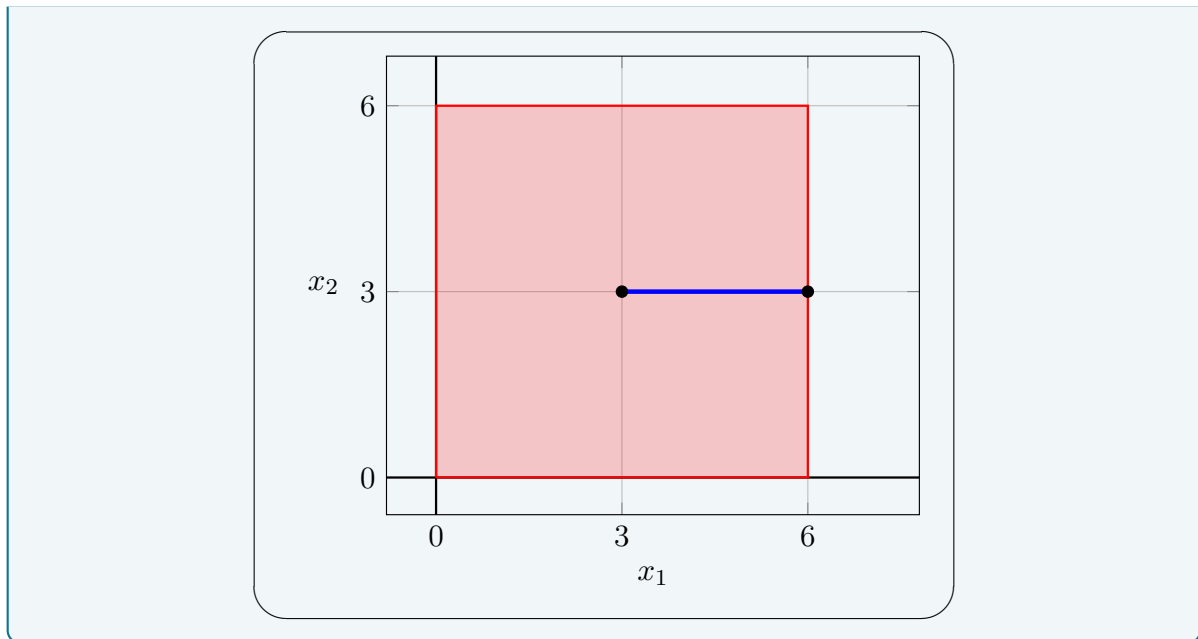
Quindi, $\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}$ appartiene all'insieme e l'insieme è convesso. $\square$

Esempio 1.5: quadrato in $\mathbb{R}^2$

Per esempio, in $\mathbb{R}^2$, con $r = 3$ e $\mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$, abbiamo il seguente insieme convesso:

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

I punti di questo insieme sono tutti quelli interni al quadrato con i lati paralleli agli assi, di lato 6, o sul suo bordo, con centro in $\begin{pmatrix} 3 \\ 3 \end{pmatrix}$, con vertici in $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 6 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 6 \\ 6 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 6 \end{pmatrix}$, rappresentati dalla regione ombreggiata in rosso nella figura seguente.



1.4 Proprietà degli insiemi convessi

Osservazione 1.7

Per ogni famiglia di m insiemi convessi $A_i \subseteq \mathbb{R}^n$ con $i \in \{1, 2, \dots, m\}$, l'insieme

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

è convesso.

È l'insieme dei punti di $\mathbb{R}^n$ che appartengono contemporaneamente a tutti gli insiemi $A_1, A_2, \dots, A_m$.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \bigcap_{i=1}^m A_i$ e per ogni $\lambda \in [0, 1]$, abbiamo:

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

(per definizione di intersezione)

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

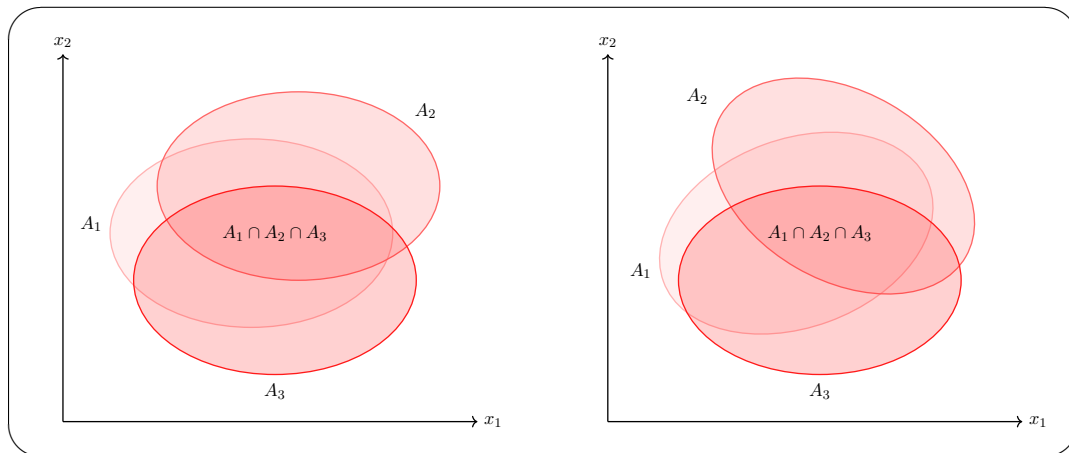
(poiché ogni A_i è convesso)

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

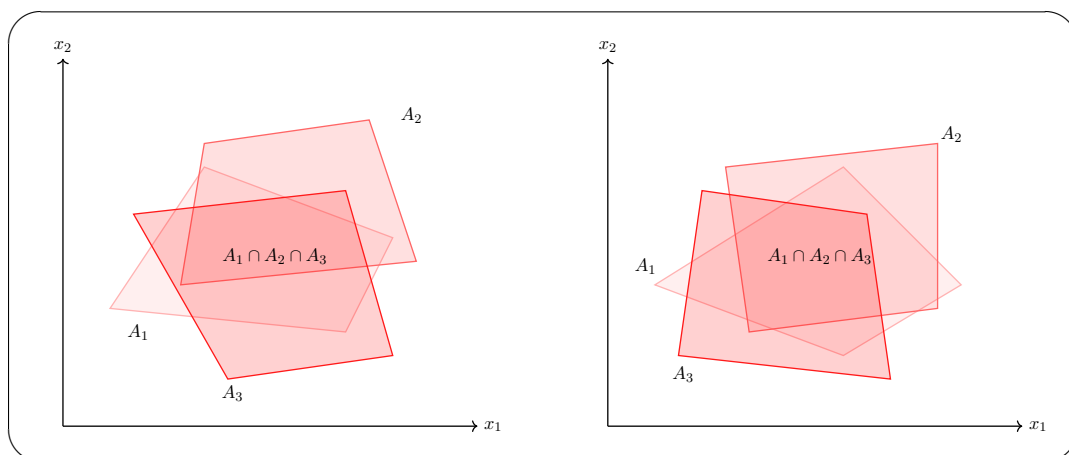
(per definizione di intersezione)

Quindi, $\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} \in \bigcap_{i=1}^m A_i$ e l'insieme è convesso. □

- In $\mathbb{R}^2$ (lo spazio a 2 dimensioni), esempi di intersezione di tre insiemi convessi A_1, A_2 e A_3 sono dati dalle figure seguenti, prima con tre ellissi:



- Lo stesso con tre quadrilateri convessi (a lati rettilinei, come i poliedri della programmazione lineare, che sono tutti convessi):



Funzioni di più variabili

2.1 Funzioni di più variabili

Definizione 2.1: funzione di più variabili reali

La funzione:

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

si chiama **funzione (scalare) di n variabili reali** e $D \subseteq \mathbb{R}^n$ è il suo **dominio**.

- Una funzione, indicata di solito con f , di n variabili reali prende in ingresso un vettore di n numeri reali, rappresentato di solito dal vettore (colonna) delle variabili $\mathbf{x}$, e restituisce un numero reale, il valore della funzione, indicato di solito con $f(\mathbf{x})$.
- I vettori si scrivono in grassetto e sono vettori colonna; il trasposto di un vettore $\mathbf{x}$ è il vettore riga $\mathbf{x}' = (x_1 \ x_2 \ \dots \ x_n)$.

Esempio 2.1: una funzione di tre variabili

Consideriamo la funzione:

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

Nel punto $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ il valore della funzione è:

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

2.2 Gradiente, matrice Hessiana e insiemi di livello

Definizione 2.2: gradiente

Data $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, il **gradiente** $\nabla f(\mathbf{x})$ di f in $\mathbf{x}$ è il vettore (colonna) delle derivate parziali prime:

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

Il gradiente $\nabla f(\mathbf{x})$ (se non è nullo) fornisce la direzione di massima crescita (massima derivata direzionale fra le direzioni unitarie), mentre $-\nabla f(\mathbf{x})$ fornisce la direzione di massima decrescita.

Definizione 2.3: matrice Hessiana

Data una funzione $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, la **matrice Hessiana** $\nabla^2 f(\mathbf{x})$ di f in $\mathbf{x}$ è la matrice delle derivate parziali seconde:

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

- L'elemento di riga i e colonna j della matrice Hessiana si ottiene derivando $\frac{\partial f(\mathbf{x})}{\partial x_j}$ rispetto a x_i .
- Se le derivate parziali seconde di f sono continue (come per tutte le funzioni di queste dispense), l'ordine di derivazione non conta, $\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i}$, e la matrice Hessiana è una matrice **simmetrica**.

La matrice Hessiana $\nabla^2 f(\mathbf{x})$ descrive la curvatura locale di f attorno a $\mathbf{x}$, cioè come varia il gradiente allontanandosi da $\mathbf{x}$.

Definizione 2.4: insieme di livello

Data una funzione $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e una costante $k \in \mathbb{R}$, l'**insieme di livello** $L(k)$ di f al livello k è:

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

Un insieme di livello $L(k)$ è l'insieme dei punti in cui la funzione assume il valore costante k . Il gradiente $\nabla f(\mathbf{x})$ (se non è nullo) è perpendicolare all'insieme di livello passante per $\mathbf{x}$ e punta verso i valori crescenti di f .

Esempio 2.2: gradiente, matrice Hessiana e insiemi di livello di una funzione quadratica

Consideriamo la funzione quadratica:

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

- Il **gradiente** della funzione è:

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

Per esempio:

- nel punto $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ abbiamo

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

- nel punto $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ abbiamo

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

- La **matrice Hessiana** della funzione è:

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

Si noti che la matrice Hessiana è costante (non dipende dal punto).

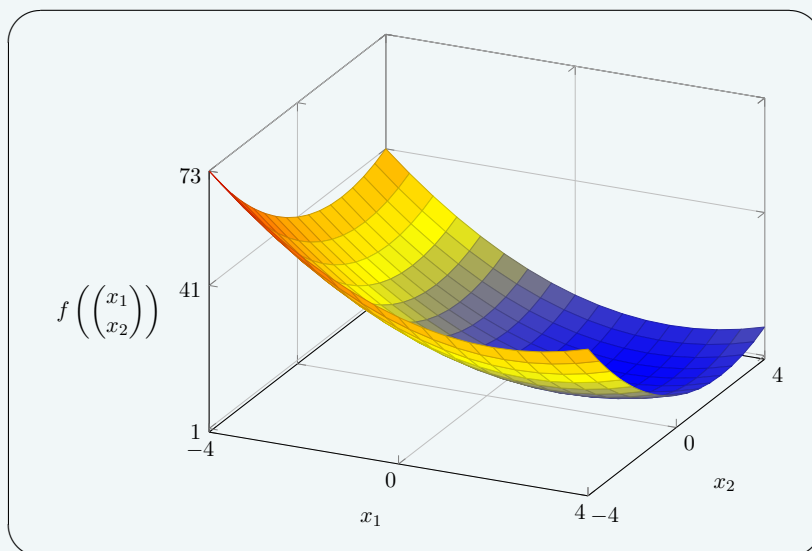
- Gli **insiemi di livello** della funzione sono:

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

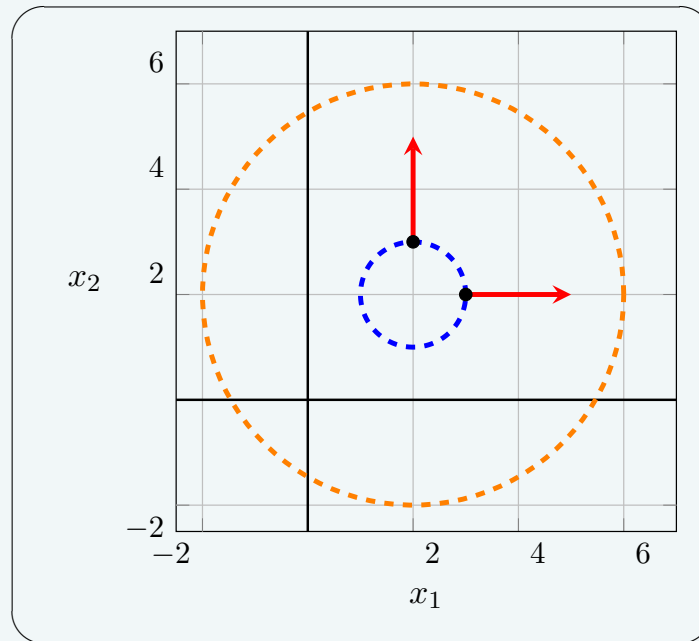
che sono circonferenze di centro $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ e raggio $\sqrt{k-1}$ per $k > 1$; l'insieme di livello $L(1)$ è il solo punto $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ e gli insiemi di livello con $k < 1$ sono vuoti.

Esempio 2.3: grafico, insiemi di livello e gradienti di una funzione quadratica

Per la funzione dell'esempio precedente, questo è il grafico nel dominio $[-4, 4] \times [-4, 4]$; il suo punto più basso è in $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$, dove $f = 1$, e il suo punto più alto è in $\begin{pmatrix} -4 \\ -4 \end{pmatrix}$, dove $f = 73$:



La figura seguente mostra gli insiemi di livello $L(2)$ (circonferenza di raggio 1, in blu) e $L(17)$ (circonferenza di raggio 4, in arancione), e i due gradienti dell'esempio precedente (in rosso), tracciati a partire dai punti $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ e $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ dell'insieme di livello $L(2)$:



I vettori gradiente sono perpendicolari agli insiemi di livello e puntano nella direzione di massima crescita della funzione (allontanandosi dal centro).

2.3 Esempi di funzioni non quadratiche

- Per una funzione quadratica la matrice Hessiana è costante (capitolo «Funzioni quadratiche»). In generale, la matrice Hessiana dipende dal punto, come nei due esempi seguenti.

Esempio 2.4: gradiente e matrice Hessiana di una funzione cubica

Consideriamo la funzione:

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

- Il **gradiente** e la **matrice Hessiana** della funzione sono:

$$\nabla f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 2x_1x_2 \\ x_1^2 \end{pmatrix}, \quad \nabla^2 f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \begin{pmatrix} 2x_2 & 2x_1 \\ 2x_1 & 0 \end{pmatrix}$$

- Nel punto $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ abbiamo:

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

- Nel punto $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$ abbiamo:

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

La matrice Hessiana cambia da punto a punto e, come previsto, è simmetrica in ogni punto.

Esempio 2.5: gradiente, matrice Hessiana e un insieme di livello di una funzione esponenziale

Consideriamo la funzione:

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

- Il **gradiente** e la **matrice Hessiana** della funzione sono:

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

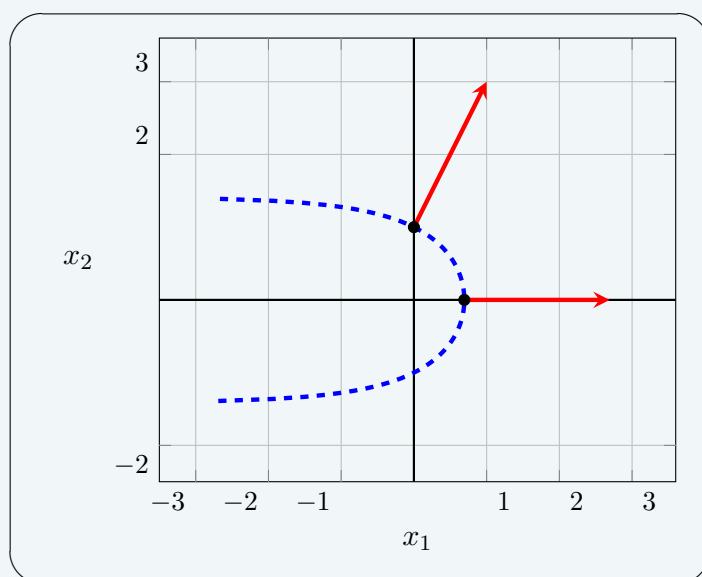
- L'insieme di livello $L(2)$ è la curva $e^{x_1} + x_2^2 = 2$, cioè $x_1 = \ln(2 - x_2^2)$ per $|x_2| < \sqrt{2}$.
 Contiene i punti $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} \ln 2 \\ 0 \end{pmatrix}$, poiché $e^0 + 1^2 = 2$ e $e^{\ln 2} + 0^2 = 2$. In questi punti:

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

e

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

La figura seguente mostra l'insieme di livello $L(2)$ (in blu) e i due vettori gradiente (in rosso):



La direzione tangente all'insieme di livello in $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ è $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ (la derivata di $x_1 = \ln(2 - x_2^2)$ rispetto a x_2 è $\frac{-2x_2}{2-x_2^2}$, uguale a -2 per $x_2 = 1$), e infatti:

$$(1 \quad 2) \begin{pmatrix} -2 \\ 1 \end{pmatrix} = -2 + 2 = 0$$

quindi il gradiente è perpendicolare all'insieme di livello anche quando l'insieme di livello non è una circonferenza.

Il gradiente $\nabla f(\mathbf{x})$ è il vettore colonna delle derivate parziali prime: punta nella direzione di massima crescita ed è perpendicolare all'insieme di livello passante per $\mathbf{x}$. La matrice Hessiana $\nabla^2 f(\mathbf{x})$ è la matrice (simmetrica) delle derivate parziali seconde e descrive la curvatura locale di f .

Funzioni convesse e concave

3.1 Funzioni convesse

Definizione 3.1: funzione convessa

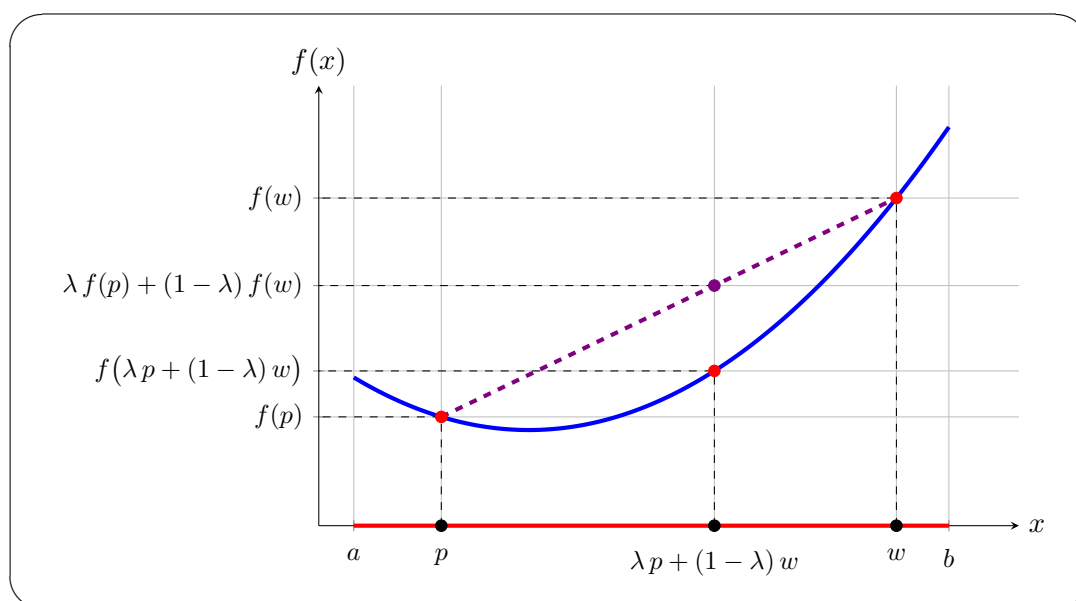
Una funzione $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è **convessa** nel suo dominio D se:

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

e il dominio D della funzione f è un insieme convesso.

Una funzione è convessa nel suo dominio se, per ogni coppia di punti del suo dominio convesso, il valore in una qualsiasi combinazione convessa dei due punti è minore o uguale alla combinazione convessa dei valori in quei due punti.

- Se $D = [a, b] \subseteq \mathbb{R}$ (un intervallo, che per natura è un insieme convesso), il grafico di una funzione convessa è dato dalla seguente figura:



3.2 Funzioni concave

Definizione 3.2: funzione concava

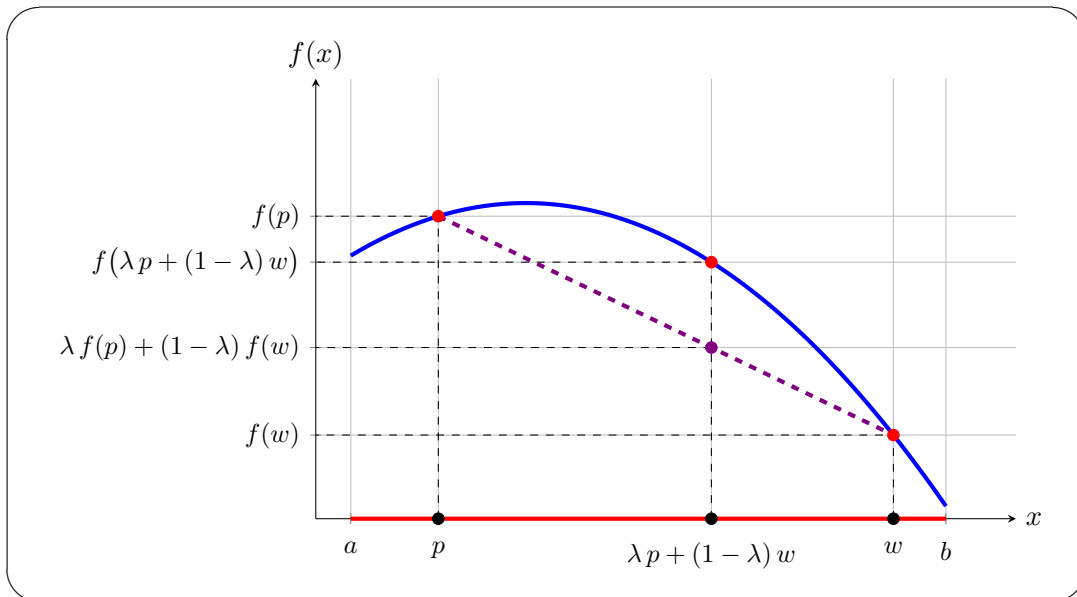
Una funzione $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è **concava** nel suo dominio D se:

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

e il dominio D della funzione f è un insieme convesso.

Una funzione è concava nel suo dominio se, per ogni coppia di punti del suo dominio convesso, il valore in una qualsiasi combinazione convessa dei due punti è maggiore o uguale alla combinazione convessa dei valori in quei due punti.

- Se $D = [a, b] \subseteq \mathbb{R}$ (un intervallo, che per natura è un insieme convesso), il grafico di una funzione concava è dato dalla seguente figura:



3.3 Esempi di funzioni convesse e concave

- In questo capitolo consideriamo anche funzioni non lineari. La convessità di una funzione si dimostra sempre verificando la definizione: per punti arbitrari $\mathbf{p}, \mathbf{w}$ del dominio e un arbitrario $\lambda \in [0, 1]$, verifichiamo la disuguaglianza della definizione. Per le funzioni quadratiche dei primi esempi, usiamo la seguente disuguaglianza.

Osservazione 3.3

Per ogni $a, b \in \mathbb{R}$ e ogni $\lambda \in [0, 1]$, abbiamo

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

Dimostrazione. Per ogni $a, b \in \mathbb{R}$ e ogni $\lambda \in [0, 1]$, abbiamo:

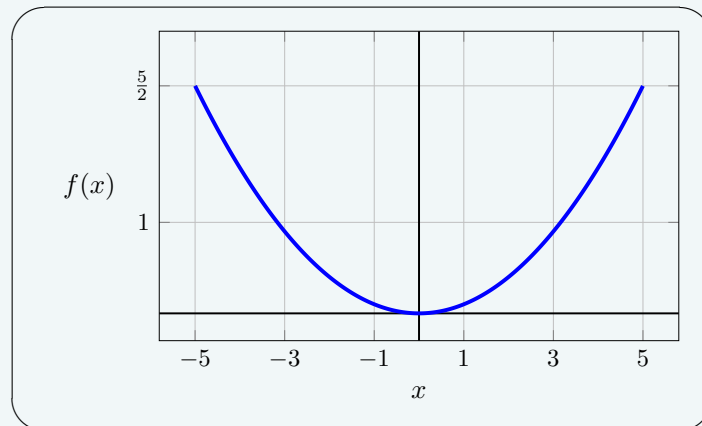
$$\begin{aligned} (\lambda a + (1 - \lambda)b)^2 &= \lambda^2 a^2 + 2\lambda(1 - \lambda)ab + (1 - \lambda)^2 b^2 \\ &= \lambda a^2 + (1 - \lambda)b^2 - \lambda(1 - \lambda)a^2 + 2\lambda(1 - \lambda)ab - \lambda(1 - \lambda)b^2 \\ &= \lambda a^2 + (1 - \lambda)b^2 - \lambda(1 - \lambda)(a^2 - 2ab +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

□

Esempio 3.1: una funzione quadratica convessa

Consideriamo la seguente funzione:

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



Per ogni coppia di punti $p, w \in \mathbb{R}$ e ogni $\lambda \in [0, 1]$, abbiamo:

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

Quindi,

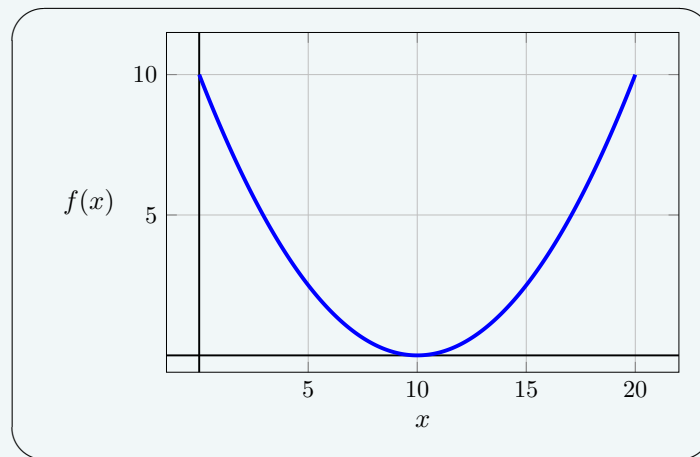
$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

Per natura, il dominio $\mathbb{R}$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

Esempio 3.2: una funzione quadratica convessa traslata

Consideriamo la seguente funzione:

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



Per ogni coppia di punti $p, w \in \mathbb{R}$ e ogni $\lambda \in [0, 1]$, abbiamo:

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

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

Quindi,

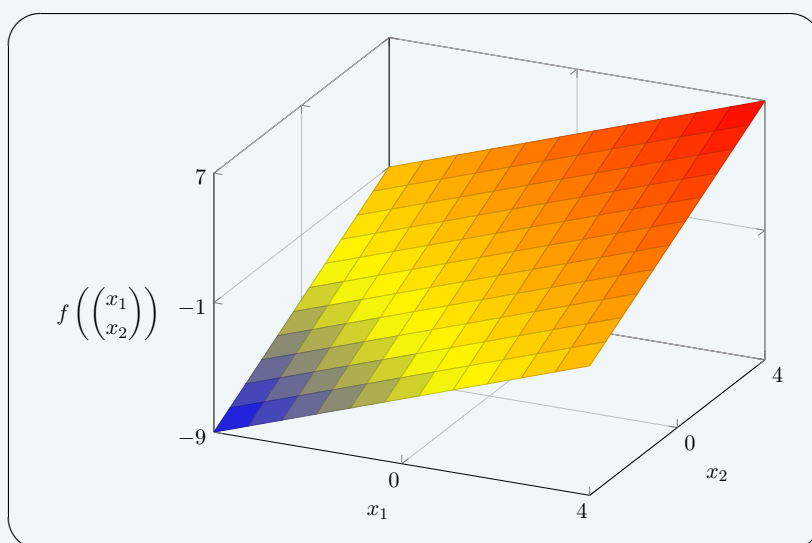
$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

Per natura, il dominio $\mathbb{R}$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

Esempio 3.3: una funzione lineare di due variabili

Consideriamo la seguente funzione:

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



Per ogni coppia di punti $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2$ e ogni $\lambda \in [0, 1]$, abbiamo:

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

Quindi,

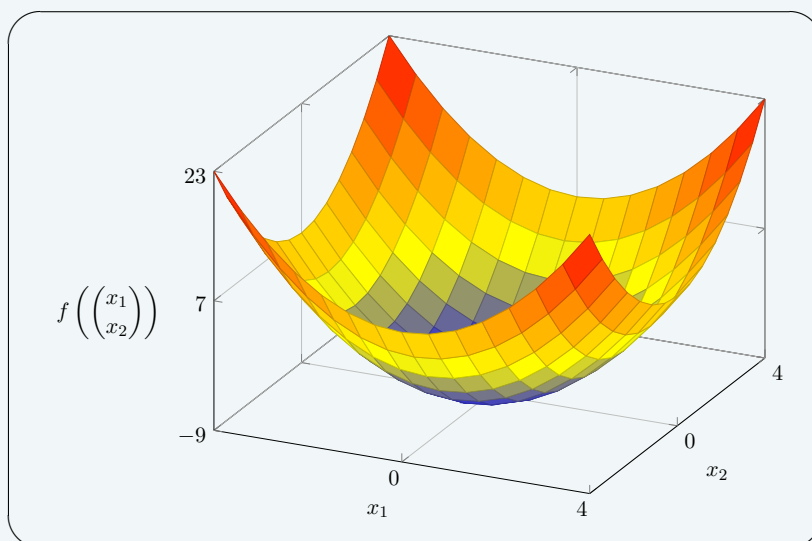
$$f(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \leq \lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w})$$

(con l'uguaglianza). Per natura, il dominio $\mathbb{R}^2$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio. Poiché vale l'uguaglianza, vale anche la disuguaglianza con $\geq$: la funzione è anche concava nel suo dominio.

Esempio 3.4: una funzione quadratica convessa di due variabili

Consideriamo la seguente funzione:

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



Per ogni coppia di punti $\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in \mathbb{R}^2$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= (\lambda p_1 + (1 - \lambda) w_1)^2 + (\lambda p_2 + (1 - \lambda) w_2)^2 - 9 \\ &\leq \lambda p_1^2 + (1 - \lambda) w_1^2 + \lambda p_2^2 + (1 - \lambda) w_2^2 - 9 \\ &\quad \text{(per la disuguaglianza (3.3), due volte)} \\ &= \lambda(p_1^2 + p_2^2 - 9) + (1 - \lambda)(w_1^2 + w_2^2 - 9) \\ &\quad \text{(poiché } \lambda + (1 - \lambda) = 1\text{)} \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^2$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

- I due esempi seguenti sono funzioni di una variabile il cui grafico è formato da due segmenti.

Esempio 3.5: il valore assoluto è una funzione convessa

Consideriamo la funzione $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$. Per ogni coppia di punti $p, w \in \mathbb{R}$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda p + (1 - \lambda)w) &= |\lambda p + (1 - \lambda)w| \\ &\leq |\lambda p| + |(1 - \lambda)w| \quad (\text{per la disuguaglianza triangolare}) \\ &= |\lambda| |p| + |1 - \lambda| |w| \quad (\text{per l'omogeneità assoluta}) \\ &= \lambda |p| + (1 - \lambda) |w| \quad (\text{poiché } \lambda \geq 0 \text{ e } 1 - \lambda \geq 0) \\ &= \lambda f(p) + (1 - \lambda)f(w) \end{aligned}$$

Quindi,

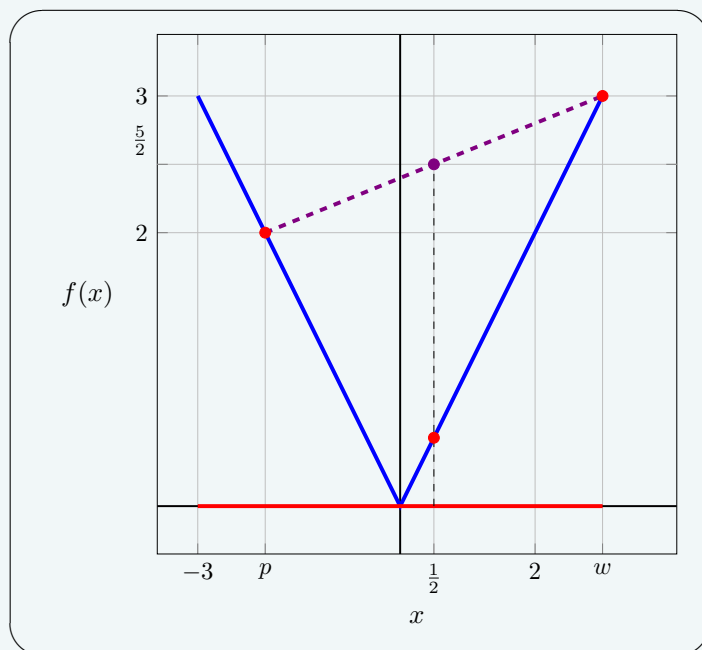
$$f(\lambda p + (1 - \lambda)w) \leq \lambda f(p) + (1 - \lambda)f(w)$$

Per natura, il dominio $\mathbb{R}$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

Ad esempio, con $p = -2$, $w = 3$ e $\lambda = \frac{1}{2}$ otteniamo $\lambda p + (1 - \lambda)w = \frac{1}{2}$ e

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

In figura, il segmento che congiunge $(p, f(p))$ e $(w, f(w))$ (in viola) si trova sopra il grafico di f (in blu).



Esempio 3.6: il minimo di due funzioni lineari è una funzione concava

Consideriamo la funzione $f : [-2, 4] \rightarrow \mathbb{R}$, $f(x) = \min\{x + 2, 4 - \frac{1}{2}x\}$. Per ogni coppia di punti $p, w \in [-2, 4]$ e ogni $\lambda \in [0, 1]$, abbiamo:

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

Quindi,

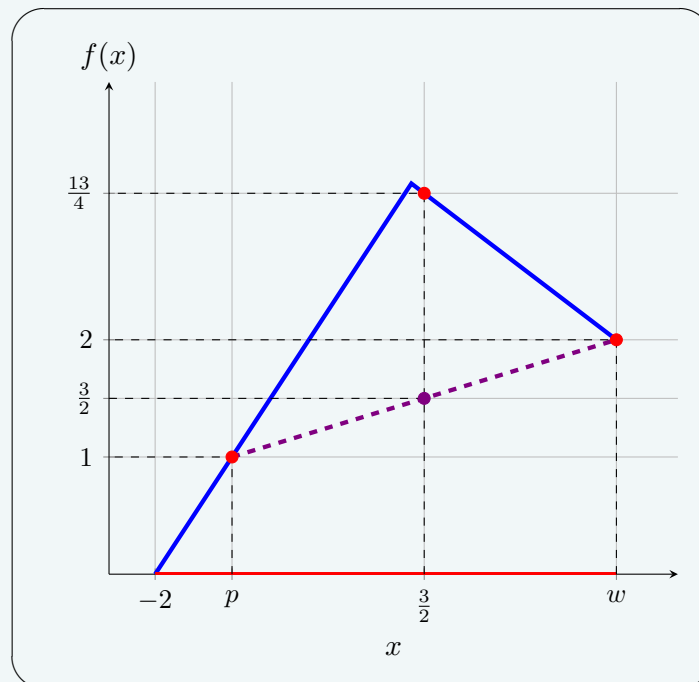
$$f(\lambda p + (1 - \lambda)w) \geq \lambda f(p) + (1 - \lambda)f(w)$$

Il dominio $[-2, 4]$ è un intervallo, quindi un insieme convesso. Di conseguenza, la funzione è concava nel suo dominio.

Le due rette si incontrano in $x = \frac{4}{3}$, dove $f(\frac{4}{3}) = \frac{10}{3}$. Ad esempio, con $p = -1$, $w = 4$ e $\lambda = \frac{1}{2}$ otteniamo $\lambda p + (1 - \lambda)w = \frac{3}{2}$ e

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

In figura, il segmento che congiunge $(p, f(p))$ e $(w, f(w))$ (in viola) si trova sotto il grafico di f (in blu).



3.4 Funzioni convesse definite da norme

Definizione 3.4: norma

Una **norma** su $\mathbb{R}^n$ è una funzione $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ tale che, per ogni $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\alpha \in \mathbb{R}$:

- $\|\mathbf{p}\| \geq 0$, e $\|\mathbf{p}\| = 0$ se e solo se $\mathbf{p} = \mathbf{0}$ (positività);
- $\|\alpha \mathbf{p}\| = |\alpha| \|\mathbf{p}\|$ (omogeneità assoluta);
- $\|\mathbf{p} + \mathbf{w}\| \leq \|\mathbf{p}\| + \|\mathbf{w}\|$ (disuguaglianza triangolare).

Definizione 3.5: le norme 1, 2, q e ∞

Per ogni $\mathbf{x} \in \mathbb{R}^n$:

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

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{j=1}^n x_j^2} \quad (\text{norma 2, o norma euclidea})$$

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

$$\|\mathbf{x}\|_\infty = \max_{j=1, \dots, n} |x_j| \quad (\text{norma } \infty, \text{ o norma del massimo})$$

Le norme 1 e 2 sono la norma q con $q = 1$ e $q = 2$, e $\|\mathbf{x}\|_\infty = \lim_{q \rightarrow \infty} \|\mathbf{x}\|_q$.

Esempio 3.7: le quattro norme di un vettore

Per $\mathbf{x} = \begin{pmatrix} 3 \\ -4 \end{pmatrix}$ abbiamo:

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

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

Osservazione 3.6: convessità della norma euclidea

La funzione

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

è convessa nel suo dominio.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}\|_2 \\ &\leq \|\lambda \mathbf{p}\|_2 + \|(1 - \lambda) \mathbf{w}\|_2 \quad (\text{per la disuguaglianza triangolare}) \\ &= |\lambda| \|\mathbf{p}\|_2 + |1 - \lambda| \|\mathbf{w}\|_2 \quad (\text{per l'omogeneità assoluta}) \\ &= \lambda \|\mathbf{p}\|_2 + (1 - \lambda) \|\mathbf{w}\|_2 \quad (\text{poiché } \lambda \geq 0 \text{ e } 1 - \lambda \geq 0) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^n$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio. $\square$

Osservazione 3.7: convessità di ogni norma

Per ogni norma $\|\cdot\|$ su $\mathbb{R}^n$ (in particolare $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_q$ con $q \geq 1$, e $\|\cdot\|_\infty$), la funzione

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

è convessa nel suo dominio.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}\| \\ &\leq \|\lambda \mathbf{p}\| + \|(1 - \lambda) \mathbf{w}\| \quad (\text{per la disuguaglianza triangolare}) \\ &= |\lambda| \|\mathbf{p}\| + |1 - \lambda| \|\mathbf{w}\| \quad (\text{per l'omogeneità assoluta}) \\ &= \lambda \|\mathbf{p}\| + (1 - \lambda) \|\mathbf{w}\| \quad (\text{poiché } \lambda \geq 0 \text{ e } 1 - \lambda \geq 0) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^n$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio. $\square$

Corollario 3.8

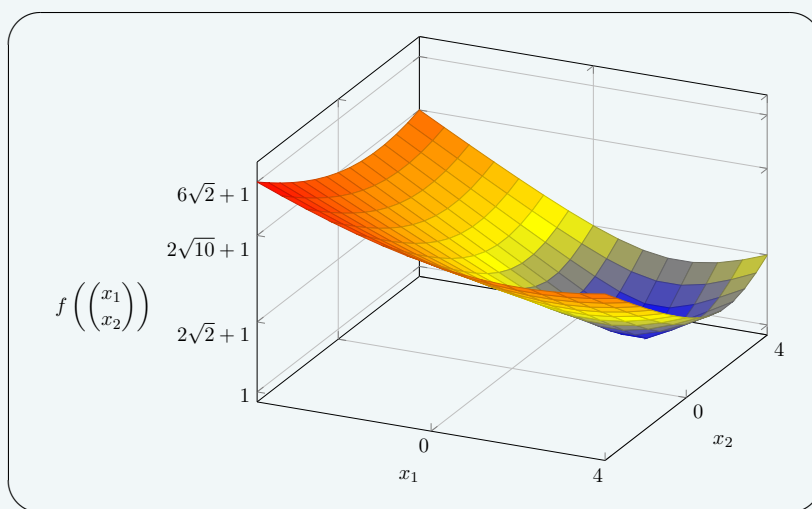
Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo

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

Esempio 3.8: la distanza da un punto

Consideriamo il vettore $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ e la funzione:

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



Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w} - \mathbf{c}\|_2 + 1 \\ &= \|\lambda(\mathbf{p} - \mathbf{c}) + (1 - \lambda)(\mathbf{w} - \mathbf{c})\|_2 + 1 \quad (\text{poiché } \lambda + (1 - \lambda) = 1) \\ &\leq \lambda \|\mathbf{p} - \mathbf{c}\|_2 + (1 - \lambda) \|\mathbf{w} - \mathbf{c}\|_2 + 1 \quad (\text{per la disuguaglianza (3.4)}) \\ &= \lambda(\|\mathbf{p} - \mathbf{c}\|_2 + 1) + (1 - \lambda)(\|\mathbf{w} - \mathbf{c}\|_2 + 1) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^2$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

Osservazione 3.9: convessità del quadrato della norma euclidea

La funzione

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

è convessa nel suo dominio.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) &= \|\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}\|_2^2 \\ &= (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})' (\lambda \mathbf{p} + (1 - \lambda) \mathbf{w}) \\ &= \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + 2\lambda(1 - \lambda) (\mathbf{p}' \mathbf{w}) \\ &\leq \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + 2\lambda(1 - \lambda) \|\mathbf{p}\|_2 \|\mathbf{w}\|_2 \\ &\quad \text{(per la disuguaglianza di Cauchy-Schwarz)} \\ &\leq \lambda^2 \|\mathbf{p}\|_2^2 + (1 - \lambda)^2 \|\mathbf{w}\|_2^2 + \lambda(1 - \lambda) \|\mathbf{p}\|_2^2 + \lambda(1 - \lambda) \|\mathbf{w}\|_2^2 \\ &\quad \text{(poiché } 2ab \leq a^2 + b^2 \text{ per ogni } a, b \in \mathbb{R}, \\ &\quad \text{con } a = \|\mathbf{p}\|_2 \text{ e } b = \|\mathbf{w}\|_2) \\ &= (\lambda^2 + \lambda(1 - \lambda)) \|\mathbf{p}\|_2^2 + ((1 - \lambda)^2 + \lambda(1 - \lambda)) \|\mathbf{w}\|_2^2 \\ &= \lambda \|\mathbf{p}\|_2^2 + (1 - \lambda) \|\mathbf{w}\|_2^2 \\ &= \lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w}) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^n$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio. $\square$

- La dimostrazione seguente è più diretta e non richiede l'uso della disuguaglianza di Cauchy-Schwarz né della disuguaglianza $2ab \leq a^2 + b^2$ per ogni $a, b \in \mathbb{R}$.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo:

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

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

Quindi,

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

Per natura, il dominio $\mathbb{R}^n$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio. $\square$

Corollario 3.10

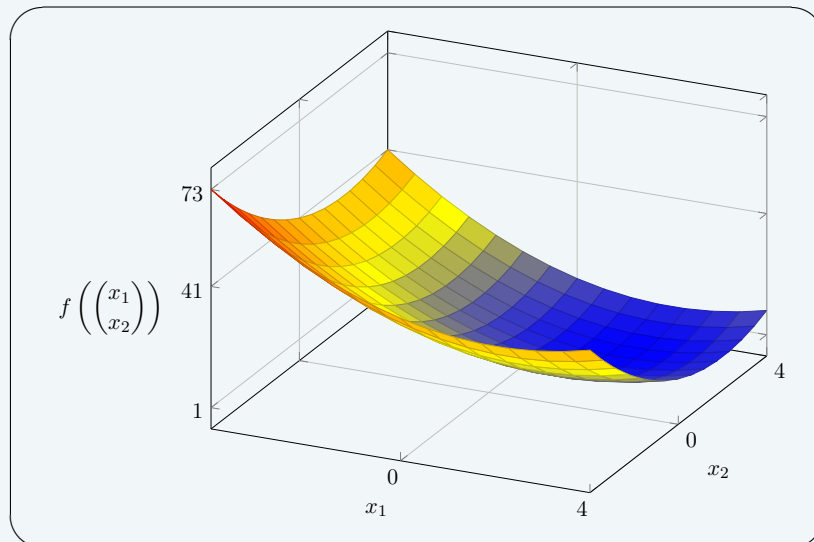
Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo

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

Esempio 3.9: il quadrato della distanza da un punto

Consideriamo il vettore $\mathbf{c} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ e la funzione:

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



Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$f(\lambda\mathbf{p} + (1-\lambda)\mathbf{w}) = \|\lambda\mathbf{p} + (1-\lambda)\mathbf{w} - \mathbf{c}\|_2^2 + 1$$

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

Quindi,

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

Per natura, il dominio $\mathbb{R}^2$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

3.5 Proprietà delle funzioni convesse e concave

Osservazione 3.11

Una funzione $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è convessa in D se e solo se la funzione $-f$ è concava in D .

Dimostrazione. Sia f convessa in D . Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in D$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned}
 (-f)(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) &= -f(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \\
 &\geq -(\lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w})) \\
 &\quad (\text{poiché } f \text{ è convessa, e } -1 \text{ inverte la disuguaglianza}) \\
 &= \lambda (-f)(\mathbf{p}) + (1 - \lambda)(-f)(\mathbf{w})
 \end{aligned}$$

Quindi,

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

Il dominio D è un insieme convesso, poiché f è convessa in D . Di conseguenza, la funzione $-f$ è concava nel suo dominio. Il viceversa si dimostra allo stesso modo, partendo dalla disuguaglianza della funzione concava $-f$ e moltiplicandola per -1 . $\square$

Osservazione 3.12

Siano $f, g : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convesse in D e siano $\alpha, \beta \geq 0$. Allora la funzione $\alpha f + \beta g$ è convessa in D . Lo stesso vale per le funzioni concave.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in D$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} (\alpha f + \beta g)(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) &= \alpha f(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) + \beta g(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \\ &\leq \alpha(\lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w})) + \beta(\lambda g(\mathbf{p}) + (1 - \lambda)g(\mathbf{w})) \\ &\quad (\text{poiché } f \text{ e } g \text{ sono convesse, e } \alpha \geq 0, \beta \geq 0) \\ &= \lambda(\alpha f(\mathbf{p}) + \beta g(\mathbf{p})) + (1 - \lambda)(\alpha f(\mathbf{w}) + \beta g(\mathbf{w})) \\ &= \lambda(\alpha f + \beta g)(\mathbf{p}) + (1 - \lambda)(\alpha f + \beta g)(\mathbf{w}) \end{aligned}$$

Quindi,

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

Il dominio D è un insieme convesso, poiché f è convessa in D . Di conseguenza, la funzione $\alpha f + \beta g$ è convessa nel suo dominio. Per le funzioni concave vale lo stesso ragionamento con $\geq$. $\square$

- Il massimo puntuale di funzioni convesse è a sua volta una funzione convessa: questo ci permette di combinare più funzioni convesse in una sola preservando la convessità.

Osservazione 3.13

Date m funzioni $f_i : D_i \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convesse nei loro domini, con $i \in \{1, 2, \dots, m\}$, la funzione

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

è convessa nel suo dominio. Analogamente, il minimo di m funzioni concave è concavo nel suo dominio.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \bigcap_{i=1}^m D_i$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} f(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) &= \max \{ f_i(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) : i \in \{1, 2, \dots, m\} \} \\ &= f_k(\lambda \mathbf{p} + (1 - \lambda)\mathbf{w}) \\ &\quad (\text{per un indice } k \in \{1, 2, \dots, m\} \text{ che realizza il massimo}) \\ &\leq \lambda f_k(\mathbf{p}) + (1 - \lambda)f_k(\mathbf{w}) \quad (\text{poiché } f_k \text{ è convessa}) \\ &\leq \lambda \max \{ f_i(\mathbf{p}) : i \in \{1, 2, \dots, m\} \} \\ &\quad + (1 - \lambda) \max \{ f_i(\mathbf{w}) : i \in \{1, 2, \dots, m\} \} \\ &\quad (\text{poiché } \lambda \geq 0 \text{ e } 1 - \lambda \geq 0) \end{aligned}$$

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

Quindi,

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

Il dominio $\bigcap_{i=1}^m D_i$ è un insieme convesso, poiché è l'intersezione di insiemi convessi. Di conseguenza, la funzione f è convessa nel suo dominio. L'enunciato per il minimo di funzioni concave segue applicando questo a $-f_1, \dots, -f_m$. $\square$

- Il valore assoluto $|x| = \max\{x, -x\}$ dell'Esempio 3.5 è il massimo delle due funzioni x e $-x$, che sono convesse (e concave), poiché la definizione vale con l'uguaglianza: $-(\lambda p + (1 - \lambda)w) = \lambda(-p) + (1 - \lambda)(-w)$. La funzione dell'Esempio 3.6 è il minimo delle due funzioni $x + 2$ e $4 - \frac{1}{2}x$, che sono concave per la stessa ragione: $(\lambda p + (1 - \lambda)w) + 2 = \lambda(p + 2) + (1 - \lambda)(w + 2)$ e $4 - \frac{1}{2}(\lambda p + (1 - \lambda)w) = \lambda(4 - \frac{1}{2}p) + (1 - \lambda)(4 - \frac{1}{2}w)$. Per l'osservazione precedente, ritroviamo che le due funzioni sono, rispettivamente, convessa e concava.

Esempio 3.10: massimo di funzioni convesse

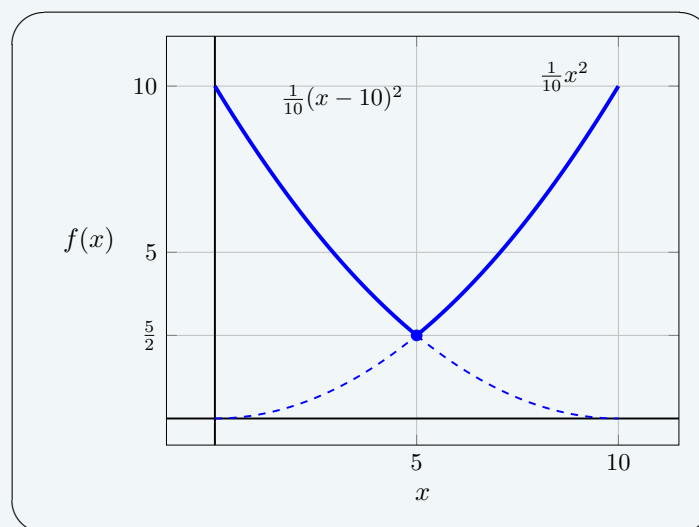
Consideriamo le seguenti due funzioni convesse (Esempi 3.2 e 3.1):

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

Per l'osservazione precedente, la seguente funzione è convessa nel suo dominio:

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

In figura, il grafico di f è la linea blu continua; le parti tratteggiate delle due parabole si trovano sotto di essa.



Osservazione 3.14

Date m funzioni $f_i : D_i \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ convesse nei loro domini, con $i \in \{1, 2, \dots, m\}$, l'insieme

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

è convesso.

Dimostrazione. Per ogni coppia di punti $\mathbf{p}$ e $\mathbf{w}$ dell'insieme e ogni $\lambda \in [0, 1]$, il punto $\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}$ appartiene a $\bigcap_{i=1}^m D_i$, poiché $\mathbf{p}, \mathbf{w} \in D_i$ e ogni dominio D_i è convesso, e per ogni $i \in \{1, 2, \dots, m\}$ abbiamo:

$$\begin{aligned} f_i(\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}) &\leq \lambda f_i(\mathbf{p}) + (1 - \lambda)f_i(\mathbf{w}) \quad (\text{poiché } f_i \text{ è convessa}) \\ &\leq \lambda 0 + (1 - \lambda) 0 \\ &\quad (\text{poiché } f_i(\mathbf{p}) \leq 0, f_i(\mathbf{w}) \leq 0, \lambda \geq 0 \text{ e } 1 - \lambda \geq 0) \\ &= 0 \end{aligned}$$

Quindi,

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

Di conseguenza, $\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}$ appartiene all'insieme e l'insieme è convesso. $\square$

Esempio 3.11: insieme convesso da funzioni convesse

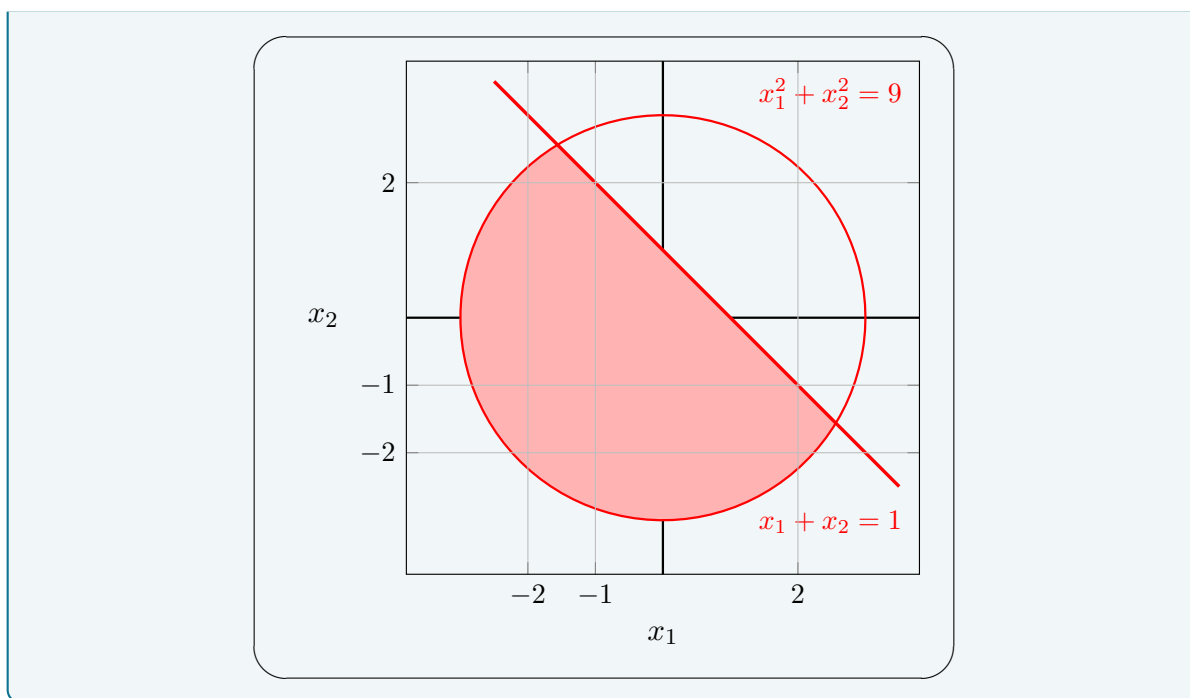
Consideriamo le seguenti due funzioni convesse (Esempi 3.4 e 3.3):

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

Per l'osservazione precedente, il seguente è un insieme convesso:

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

È il disco di raggio 3 centrato nell'origine (circonferenza rossa) tagliato dal semipiano $x_1 + x_2 \leq 1$ (retta rossa), ombreggiato in rosso.



Esempio 3.12: insieme convesso da funzioni convesse

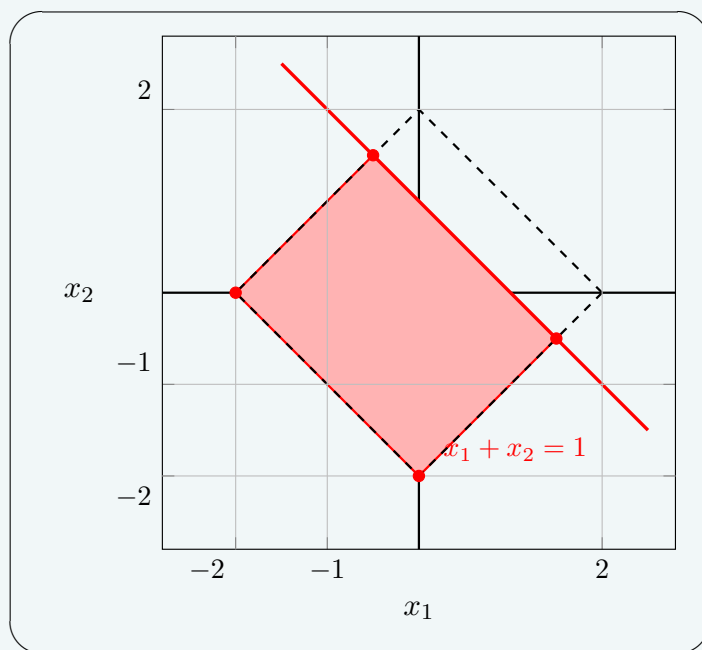
Consideriamo le seguenti due funzioni convesse:

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

La funzione f_1 è convessa perché $|x_1| = \max\{x_1, -x_1\}$ e $|x_2| = \max\{x_2, -x_2\}$ sono massimi di funzioni convesse (come $|x| = \max\{x, -x\}$ sopra), e la somma di funzioni convesse è convessa; f_2 è convessa per l'Esempio 3.3. Anche in questo caso, per la stessa osservazione, il seguente è un insieme convesso:

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

È il quadrato di vertici $(\pm 2, 0)$ e $(0, \pm 2)$ (tratteggiato) tagliato dal semipiano $x_1 + x_2 \leq 1$ (retta rossa): il quadrilatero di vertici $(-2, 0)$, $(0, -2)$, $(\frac{3}{2}, -\frac{1}{2})$ e $(-\frac{1}{2}, \frac{3}{2})$ ombreggiato in rosso.



I punti che soddisfano un numero finito di vincoli $f_i(\mathbf{x}) \leq 0$ con funzioni convesse f_i formano un insieme convesso. Ad esempio, ogni semispazio $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}'\mathbf{x} \leq b\}$, con $\mathbf{a} \in \mathbb{R}^n$ e $b \in \mathbb{R}$, è un insieme convesso di questo tipo, poiché la funzione lineare $\mathbf{a}'\mathbf{x} - b$ è convessa; lo è anche ogni intersezione di un numero finito di semispazi, come i poliedri della programmazione lineare.

Funzioni quadratiche

4.1 Funzioni quadratiche

Definizione 4.1: funzione quadratica

Data una matrice quadrata simmetrica

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

la funzione (scalare) associata

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

si chiama **funzione quadratica**.

- Se la matrice quadrata di una funzione quadratica non è simmetrica, la si può rendere simmetrica sostituendo ogni coppia di elementi fuori diagonale con la loro media:

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

Ciò è lecito perché per ogni $i, j \in \{1, 2, \dots, n\}$ con $i < j$, abbiamo:

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

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

Quindi, le funzioni quadratiche definite da $\mathbf{Q}$ e da $\frac{1}{2}(\mathbf{Q} + \mathbf{Q}')$ sono equivalenti, poiché assumono lo stesso valore per ogni $\mathbf{x} \in \mathbb{R}^n$.

Senza perdita di generalità, possiamo supporre che la matrice $\mathbf{Q}$ di una funzione quadratica sia simmetrica.

Esempio 4.1: matrici equivalenti per le funzioni quadratiche

Consideriamo la seguente matrice 3×3 non simmetrica:

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

Si ottiene una funzione quadratica equivalente usando la seguente matrice simmetrica:

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

Infatti, le due matrici definiscono la stessa funzione:

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

4.2 Gradiente, matrice Hessiana e insiemi di livello

Osservazione 4.2: gradiente di una funzione quadratica

Il **gradiente** (vettore delle derivate parziali) di una funzione quadratica è:

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

Dimostrazione. Per ogni $k \in \{1, 2, \dots, n\}$, solo i termini della doppia sommatoria con $i = k$ o $j = k$ contengono la variabile x_k : i termini $q_{kj} x_k x_j$ contribuiscono alla derivata con $q_{kj} x_j$ e i termini $q_{ik} x_i x_k$ con $q_{ik} x_i$ (il termine $q_{kk} x_k^2$, con $i = j = k$, contribuisce con $2 q_{kk} x_k$, una volta in ciascuna sommatoria). Quindi:

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

dove l'ultima uguaglianza vale perché la matrice $\mathbf{Q}$ è simmetrica ($q_{jk} = q_{kj}$). $\square$

Osservazione 4.3: matrice Hessiana di una funzione quadratica

La **matrice Hessiana** (matrice delle derivate parziali del secondo ordine) di una funzione quadratica è:

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

Dimostrazione. Per l'osservazione precedente, per ogni $k \in \{1, 2, \dots, n\}$ abbiamo $\frac{\partial f(\mathbf{x})}{\partial x_k} = 2 \sum_{j=1}^n q_{kj} x_j$, la cui derivata rispetto a x_h è $2 q_{kh}$, per ogni $h \in \{1, 2, \dots, n\}$. $\square$

Definizione 4.4: insiemi di livello di una funzione quadratica

Gli **insiemi di livello** di una funzione quadratica (gli insiemi dei punti in cui la funzione è **costante**) sono:

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

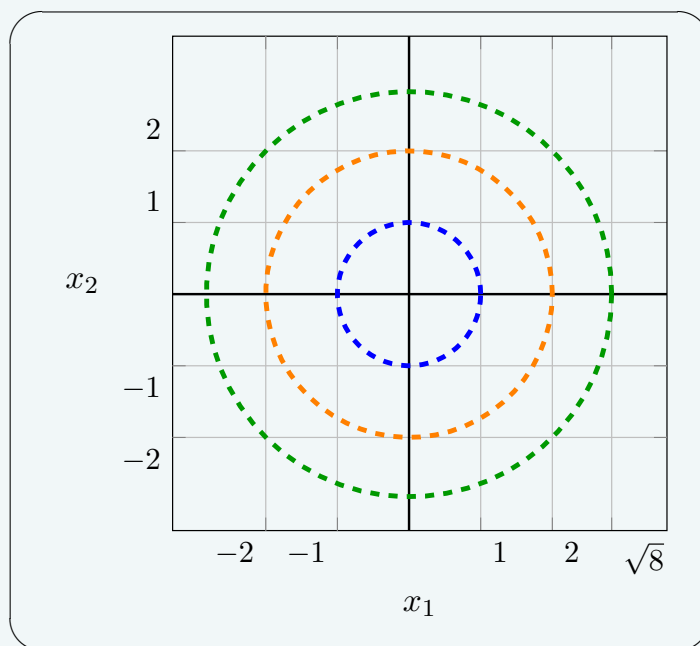
- A seconda della matrice $\mathbf{Q}$, gli insiemi di livello non vuoti possono essere **ellissoidi** (ellissi in $\mathbb{R}^2$), **iperboloidi** (iperboli in $\mathbb{R}^2$) oppure, quando alcune direzioni non cambiano il valore della funzione, **cilindri** (coppie di rette parallele in $\mathbb{R}^2$). Per $k = 0$ ci sono anche casi degeneri: il solo punto $\mathbf{0}$ al posto di un ellissoide, due rette incidenti nell'origine in $\mathbb{R}^2$ (un cono in $\mathbb{R}^n$) al posto di un iperboloide, e una sola retta passante per l'origine in $\mathbb{R}^2$ (un sottospazio lineare in $\mathbb{R}^n$) al posto di un cilindro. Il grafico delle funzioni quadratiche degli esempi seguenti è un **paraboloide**.

Per una funzione quadratica con matrice simmetrica $\mathbf{Q}$, il gradiente è $\nabla f(\mathbf{x}) = 2 \mathbf{Q} \mathbf{x}$ e la matrice Hessiana è la matrice costante $\nabla^2 f(\mathbf{x}) = 2 \mathbf{Q}$. In ogni punto in cui il gradiente non è nullo, il gradiente è perpendicolare all'insieme di livello passante per quel punto.

Esempio 4.2: insiemi di livello di funzioni quadratiche

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, la funzione quadratica associata è:

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



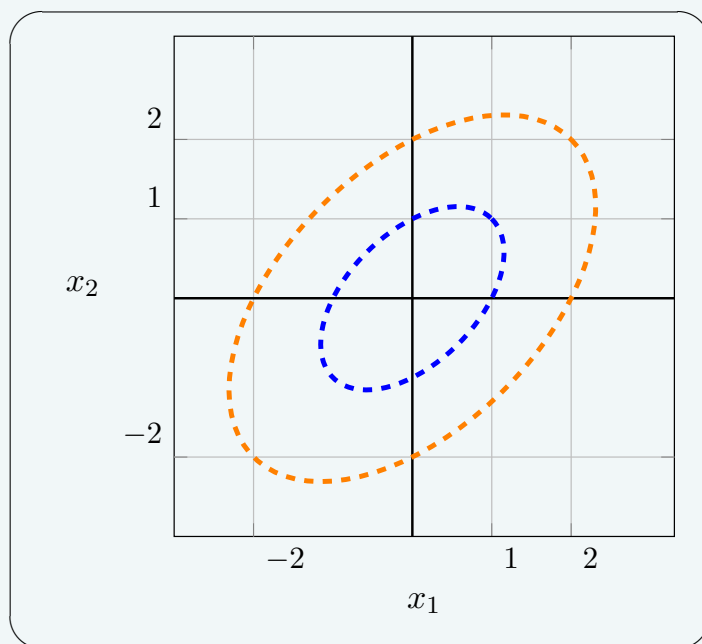
La figura mostra:

- Circonferenza blu: insieme di livello $x_1^2 + x_2^2 = 1$ (raggio 1)
- Circonferenza arancione: insieme di livello $x_1^2 + x_2^2 = 4$ (raggio 2)
- Circonferenza verde: insieme di livello $x_1^2 + x_2^2 = 8$ (raggio $\sqrt{8}$)

Esempio 4.3: insiemi di livello di funzioni quadratiche

With $\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, la funzione quadratica associata è:

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



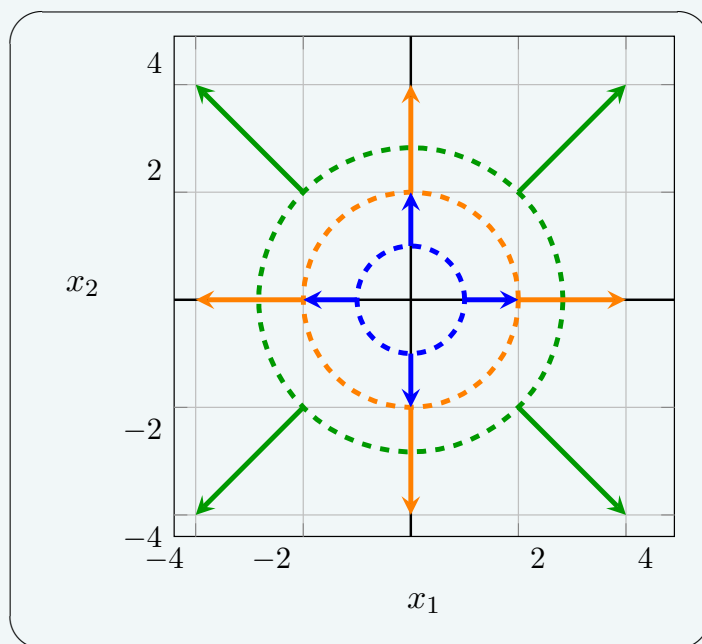
La figura mostra:

- Ellisse blu: insieme di livello $2x_1^2 - 2x_1x_2 + 2x_2^2 = 2$
- Ellisse arancione: insieme di livello $2x_1^2 - 2x_1x_2 + 2x_2^2 = 8$

Esempio 4.4: gradienti di funzioni quadratiche

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, il gradiente della funzione quadratica è:

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



La figura mostra:

- Vettori blu (gradienti di norma 2 sull'insieme di livello $x_1^2 + x_2^2 = 1$):

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

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

- Vettori arancioni (gradienti di norma 4 sull'insieme di livello $x_1^2 + x_2^2 = 4$):

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

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

- Vettori verdi (gradienti di norma $4\sqrt{2}$ sull'insieme di livello $x_1^2 + x_2^2 = 8$):

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

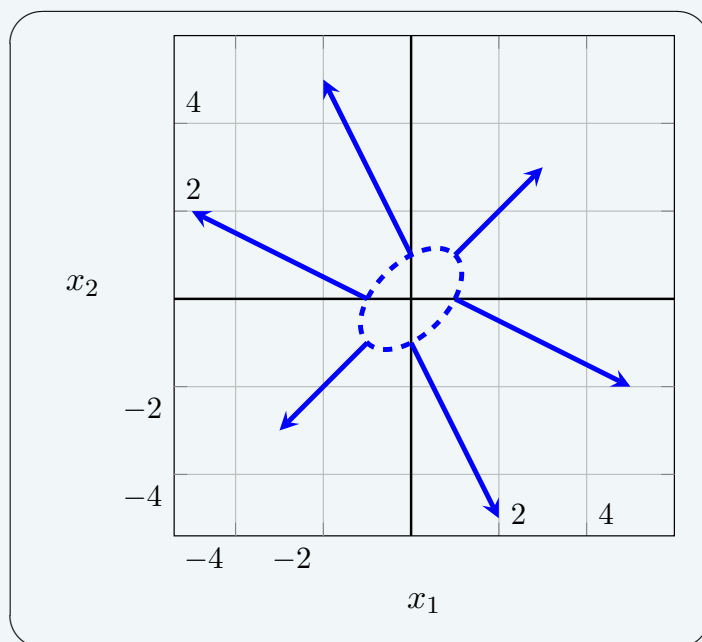
$$\nabla f \left(\begin{pmatrix} -2 \\ 2 \end{pmatrix} \right) = \begin{pmatrix} -4 \\ 4 \end{pmatrix}, \quad \nabla f \left(\begin{pmatrix} 2 \\ -2 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ -4 \end{pmatrix}$$

Ogni gradiente è perpendicolare all'insieme di livello passante per il suo punto.

Esempio 4.5: gradienti di funzioni quadratiche

With $Q = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, il gradiente della funzione quadratica è:

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



La figura mostra:

- Ellisse blu: insieme di livello $2x_1^2 - 2x_1x_2 + 2x_2^2 = 2$
- Vettori blu (gradienti sull'insieme di livello, di norma $2\sqrt{2}$ nei punti sulla diagonale e di norma $2\sqrt{5}$ nei punti sugli assi):

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

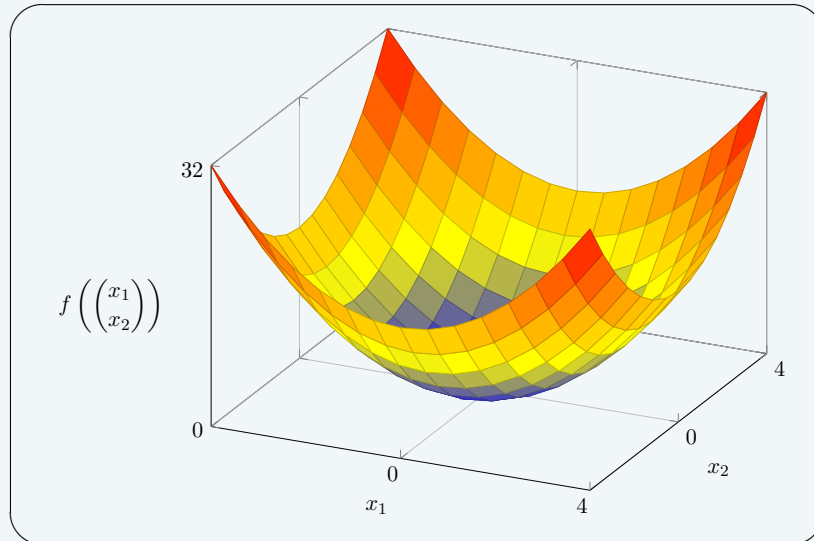
$$\nabla f \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \quad \nabla f \left(\begin{pmatrix} -1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} -4 \\ 2 \end{pmatrix},$$

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

Anche in questo caso, ogni gradiente è perpendicolare all'insieme di livello passante per il suo punto.

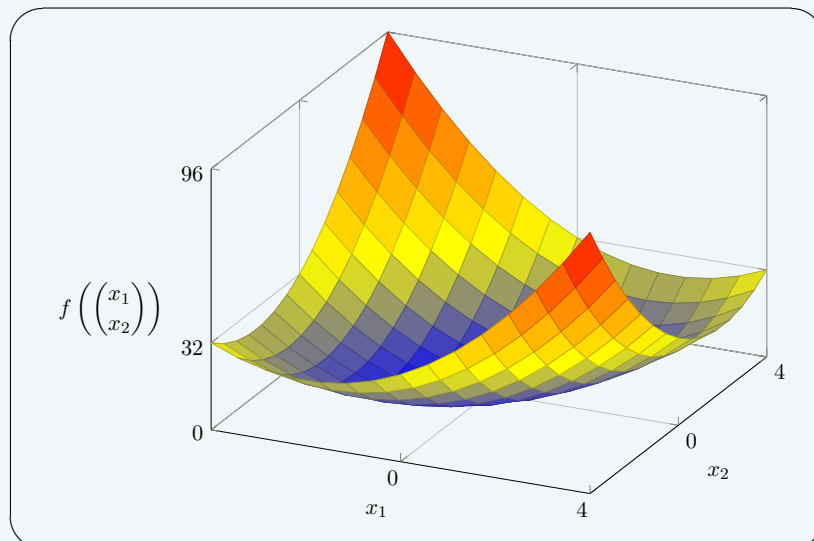
Esempio 4.6: grafici di funzioni quadratiche

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, abbiamo $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + x_2^2$. Il suo grafico è:



Esempio 4.7: grafici di funzioni quadratiche

With $\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, abbiamo $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = 2x_1^2 - 2x_1x_2 + 2x_2^2$. Il suo grafico è:



4.3 Convessità delle funzioni quadratiche

Definizione 4.5: matrice semidefinita positiva

Una matrice quadrata simmetrica $Q \in \mathbb{R}^{n \times n}$ è **semidefinita positiva** se:

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

Proposizione 4.6

Per ogni matrice quadrata Q , se Q è semidefinita positiva allora la funzione quadratica associata è convessa nel suo dominio.

Dimostrazione. Se la matrice Q è semidefinita positiva, per ogni coppia di punti $p, w \in \mathbb{R}^n$ e per ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} & f(\lambda p + (1 - \lambda) w) \\ &= (\lambda p + (1 - \lambda) w)' Q (\lambda p + (1 - \lambda) w) \\ &= \lambda^2 p'Qp + (1 - \lambda)^2 w'Qw \\ &\quad + \lambda (1 - \lambda) p'Qw + \lambda (1 - \lambda) w'Qp \\ &= \lambda p'Qp + (1 - \lambda) w'Qw \\ &\quad - \lambda (1 - \lambda) (p'Qp + w'Qw - p'Qw - w'Qp) \\ &\quad (\text{poiché } \lambda^2 = \lambda - \lambda(1 - \lambda) \text{ e } (1 - \lambda)^2 = (1 - \lambda) - \lambda(1 - \lambda)) \\ &= \lambda p'Qp + (1 - \lambda) w'Qw - \lambda (1 - \lambda) (p - w)'Q(p - w) \\ &\leq \lambda p'Qp + (1 - \lambda) w'Qw \\ &\quad (\text{poiché } Q \text{ è semidefinita positiva e } \lambda(1 - \lambda) \geq 0) \\ &= \lambda f(p) + (1 - \lambda) f(w) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^n$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio. $\square$

- Abbiamo dimostrato che:

$$Q \text{ semidefinita positiva} \implies \text{funzione quadratica convessa}$$

quindi “ Q semidefinita positiva” è una condizione sufficiente per “funzione quadratica convessa” e “funzione quadratica convessa” è una condizione necessaria per “ Q semidefinita positiva”.

- L’implicazione vale anche nella direzione opposta, come afferma la seguente proposizione.

Proposizione 4.7

Per ogni funzione quadratica, se la funzione è convessa nel suo dominio allora la matrice quadrata associata Q è semidefinita positiva.

Dimostrazione. Se la funzione quadratica è convessa, per ogni coppia di vettori colonna $p, w \in \mathbb{R}^n$ e ogni scalare $\lambda \in (0, 1)$, abbiamo:

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

dove il membro sinistro è:

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

Riordinando e semplificando, ciò implica che:

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

che, poiché $\lambda (1 - \lambda) > 0$, implica:

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

Ogni vettore $x \in \mathbb{R}^n$ si può scrivere come $x = p - w$ (per esempio, con $p = x$ e $w = 0$). Quindi, $x' Q x \geq 0$ per ogni $x \in \mathbb{R}^n$, e la matrice quadrata Q è semidefinita positiva. $\square$

- Abbiamo dimostrato che:

$$\text{funzione quadratica convessa} \implies Q \text{ semidefinita positiva}$$

quindi “funzione quadratica convessa” è una condizione sufficiente per “ Q semidefinita positiva” e “ Q semidefinita positiva” è una condizione necessaria per “funzione quadratica convessa”.

- Abbiamo quindi, infine:

$$Q \text{ semidefinita positiva} \iff \text{funzione quadratica convessa}$$

quindi “ Q semidefinita positiva” è una condizione necessaria e sufficiente per “funzione quadratica convessa” e viceversa.

Corollario 4.8

Per ogni matrice quadrata $\mathbf{Q}$, la funzione quadratica associata è convessa nel suo dominio se e solo se $\mathbf{Q}$ è semidefinita positiva.

- Con $\mathbf{Q}$ uguale alla matrice identità, la funzione quadratica è $f(\mathbf{x}) = \mathbf{x}'\mathbf{x} = \|\mathbf{x}\|_2^2$, il quadrato della norma euclidea, la cui convessità è dimostrata nel capitolo «Funzioni convesse e concave» direttamente dalla definizione. La matrice identità è semidefinita positiva, poiché $\mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 + \dots + x_n^2 \geq 0$, in accordo con il corollario.
- Quando $\mathbf{x}'\mathbf{Q}\mathbf{x}$ si può scrivere come somma di quadrati di funzioni lineari, come nell'esempio seguente, la matrice $\mathbf{Q}$ è semidefinita positiva, e la convessità si può dimostrare anche direttamente dalla definizione, con la disuguaglianza $(\lambda a + (1 - \lambda)b)^2 \leq \lambda a^2 + (1 - \lambda)b^2$ del capitolo «Funzioni convesse e concave».

Esempio 4.8: una funzione quadratica convessa

With $\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$, per ogni $\mathbf{x} \in \mathbb{R}^2$ abbiamo:

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

Quindi, la matrice $\mathbf{Q}$ è semidefinita positiva e, per il corollario, la funzione quadratica degli esempi precedenti è convessa nel suo dominio.

La stessa conclusione segue dalla definizione: per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^2$ e ogni $\lambda \in [0, 1]$, abbiamo:

$$\begin{aligned} & f(\lambda\mathbf{p} + (1 - \lambda)\mathbf{w}) \\ &= (\lambda p_1 + (1 - \lambda)w_1)^2 + (\lambda p_2 + (1 - \lambda)w_2)^2 \\ & \quad + (\lambda(p_1 - p_2) + (1 - \lambda)(w_1 - w_2))^2 \\ &\leq \lambda p_1^2 + (1 - \lambda)w_1^2 + \lambda p_2^2 + (1 - \lambda)w_2^2 \\ & \quad + \lambda(p_1 - p_2)^2 + (1 - \lambda)(w_1 - w_2)^2 \\ & \quad \text{(per la disuguaglianza del capitolo «Funzioni convesse e concave»,} \\ & \quad \text{tre volte)} \\ &= \lambda(p_1^2 + p_2^2 + (p_1 - p_2)^2) \\ & \quad + (1 - \lambda)(w_1^2 + w_2^2 + (w_1 - w_2)^2) \\ &= \lambda f(\mathbf{p}) + (1 - \lambda)f(\mathbf{w}) \end{aligned}$$

Quindi,

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

Per natura, il dominio $\mathbb{R}^2$ è un insieme convesso. Di conseguenza, la funzione è convessa nel suo dominio.

Esempio 4.9: una funzione quadratica non convessa

With $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, la funzione quadratica è

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

e i suoi insiemi di livello non vuoti con $k \neq 0$ sono iperboli. La matrice $\mathbf{Q}$ non è semidefinita positiva, poiché:

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

Per il corollario, la funzione non è convessa nel suo dominio. Infatti, con

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

la disuguaglianza della definizione di funzione convessa non è verificata:

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

4.4 Funzioni quadratiche con una parte lineare

- Nei problemi di ottimizzazione si incontrano spesso funzioni che sono la somma di una funzione quadratica, di una funzione lineare e di una costante: data una matrice quadrata simmetrica $\mathbf{Q} \in \mathbb{R}^{n \times n}$, un vettore $\mathbf{c} \in \mathbb{R}^n$ e uno scalare $d \in \mathbb{R}$,

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

- Poiché la derivata di $\mathbf{c}'\mathbf{x} + d$ rispetto a x_k è c_k , il gradiente e la matrice Hessiana sono:

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

- La funzione $h(\mathbf{x}) = \mathbf{c}'\mathbf{x} + d$ è convessa (e concava) nel suo dominio, poiché la definizione

vale con l'uguaglianza: per ogni coppia di punti $\mathbf{p}, \mathbf{w} \in \mathbb{R}^n$ e ogni $\lambda \in [0, 1]$, abbiamo:

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

Quindi, entrambe le disuguaglianze delle definizioni di funzione convessa e concava valgono con l'uguaglianza. Per natura, il dominio $\mathbb{R}^n$ è un insieme convesso. Di conseguenza, la funzione h è convessa e concava nel suo dominio. Poiché la somma di funzioni convesse è convessa (capitolo «Funzioni convesse e concave»), se la matrice $\mathbf{Q}$ è semidefinita positiva, la funzione g è convessa nel suo dominio. Viceversa, se g è convessa, la funzione quadratica $\mathbf{x}'\mathbf{Q}\mathbf{x} = g(\mathbf{x}) + (-\mathbf{c}'\mathbf{x} - d)$ è convessa, in quanto somma di due funzioni convesse, e la matrice $\mathbf{Q}$ è semidefinita positiva per il corollario.

Una funzione quadratica $\mathbf{x}'\mathbf{Q}\mathbf{x}$ è convessa nel suo dominio se e solo se la matrice $\mathbf{Q}$ è semidefinita positiva. L'aggiunta di una funzione lineare e di una costante non ne modifica la convessità.