Classificazione dei problemi di ottimizzazione

- Ottimizzazione
  - Non vincolata ($\Omega = \mathbb{R}^n$)
    - Differenziabile
    - Non differenziabile
  - Vincolata ($\Omega \subset \mathbb{R}^n$)
    - Non differenziabile
    - Differenziabile
      - Lineare
      - Non Lineare

Ottimizzazione statica (Cerchiamo una soluzione ottima in $\mathbb{R}^n$)

Controllo ottimo (Cerchiamo una soluzione ottima in uno spazio di funzioni)

Sia dato un sistema $\dot{x}(t) = f(x(t), u(t))$ Si cerca un controllo $u(t)$, $t \in [t_0, t_f]$ tale che sia minimizzato

$$J = \beta(x(t_f), t_f) + \int_{t_0}^{t_f} h(x(t), u(t))\, dt$$

Richiami sulle funzioni di più variabili

$f : \mathbb{R}^n \to \mathbb{R}$; Derivata parziale $\dfrac{\partial f}{\partial x_i}(x) = \lim_{h \to 0} \dfrac{f(x + h e_i) - f(x)}{h}$

$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \to i$ i-mo versore (vettore unitario)

$\nabla f(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(x) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(x) \end{pmatrix}$; Es. $f(x_1, x_2) = \sqrt{x_1^2 + x_2^2} = (x_1^2 + x_2^2)^{\frac{1}{2}} = \|x\|_2$

La funzione è derivabile per $(x_1, x_2) \neq (0,0)$

ed è $\dfrac{\partial f}{\partial x_1}(x) = \dfrac{x_1}{\sqrt{x_1^2 + x_2^2}}$; $\dfrac{\partial f}{\partial x_2}(x) = \dfrac{x_2}{\sqrt{x_1^2 + x_2^2}}$

Calcoliamo il rapporto incrementale in $(0,0)$ lungo x_1

$$\frac{f(0+h, 0) - f(0)}{h} = \frac{\sqrt{h^2}}{h} \begin{cases} 1 & \text{se } h > 0 \\ -1 & \text{se } h < 0 \end{cases} \Rightarrow \text{il limite non esiste}$$

situazione analoga lungo l'asse x_2

Differenza tra funzioni di una sola e di più variabili

Per funzioni di una sola variabile: derivabilità $\Rightarrow$ continuità ($\not\Leftarrow$)

Una funzione di più variabili può essere derivabile (cioè ammettere derivate parziali) in un punto senza essere continua.

Es. $$f(x_1, x_2) = \begin{cases} 0 & \text{per } x = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{per } x \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$$

Verifichiamo se esiste il limite del rapporto incrementale in $(0,0)$ lungo la direzione x_1

$$\frac{f(0+h, 0) - f(0,0)}{h} = \frac{(0+h)\cdot 0}{(0+h)^2 + 0} / h = \frac{0}{h^3} = 0$$

$\Rightarrow$ Rapporto incrementale costante e uguale a zero

$\Rightarrow$ limite esiste e vale 0 $\Rightarrow \dfrac{\partial f}{\partial x_1}(0,0) = 0$

Analogamente si può verificare che $\dfrac{\partial f}{\partial x_2}(0,0) = 0$

La funzione è <u>derivabile</u> nel punto $(0,0)$, ma in tale punto non è continua. Infatti si consideri la retta di equazione $x_1 = x_2$. Su di essa f è costante e vale $\frac{1}{2}$, anche infinitamente vicino a $(0,0)$, dove invece f vale 0.

$$\Downarrow$$

In questo caso le derivate parziali non sono utilizzabili come <u>strumento di approssimazione</u> della funzione, come invece accade per le funzioni di <u>una sola variabile</u>:

$$f(x+h) = \underbrace{f(x) + f'(x)h}_{\text{funzione lineare}} + o(h) \quad \text{con} \quad \lim_{h \to 0} \frac{o(h)}{h} = 0$$

Dimostrare. Verificare che se si considera un'approssimazione del tipo $f(x) + ah$ con $a \neq f'(x)$ sussiste $\lim_{h \to 0} f(x) + ah = f(x)$ <u>ma non</u> $\lim \frac{o(h)}{h} = 0$

Definizione. $f: \mathbb{R}^n \to \mathbb{R}$ è differenziabile in x se

a) $\exists\, \nabla f(x)$ b) $\displaystyle\lim_{\|d\| \to 0} \frac{f(x+d) - f(x) - \nabla f(x)^T d}{\|d\|} = 0$

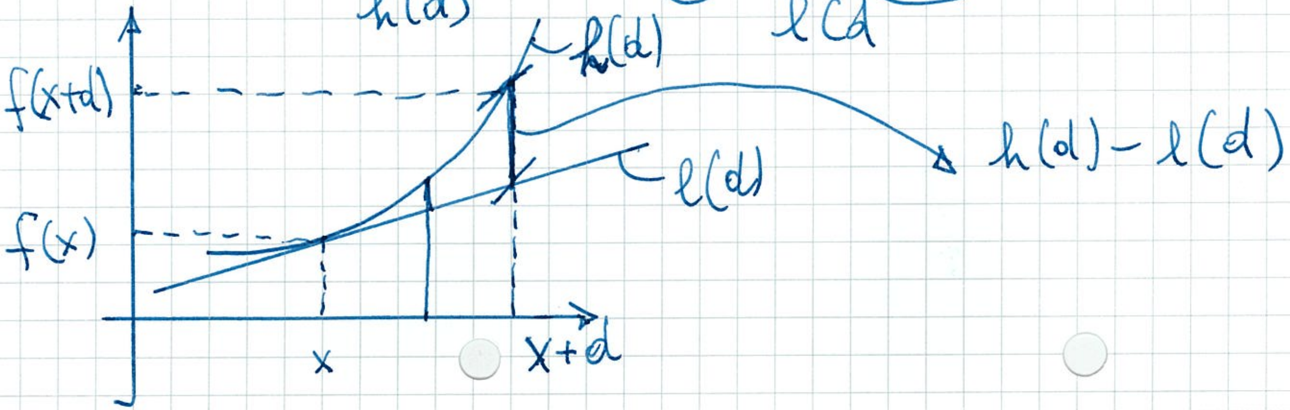
In altri termini f è differenziabile in x se essa è approssimabile da una forma lineare.

Fissato x, $f(x+d)$ dipende solo da $d \in \mathbb{R}^n$, poniamo quindi $h(d) = f(x+d)$. Definiamo una funzione lineare (affine) di d

$\ell(d) = a + b^T d$ con $a = f(x)$, $b = \nabla f(x)$

La differenziabilità di f in x ci dice che la differenza $h(d) - \ell(d)$ è un infinitesimo di ordine superiore rispetto a $\|d\|$

cioè $$\underbrace{f(x+d)}_{h(d)} - \Big[\underbrace{f(x) + \nabla f(x)^T d}_{\ell(d)}\Big] = O(\|d\|) \quad (*)$$



Nota differenziabilità $\Rightarrow$ continuità.

Infatti $\lim_{\|d\|\to 0} f(x+d) = \lim_{\|d\|\to 0} f(x) + \nabla f(x)^T d + o(\|d\|) = f(x)$

da (*)

La funzione lineare $\nabla f(x)^T d$ prende il nome di differenziale di f in x

Teorema del differenziale. $f: \mathbb{R}^n \to \mathbb{R}$; derivate parziali continue in x $\left(\frac{\partial f}{\partial x_i}(y) \xrightarrow[y \to x]{} \frac{\partial f}{\partial x_i}(x) \quad \forall i = 1, \dots, n \right)$

$\Rightarrow$ f è differenziabile in x

Significato intuitivo: se le derivate parziali sono continue in x, esse "catturano" tutto quello che accade nell'intorno di x e non solo lungo gli assi coordinati. Quindi sono in grado di fornire un'approssimazione complessiva di f.

Esercizio. Verificare che per la funzione nondifferenziabile in $(0,0)$

$$f(x_1,x_2) = \begin{cases} 0 & \text{se } x_1 = x_2 = 0 \\ \dfrac{x_1 x_2}{x_1^2 + x_2^2} & \text{altrimenti} \end{cases}$$

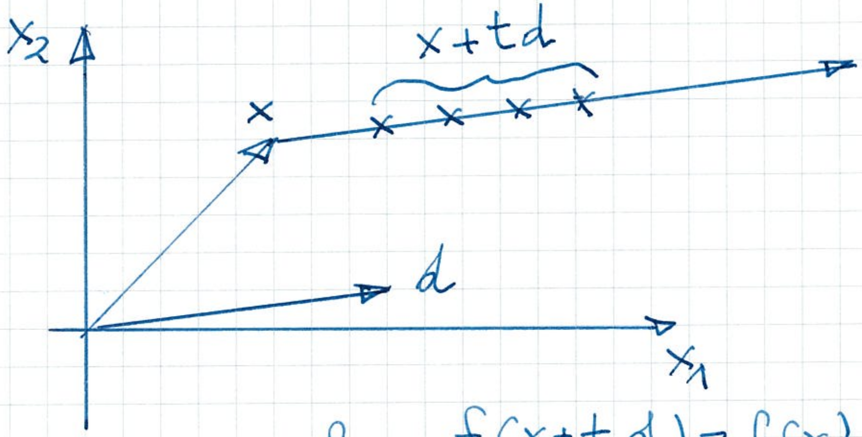
Le derivate parziali non sono continue in $(0,0)$.

Nota. Il teorema del differenziale esprime una condizione sufficiente ma non necessaria: una funzione può essere differenziabile senza che le derivate parziali siano continue.

Se una funzione è differenziabile in tutti i punti si dirà semplicemente differenziabili. Per indicare una funzione differenziabile è usuale la notazione "$f \in C^{(1)}$" dove con $C^{(1)}$ si indica la classe delle funzioni continue e con derivate parziali continue.

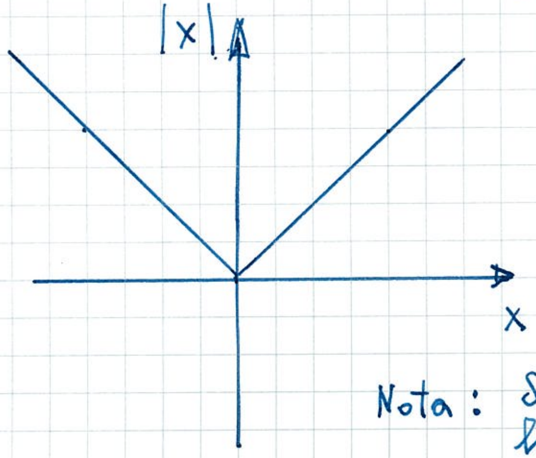
Con $C^{(n)}$ si intenderà la classe delle funzioni continue "fino alle derivate n-me")

Derivata direzionale



Si considera la semiretta
$$\{y \mid y = x+td,\ t \geq 0\}$$
Si definisce "derivata direzionale in x lungo la direzione d" il limite (se esiste finito)
$$\lim_{t \to 0^+} \frac{f(x+td)-f(x)}{t}$$
; Per la derivata direzionale si usa la notazione $f'(x,d)$

Esempio. La funzione $f(x) = |x|$. Non è derivabile in 0
$$\left(\lim_{t \to 0} \frac{|0+t|-|0|}{t} = \begin{cases} 1 & \text{se } t>0 \\ -1 & \text{se } t<0 \end{cases} \right)$$
Ammette derivata direzionale sia nella direzione $d=1$ che $d=-1$



$$d = 1 \qquad f'(0,d) = \lim_{t \to 0^+} \frac{|0+t|-|0|}{t} = 1$$

$$d = -1 \qquad f'(0,d) = \lim_{t \to 0^+} \frac{|0-t|-|0|}{t} = 1$$

Nota: Se da 0 ci si muove sia in direzione $d=1$ che $d=-1$ la funzione cresce.

Derivate parziali e derivate direzionali lungo gli assi

Sia $f : \mathbb{R}^n \to \mathbb{R}$ derivabile in $x \Rightarrow \exists \; \forall i \quad \frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0} \frac{f(x+te_i)-f(x)}{t}$

$$= \lim_{t \to 0^+} \frac{f(x+te_i)-f(x)}{t} = \lim_{t \to 0^-} \frac{f(x+te_i)-f(x)}{t}$$

Consideriamo ora

$$f'(x, e_i) = \lim_{t \to 0^+} \frac{f(x+te_i)-f(x)}{t} = \frac{\partial f}{\partial x_i}(x)$$

$$f'(x, -e_i) = \lim_{t \to 0^+} \frac{f(x-te_i)-f(x)}{t} = \lim_{h \to 0^-} \frac{f(x+he_i)-f(x)}{-h} =$$

ponendo $h = -t$

$$= -\lim_{h \to 0^-} \frac{f(x+he_i)-f(x)}{h} = -\frac{\partial f}{\partial x_i}(x)$$

Se f è differenziabile, la derivata direzionale esiste in tutte le direzioni ed è

$$f'(x, d) = \nabla f(x)^T d$$

Infatti

$$f'(x, d) = \lim_{t \to 0^+} \frac{f(x+td) - f(x)}{t} = \lim_{t \to 0^+} \frac{\cancel{f(x)} + t\,\nabla f(x)^T d + o(t\|d\|) - \cancel{f(x)}}{t} =$$

$$= \nabla f(x)^T d + \lim_{t \to 0^+} \frac{o(t\|d\|)}{t} = \nabla f(x)^T d$$

Se la funzione è differenziabile, il gradiente contiene informazioni capaci di "ricostruire" l'andamento di f in tutto l'intorno di x, cioè lungo tutte le direzioni d.

Matrice Jacobiana. Sia $g : \mathbb{R}^n \to \mathbb{R}^m$ (g è un vettore di m funzioni) $g(x) = \begin{pmatrix} g_1(x) \\ \vdots \\ g_m(x) \end{pmatrix} = \begin{pmatrix} g_1(x_1, \dots, x_n) \\ \vdots \\ g_m(x_1, \dots, x_n) \end{pmatrix}$

si definisca "Jacobiano" di g in x la matrice $J(x) = \begin{pmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_m(x)^T \end{pmatrix}$

Differenziale di un vettore di funzioni $g: \mathbb{R}^n \to \mathbb{R}^m$.

g è differenziabile in x se

a) esiste lo Jacobiano $J(x)$; b) $\lim_{\|d\| \to 0} \frac{\| g(x+d) - g(x) - J(x)\, d \|}{\|d\|} = 0$

Differenziabilità: $g(x+d)$ è approssimabile dalla forma lineare $h(d) = g(x) + J(x)\, d$ $\quad (h: \mathbb{R}^n \to \mathbb{R}^m)$

Matrice delle derivate seconde (Matrice Hessiana)

Consideriamo nel punto x il gradiente $\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$
e supponiamo che ciascuna delle derivate parziali sia a sua volta derivabile. Cioè, $\forall i, j = 1, .., n$
esiste $\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}(x)\right)$ che indichiamo con $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$

"Organizziamo" queste derivate nella matrice $\nabla^2 f(x)$

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

Nota $\nabla^2 f$ coincide con la trasposta dello Jacobiano di $\nabla f(x)$

Funzione due volte differenziabile $\left(f \in C^{(2)}\right)$

Def. $f: \mathbb{R}^n \to \mathbb{R}$ è due volte differenziabile se il suo gradiente (che è una funzione $\mathbb{R}^n \to \mathbb{R}^n$) è differenziabile

Teorema del differenziale. Se in x esiste la matrice Hessiana ed è continua $\Longrightarrow$ f è due volte differenziabile e $\nabla^2 f(x)$ è simmetrica $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j} = \dfrac{\partial^2 f}{\partial x_j \partial x_i} \quad \forall\, i,j \right)$

Es. $f: \mathbb{R}^2 \to \mathbb{R} \qquad f(x) = f(x_1, x_2) = x_1^2 \operatorname{sen} x_2$

$$\nabla f(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 \operatorname{sen} x_2 \\ x_1^2 \cos x_2 \end{pmatrix}$$

$$\nabla^2 f(x)^T = \left(\nabla\left(\frac{\partial f}{\partial x_1}\right) \;\Big|\; \nabla\left(\frac{\partial f}{\partial x_2}\right) \right) = \begin{pmatrix} 2 \operatorname{sen} x_2 & 2x_1 \cos x_2 \\ 2x_1 \cos x_2 & -x_1^2 \operatorname{sen} x_2 \end{pmatrix} = \nabla^2 f(x)$$

Simmetrica

Applicando il teorema di derivazione delle funzioni composte (Suggerimento: si studi la funzione $h(t) = f(x+td)$)

Si ottiene la formula di Taylor, nelle due forme:

1) $f(x+d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x) d + o(\|d\|^2)$

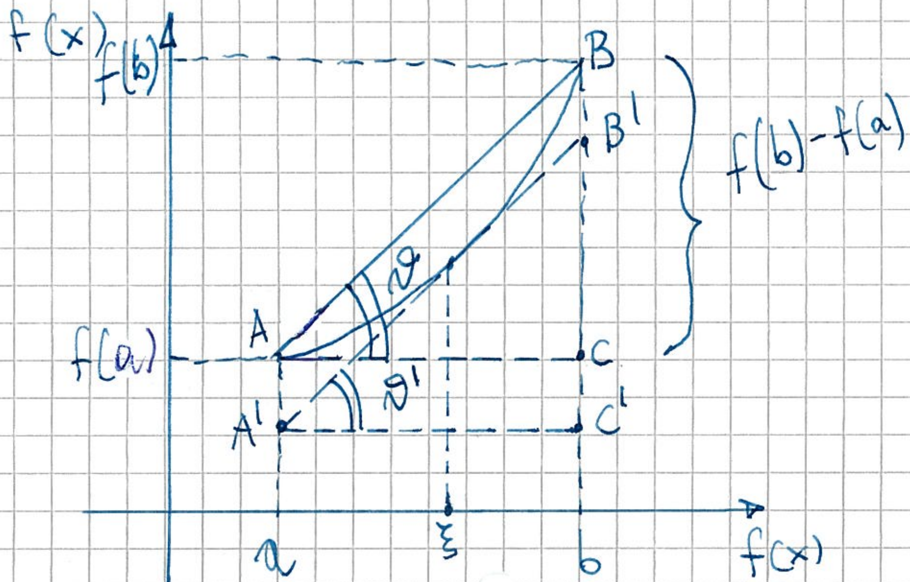
2) $f(x+d) = f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(\xi) d$

con $\xi = x + td$ con $t \in (0,1)$

Teorema del valor medio per funzioni di una sola variabile

Sia $f: \mathbb{R} \to \mathbb{R}$ continua in $[a,b]$ e derivabile in (a,b)

$\Rightarrow \exists\, \xi \in (a,b)$ tale che $f(b) = f(a) + f'(\xi)(b-a)$



$$\Downarrow$$

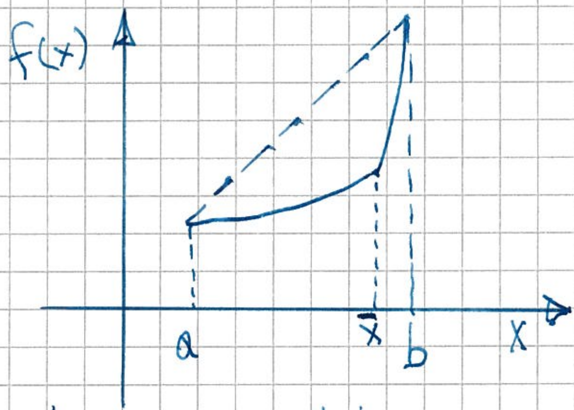
$$f'(\xi) = \frac{f(b)-f(a)}{b-a}$$

$\vartheta = \vartheta' \Rightarrow$

$$\mathrm{tg}\,\vartheta = \frac{f(b)-f(a)}{b-a} = \mathrm{tg}\,\vartheta' = f'(\xi)$$

"La pendenza della tangente in ξ è uguale alla pendenza della secante"

Nota l'ipotesi di derivabilità ovunque in (a,b) è essenziale!



Lagrange vs. Lebourg

La funzione è derivabile quasi ovunque (esiste un unico punto in cui non lo è) Tuttavia in nessun punto (dove esiste) la pendenza della tangente è uguale a quella della secante (a sinistra di $\bar{x}$ è troppo piccola e a destra troppo grande)

Teorema di Lagrange del valor medio. Sia $f: \mathbb{R}^n \to \mathbb{R} \in C^{(1)}$

$\Rightarrow \quad \forall (x,y) \qquad f(y) = f(x) + \nabla f(\xi)^T (y-x)$ per qualche $\xi = x + t(y-x)$ con $t \in (0,1)$

Formulazione equivalente (ponendo $y = x + d$)

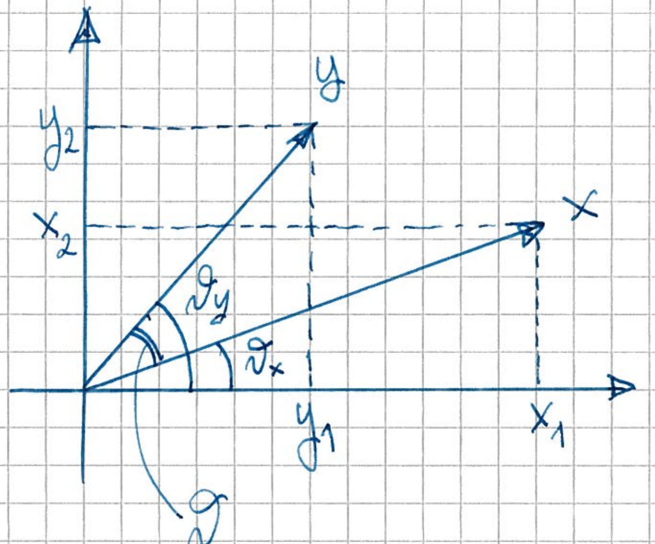
$$f(x+d) = f(x) + \nabla f(\xi)^T d$$

con $\xi = x + td, \quad t \in (0,1)$

Definizione. Angolo tra due vettori $x \in \mathbb{R}^n$, $y \in \mathbb{R}^n$ (entrambi non nulli)

$$\cos \vartheta = \frac{x^T y}{\|x\| \, \|y\|}$$

In $\mathbb{R}^n$ non c'è una interpretazione geometrica del concetto di angolo. Si estende quindi a $\mathbb{R}^n$ una proprietà di $\mathbb{R}^2$



$$\|x\| = \sqrt{x_1^2 + x_2^2} \quad ; \quad \|y\| = \sqrt{y_1^2 + y_2^2}$$

$$x_1 = \|x\| \cos \vartheta_x \quad ; \quad x_2 = \|x\| \operatorname{sen} \vartheta_x$$

$$y_1 = \|y\| \cos \vartheta_y \quad ; \quad y_2 = \|y\| \operatorname{sen} \vartheta_y$$

$$\Downarrow$$

$$x^T y = x_1 y_1 + x_2 y_2 = \|x\| \, \|y\| \cos \vartheta_x \cos \vartheta_y + \|x\| \, \|y\| \operatorname{sen} \vartheta_x \operatorname{sen} \vartheta_y =$$

$$= \|x\| \, \|y\| \left(\cos \vartheta_x \cos \vartheta_y + \operatorname{sen} \vartheta_x \operatorname{sen} \vartheta_y \right)$$

$$= \|x\| \, \|y\| \left(\cos (\vartheta_y - \vartheta_x \right) = \|x\| \, \|y\| \cos \vartheta$$

$$\Longrightarrow \quad \cos \vartheta = \frac{x^T y}{\|x\| \, \|y\|}$$

Due vettori di $\mathbb{R}^n$, x e y, sono *ortogonali* se $\cos \vartheta = 0$ $(x^T y = 0)$
formano un angolo *acuto* se $x^T y > 0$; formano un angolo *ottuso* se $x^T y < 0$

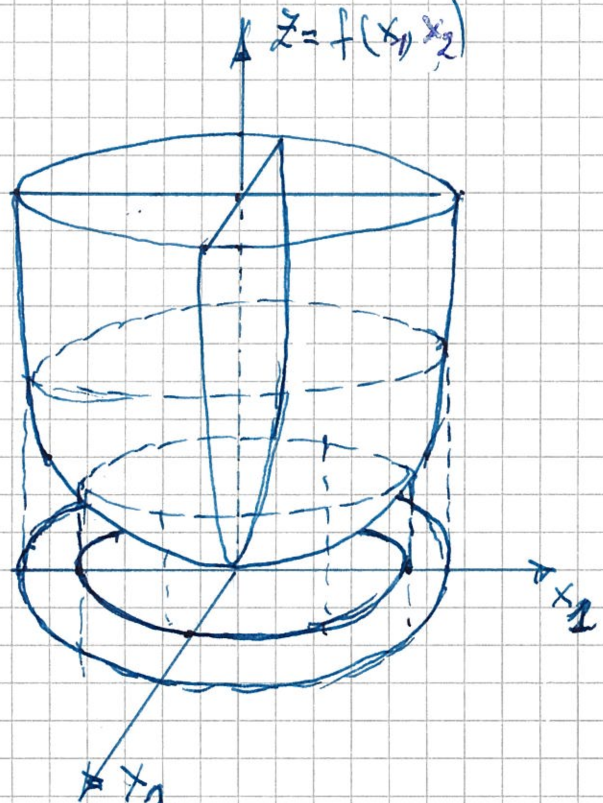


Grafico della funzione $f(x_1, x_2) = \frac{1}{2}(x_1^2 + 2x_2^2)$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad , \quad Q = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

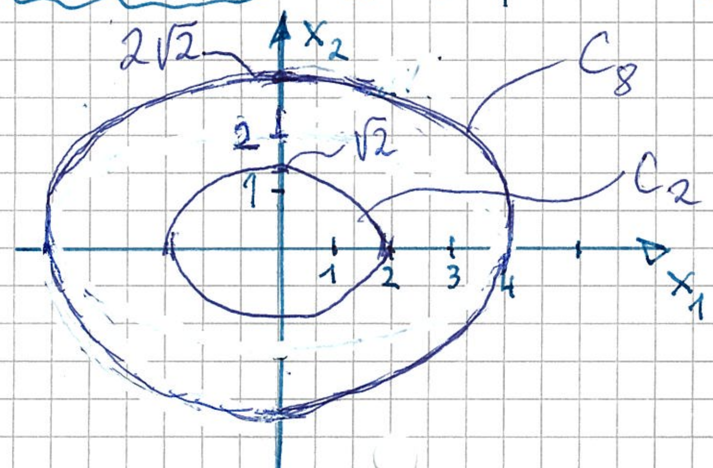
$$f(x) = \frac{1}{2} x^T Q x = \frac{1}{2}(x_1, x_2) \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\frac{\partial f}{\partial x_1} = x_1 \, , \quad \frac{\partial f}{\partial x_2} = 2x_2$$

$$\Rightarrow \quad \nabla f(x) = Q x$$

(Verificare che per una funzione quadratica $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2} x^T Q x + b^T x$ è $\nabla f(x) = Qx + b$)

(Linea)
Curva di livello $C_\alpha = \{x \mid f(x) = \alpha\}$; Insieme di livello $S_\alpha = \{x \mid f(x) \leq \alpha\}$



Linee di livello $C_8 = \{x \mid \frac{1}{2}(x_1^2 + 2x_2^2) = 8\}$

$C_2 = \{x \mid \frac{1}{2}(x_1^2 + 2x_2^2) = 2\}$

Definizione. Direzione di discesa

Data una funzione $f: \mathbb{R}^n \to \mathbb{R}$ e un punto $x \in \mathbb{R}^n$, il vettore $d \in \mathbb{R}^n$ è una direzione di discesa per f in x se esiste $\bar{t} > 0$ tale che

$$f(x + td) < f(x) \quad \forall\, t \in (0, \bar{t})$$

Concetto fondamentale nell'ottimizzazione: se in un punto x conosco una direzione di discesa, mi posso "muovere" da x lungo d riuscendo a ridurre il valore della funzione.

Teorema. Sia $f \in C^{(1)}$. Se in x per una certa direzione d sussiste $\nabla f(x)^T d < 0 \Rightarrow d$ è una direzione di discesa in x.

Prova. Poiché f è differenziabile, vale $f'(x, d) = \nabla f(x)^T d < 0$, cioè

$$\lim_{t \to 0^+} \underbrace{\frac{f(x + td) - f(x)}{t}}_{h(t)} < 0$$

Per il teorema della permanenza del segno nell'intorno (destro) di zero, $\exists\, \bar{t}$ tale che $f(x + td) - f(x) < 0$ $\forall\, t < \bar{t}$.

$f \in C^{(1)}$. In un punto x il gradiente è perpendicolare alla tangente alla curva di livello.

Prova (in $\mathbb{R}^2$). Sia $f(x) = f(x_1, x_2) = \alpha$ $(x \in C_\alpha)$ e supponiamo che la curva C_α ammetta una rappresentazione parametrica $\left(C_\alpha = r(t) = \begin{pmatrix} r_1(t) \\ r_2(t) \end{pmatrix}, \ t \in [0, T]\right)$.

Consideriamo la funzione di una sola variabile $h : \mathbb{R} \to \mathbb{R}$ definita come

$$h(t) = f(r(t))$$

Siccome $r(t)$ "disegna" la curva di livello C_α, allora f è costante e vale α, sussiste $h(t) = \alpha =$ costante $\forall t \in [0, T]$, quindi è $h'(t) = 0$.

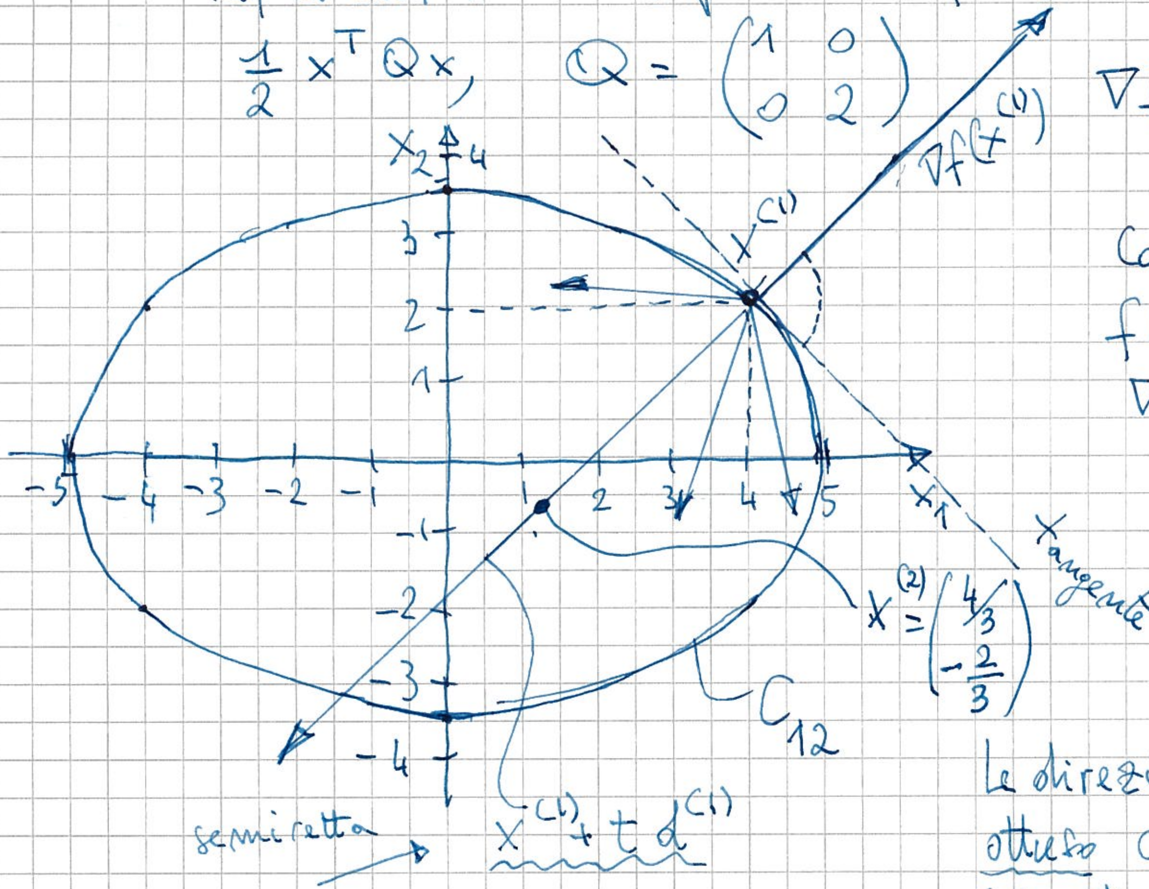
Dalla formula di derivazione delle funzioni composte sussiste

$$h'(t) = \nabla f(r(t))^T r'(t) = 0$$

dove $r'(t) = \begin{pmatrix} r_1'(t) \\ r_2'(t) \end{pmatrix}$ è la derivata della curva e ne rappresenta la tangente. Poiché $x = r(\bar{t})$ per qualche $\bar{t} \in [0, T]$ vale $\nabla f(x)^T r'(\bar{t}) = 0$, cioè il gradiente è ortogonale alla curva di livello nel punto considerato

Riprendiamo l'esempio della funzione quadratica $f(x) = \frac{1}{2}(x_1^2 + 2x_2^2) = \frac{1}{2} x^T Q x$, $Q = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$

$$\nabla f(x) = Qx = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 2x_2 \end{pmatrix}$$



Consideriamo il punto $x^{(1)} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$

$f(x^{(1)}) = 12$

$\nabla f(x^{(1)}) = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$

Trasliamo il gradiente in x e verifichiamo che è ortogonale alla tangente alla curva di livello

Le direzioni d che formano un angolo ottuso con il gradiente sono di discesa (puntano verso l'interno della curva di livello)

La direzione $d^{(1)} = -\nabla f(x^{(1)})$ è sicuramente di discesa $\left(\nabla f(x^{(1)})^T d = -\|\nabla f(x^{(1)})\|^2 < 0\right)$

Calcoliamo $f(x^{(1)} + t d^{(1)}) = f\begin{pmatrix} 4 - t4 \\ 2 - t4 \end{pmatrix} = \frac{1}{2}\left((4-t4)^2 + 2(2-t4)^2\right) =$

$= 24t^2 - 32t + 12 = h(t)$

La funzione lungo la direzione $d^{(1)}$ dipende solo da t $\Rightarrow$ "Passo"

$\Rightarrow$ Possiamo determinarne il minimo

$$h'(t) = 48t - 32 = 0 \Big|_{t=t^*} \Rightarrow t^* = \frac{32}{48} = \frac{2}{3}$$

Calcoliamo quindi il punto $x^{(2)} = x^{(1)} + t^* d^{(1)}$

$$x^{(2)} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} -4 \\ -4 \end{pmatrix} = \begin{pmatrix} \frac{4}{3} \\ -\frac{2}{3} \end{pmatrix}$$

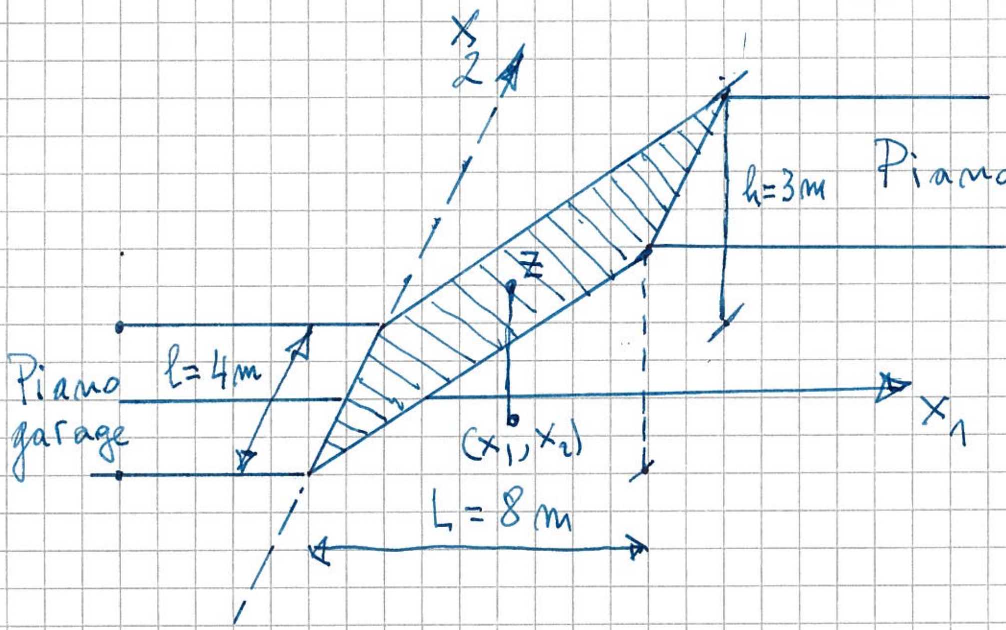
Calcoliamo $f(x^{(2)})$

$$f(x^{(2)}) = \frac{1}{2}\left(\left(\frac{4}{3}\right)^2 + 2\left(-\frac{2}{3}\right)^2\right) = \frac{24}{18} = \frac{4}{3}$$

Attraverso la minimizzazione di una funzione di una sola variabile, sfruttando una direzione di discesa abbiamo ridotto la funzione dal valore 12 a $\frac{4}{3}$ e ci siamo avvicinati al minimo (Per questa funzione è il punto $(0,0)$)

Sono stati introdotti gli elementi essenziali di un algoritmo iterativo di discesa per determinare il minimo di una funzione (differenziabile) di più variabili

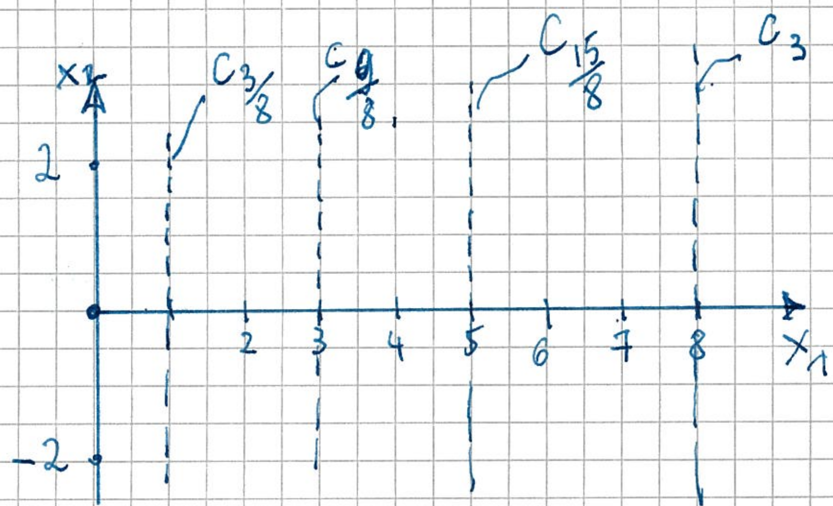
La rampa del garage



Trasportare una carriola di legna dal garage al piano strada (o viceversa !!)

Line di livello della funzione $z = f(x_1, x_2)$ (Quota)

$f(x_1, x_2) = \frac{3}{8} x_1$ (la quota è nulla lungo l'asse x_2 e vale 3 in corrispondenza di $x_1 = 8$



Il gradiente di f è costante e vale $\begin{pmatrix} 3/8 \\ 0 \end{pmatrix}$

è un vettore orientato lungo l'asse x_1, nel verso positivo.

Ricordando $f'(x, d) = \nabla f(x)^T d$, nell'origine $(0, 0)$ la direzione di salita più ripida coincide con l'asse x_1

Chi spingerebbe la carriola in direzione x_1 ?