

Funzioni quadratiche



Indice

Funzioni quadratiche

Gradiente, matrice Hessiana e insiemi di livello

Convessità delle funzioni quadratiche

Funzioni quadratiche con una parte lineare

Funzioni quadratiche



Introduzione: Funzioni quadratiche

- ▶ In questa sezione definiamo le **funzioni quadratiche**: funzioni di più variabili costruite a partire da una **matrice quadrata** Q , scritte in forma compatta come $x'Qx$.
- ▶ Dopo le funzioni lineari, sono le funzioni di più variabili più semplici, e compaiono in molti problemi di ottimizzazione.
- ▶ Vedremo che la matrice Q si può sempre supporre **simmetrica**: questa scelta semplifica il gradiente, la matrice Hessiana e lo studio della convessità nelle prossime sezioni.

Funzione quadratica (1/2)

Definizione 1: funzione quadratica

Data una matrice quadrata simmetrica

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

Funzione quadratica (2/2)

Definizione 1: funzione quadratica

la funzione (scalare) associata

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

*si chiama **funzione quadratica**.*

La definizione richiede una matrice simmetrica.
È una restrizione? Vediamo ora che ogni matrice quadrata può essere sostituita da una simmetrica, senza cambiare la funzione.

Matrice simmetrica (1/2)

- 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:

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

Matrice simmetrica (2/2)

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

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

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 $\mathbf{x} \in \mathbb{R}^n$.

Nessuna perdita di generalità

Senza perdita di generalità, possiamo supporre che la matrice Q di una funzione quadratica sia simmetrica.

Esempio: matrici equivalenti (1/2)

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:

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

Esempio: matrici equivalenti (2/2)

Infatti, le due matrici definiscono la stessa funzione:

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

Funzione quadratica:

- ▶ Data una matrice quadrata $\mathbf{Q} \in \mathbb{R}^{n \times n}$:
$$f(\mathbf{x}) = \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j = \mathbf{x}' \mathbf{Q} \mathbf{x}$$

Matrice simmetrica:

- ▶ $\mathbf{Q}$ e $\frac{1}{2}(\mathbf{Q} + \mathbf{Q}')$ definiscono la stessa funzione quadratica
- ▶ Senza perdita di generalità, $\mathbf{Q}$ è simmetrica

Gradiente, matrice Hessiana e insiemi di livello

Introduzione: Gradiente, matrice Hessiana e insiemi di livello

- ▶ In questa sezione calcoliamo il **gradiente** e la **matrice Hessiana** di una funzione quadratica: hanno una forma molto semplice, 2×2 Q e 2×2 Q .
- ▶ Poi studiamo gli **insiemi di livello**, gli insiemi dei punti in cui la funzione è costante, e le loro forme (ellissi, iperboli, coppie di rette).
- ▶ Con esempi in due variabili disegniamo insiemi di livello, gradienti e grafici: aiutano a vedere come si comporta una funzione quadratica.

Osservazione 1: gradiente di una funzione quadratica

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

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

Osservazione 1: gradiente di una funzione quadratica

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

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 $q_{kj} x_j$;
- ▶ i termini $q_{ik} x_i x_k$ contribuiscono 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:

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

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

Osservazione 2: matrice Hessiana di una funzione quadratica

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

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

Matrice Hessiana (2/2)

Osservazione 2: matrice Hessiana di una funzione quadratica

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

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 $h \in \{1, 2, \dots, n\}$. □

Gradiente e matrice Hessiana descrivono come cambia la funzione.

Ci poniamo ora la domanda opposta:
dove la funzione **non** cambia?

Definizione 2: 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:*

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

Forme degli insiemi di livello (1/2)

A seconda della matrice $\mathbf{Q}$, gli insiemi di livello non vuoti possono essere:

- ▶ **ellissoidi** (ellissi in $\mathbb{R}^2$);
- ▶ **iperboloidi** (iperboli in $\mathbb{R}^2$);
- ▶ quando alcune direzioni non cambiano il valore della funzione, **cilindri** (coppie di rette parallele in $\mathbb{R}^2$).

Forme degli insiemi di livello (2/2)

Per $k = 0$ ci sono anche casi degeneri:

- ▶ il solo punto 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;
- ▶ 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**.

Gradiente, matrice Hessiana e insiemi di livello

- ▶ 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.

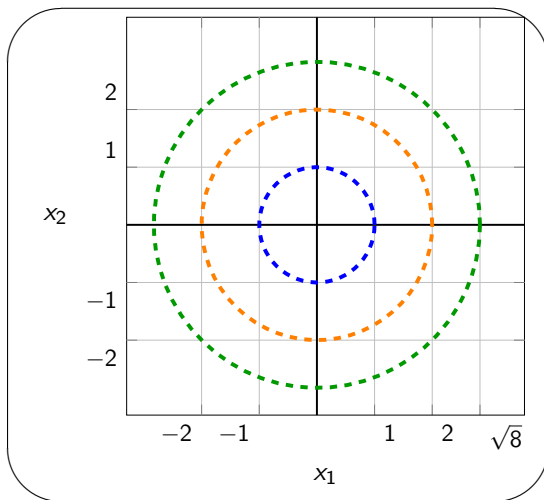
Vediamo ora queste idee su due funzioni quadratiche di due variabili, con la matrice identità e con una matrice generica:
insiemi di livello, gradienti e grafici.

Esempio: insiemi di livello, matrice identità (1/3)

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

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

Esempio: insiemi di livello, matrice identità (2/3)



Esempio: insiemi di livello, matrice identità (3/3)

La figura mostra:

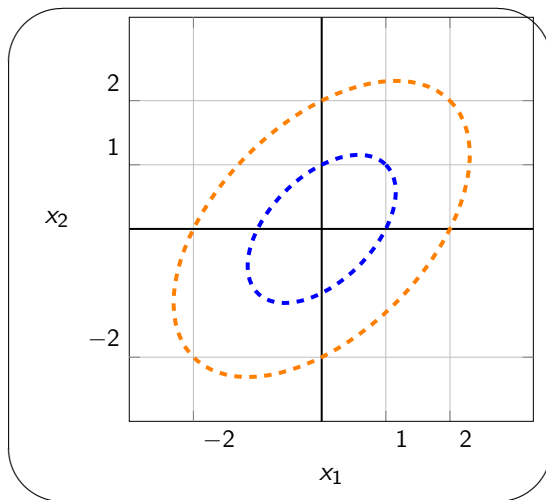
- ▶ Circonferenza blu: insieme di livello $x_1^2 + x_2^2 = 1$ (raggio 1)
- ▶ Circonferenza arancione: insieme di livello $x_1^2 + x_2^2 = 4$ (raggio 2)
- ▶ Circonferenza verde: insieme di livello $x_1^2 + x_2^2 = 8$ (raggio $\sqrt{8}$)

Esempio: insiemi di livello, matrice generica (1/3)

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

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

Esempio: insiemi di livello, matrice generica (2/3)



Esempio: insiemi di livello, matrice generica (3/3)

La figura mostra:

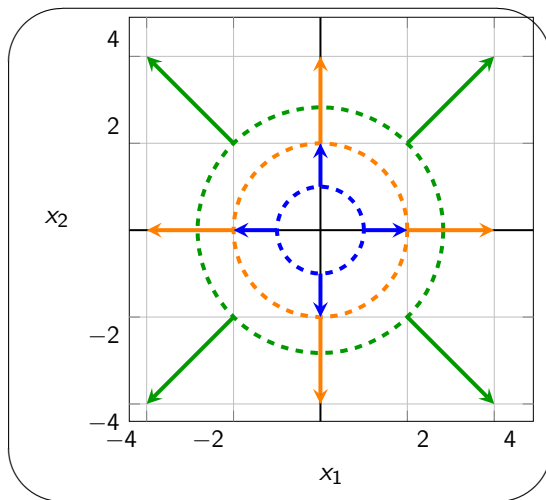
- ▶ Ellisse blu: insieme di livello $2 x_1^2 - 2 x_1 x_2 + 2 x_2^2 = 2$
- ▶ Ellisse arancione: insieme di livello $2 x_1^2 - 2 x_1 x_2 + 2 x_2^2 = 8$

Esempio: gradienti, matrice identità (1/4)

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

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

Esempio: gradienti, matrice identità (2/4)



Esempio: gradienti, matrice identità (3/4)

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 $x_1^2 + x_2^2 = 4$):

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

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

Esempio: gradienti, matrice identità (4/4)

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

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

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

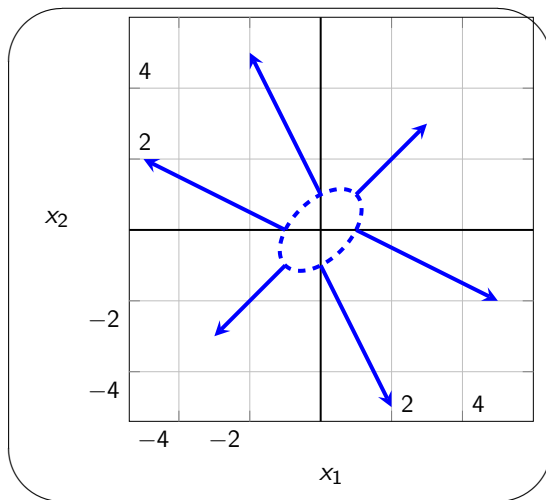
Ogni gradiente è perpendicolare all'insieme di livello passante per il suo punto.

Esempio: gradienti, matrice generica (1/3)

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

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

Esempio: gradienti, matrice generica (2/3)



Esempio: gradienti, matrice generica (3/3)

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):

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

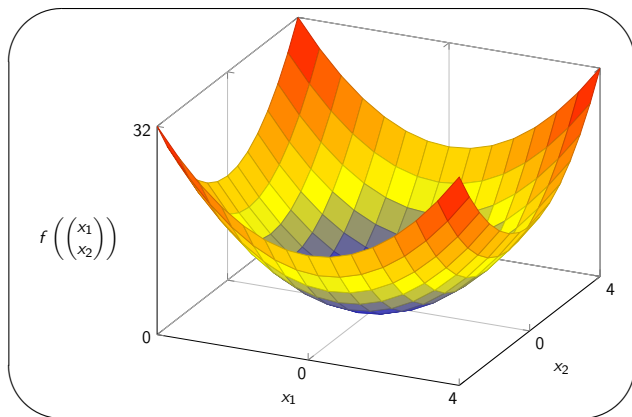
$$\nabla f \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \quad \nabla f \left(\begin{pmatrix} -1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} -4 \\ 2 \end{pmatrix},$$

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

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

Esempio: grafico, matrice identità

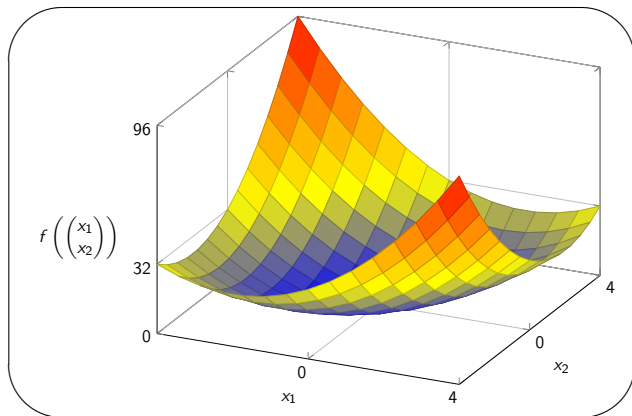
Con $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, abbiamo $f\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = x_1^2 + x_2^2$. Il suo grafico è:



Esempio: grafico, matrice generica

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



Riepilogo (1/2)

Gradiente e matrice Hessiana:

- ▶ Gradiente: $\nabla f(\mathbf{x}) = 2 \mathbf{Q} \mathbf{x}$
- ▶ Matrice Hessiana: $\nabla^2 f(\mathbf{x}) = 2 \mathbf{Q}$, una matrice costante

Insiemi di livello:

- ▶ $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}' \mathbf{Q} \mathbf{x} = k\}$, per ogni $k \in \mathbb{R}$
- ▶ Ellissoidi, iperboloidi o cilindri
- ▶ Per $k = 0$: casi degeneri (un punto, un cono, un sottospazio lineare)

Riepilogo (2/2)

Esempi in due variabili:

- ▶ Matrice identità: gli insiemi di livello sono circonferenze
- ▶ $Q = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$: gli insiemi di livello sono ellissi
- ▶ Il gradiente è perpendicolare all'insieme di livello passante per il suo punto
- ▶ I grafici sono paraboloidi

Convessità delle funzioni quadratiche



Introduzione: Convessità delle funzioni quadratiche

- ▶ In questa sezione rispondiamo alla domanda: **quando una funzione quadratica è convessa?**
- ▶ La risposta dipende solo dalla matrice Q : la funzione è convessa se e solo se Q è **semidefinita positiva**.
- ▶ Dimostriamo entrambe le implicazioni a partire dalla definizione di funzione convessa, e vediamo un esempio di funzione quadratica convessa e uno di funzione quadratica non convessa.

Matrice semidefinita positiva

Definizione 3: matrice semidefinita positiva

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

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

Mostriamo ora che questa proprietà della matrice
è esattamente ciò che rende convessa la funzione quadratica.

Prima: semidefinita positiva $\Rightarrow$ convessa.

Matrice semidefinita positiva implica funzione convessa

Proposizione 1

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

- ▶ La dimostrazione parte da $f(\lambda \mathbf{p} + (1 - \lambda) \mathbf{w})$ e arriva a $\lambda f(\mathbf{p}) + (1 - \lambda) f(\mathbf{w})$ in un'unica catena.
- ▶ Il passaggio chiave: la differenza fra i due membri è $\lambda(1 - \lambda)(\mathbf{p} - \mathbf{w})' \mathbf{Q} (\mathbf{p} - \mathbf{w})$.

Dimostrazione

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

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

Dimostrazione (continua)

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

$$\begin{aligned} &(\text{poiché } \lambda^2 = \lambda - \lambda (1 - \lambda) \text{ e} \\ &\quad (1 - \lambda)^2 = (1 - \lambda) - \lambda (1 - \lambda)) \end{aligned}$$

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

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

$$(\text{poiché } \mathbf{Q} \text{ è semidefinita positiva e } \lambda (1 - \lambda) \geq 0)$$

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

Dimostrazione (continua)

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. □

Condizione sufficiente e condizione necessaria

- ▶ Abbiamo dimostrato che:

Q semidefinita positiva $\implies$ 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.

Funzione convessa implica matrice semidefinita positiva

Proposizione 2

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

- ▶ La dimostrazione parte dalla disuguaglianza della definizione di funzione convessa, ne sviluppa il membro sinistro e semplifica.
- ▶ Ciò che resta è $(\mathbf{p} - \mathbf{w})' \mathbf{Q} (\mathbf{p} - \mathbf{w}) \geq 0$, per ogni $\mathbf{p}$ e $\mathbf{w}$.

Dimostrazione

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

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

dove il membro sinistro è:

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

Dimostrazione (continua)

Riordinando e semplificando, ciò implica che:

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

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

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

Dimostrazione (continua)

Ogni vettore $\mathbf{x} \in \mathbb{R}^n$ si può scrivere come $\mathbf{x} = \mathbf{p} - \mathbf{w}$ (per esempio, con $\mathbf{p} = \mathbf{x}$ e $\mathbf{w} = 0$).

Quindi, $\mathbf{x}'\mathbf{Q}\mathbf{x} \geq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$, e la matrice quadrata $\mathbf{Q}$ è semidefinita positiva. □

Condizione necessaria e sufficiente (1/2)

- ▶ Abbiamo dimostrato che:

funzione quadratica convessa $\implies Q$ 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”.

Condizione necessaria e sufficiente (2/2)

- ▶ Abbiamo quindi, infine:

Q semidefinita positiva $\iff$ funzione quadratica convessa

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

Convessità di una funzione quadratica

Corollario 1

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

Due osservazioni (1/2)

- 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

$$\mathbf{x}'\mathbf{x} = x_1^2 + x_2^2 + \cdots + x_n^2 \geq 0$$

in accordo con il corollario.

Due osservazioni (2/2)

- ▶ 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.
- ▶ 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: una funzione quadratica convessa (1/3)

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

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

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

Esempio: una funzione quadratica convessa (2/3)

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:

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

Esempio: una funzione quadratica convessa (3/3)

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

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: una funzione quadratica non convessa (1/2)

Con $\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, 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:

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

Esempio: una funzione quadratica non convessa (2/2)

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:

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

Matrice semidefinita positiva:

- ▶ $\mathbf{x}' \mathbf{Q} \mathbf{x} \geq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$

Convessità delle funzioni quadratiche:

- ▶ $\mathbf{Q}$ semidefinita positiva $\iff$ la funzione quadratica è convessa
- ▶ Entrambe le implicazioni si dimostrano dalla definizione di funzione convessa
- ▶ Somma di quadrati di funzioni lineari: $\mathbf{Q}$ è semidefinita positiva
- ▶ $x_1^2 - x_2^2$ non è convessa: $\mathbf{Q}$ non è semidefinita positiva

Funzioni quadratiche con una parte lineare

Introduzione: Funzioni quadratiche con una parte lineare

- ▶ Nei problemi di ottimizzazione l'obiettivo è spesso una funzione quadratica **più** una funzione lineare e una costante.
- ▶ In questa sezione calcoliamo il gradiente e la matrice Hessiana di queste funzioni.
- ▶ Mostriamo che la parte lineare **non** cambia la convessità: la funzione è ancora convessa se e solo se **Q** è semidefinita positiva.

Funzione quadratica 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 $d \in \mathbb{R}$,

$$g : \mathbb{R}^n \rightarrow \mathbb{R}, \quad g : \mathbf{x} \mapsto \mathbf{x}' \mathbf{Q} \mathbf{x} + \mathbf{c}' \mathbf{x} + d$$

$$= \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j + \sum_{j=1}^n c_j x_j + d$$

Gradiente e matrice Hessiana

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 matrice Hessiana non dipende dalla parte lineare: è la stessa di $\mathbf{x}' \mathbf{Q} \mathbf{x}$.

La parte lineare è convessa e concava

- ▶ La funzione $h(\mathbf{x}) = \mathbf{c}'\mathbf{x} + d$ è convessa (e concava) nel suo dominio, poiché la definizione vale con l'uguaglianza.
- ▶ La dimostrazione parte da $h(\lambda\mathbf{p} + (1 - \lambda)\mathbf{w})$ e arriva a $\lambda h(\mathbf{p}) + (1 - \lambda)h(\mathbf{w})$ usando solo uguaglianze.

Dimostrazione

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. □

Convessità di g

- ▶ 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.

La parte lineare non cambia la convessità

Una funzione quadratica $x'Qx$ è convessa nel suo dominio se e solo se la matrice Q è semidefinita positiva.

L'aggiunta di una funzione lineare e di una costante non ne modifica la convessità.

Funzione quadratica con una parte lineare:

- ▶ $g(\mathbf{x}) = \mathbf{x}' \mathbf{Q} \mathbf{x} + \mathbf{c}' \mathbf{x} + d$
- ▶ $\nabla g(\mathbf{x}) = 2 \mathbf{Q} \mathbf{x} + \mathbf{c}$ e $\nabla^2 g(\mathbf{x}) = 2 \mathbf{Q}$

Convessità:

- ▶ $\mathbf{c}' \mathbf{x} + d$ è convessa e concava
- ▶ g è convessa se e solo se $\mathbf{Q}$ è semidefinita positiva

Messaggi chiave

Riepilogo (1/2)

1. Funzioni quadratiche:

- ▶ $f(\mathbf{x}) = \mathbf{x}' \mathbf{Q} \mathbf{x}$, con $\mathbf{Q}$ simmetrica senza perdita di generalità

2. Gradiente, matrice Hessiana e insiemi di livello:

- ▶ $\nabla f(\mathbf{x}) = 2 \mathbf{Q} \mathbf{x}$ e $\nabla^2 f(\mathbf{x}) = 2 \mathbf{Q}$
- ▶ Insiemi di livello: ellissoidi, iperboloidi, cilindri (e casi degeneri per $k = 0$)
- ▶ Il gradiente è perpendicolare agli insiemi di livello

Riepilogo (2/2)

3. Convessità:

- ▶ La funzione quadratica è convessa $\iff Q$ è semidefinita positiva

4. Parte lineare:

- ▶ $x'Qx + c'x + d$ è convessa $\iff Q$ è semidefinita positiva