

Angolo tra due vettori

$$x, y \in \mathbb{R}^n$$

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

$$\langle x, y \rangle = \|x\| \cdot \|y\| \cdot \cos \vartheta$$

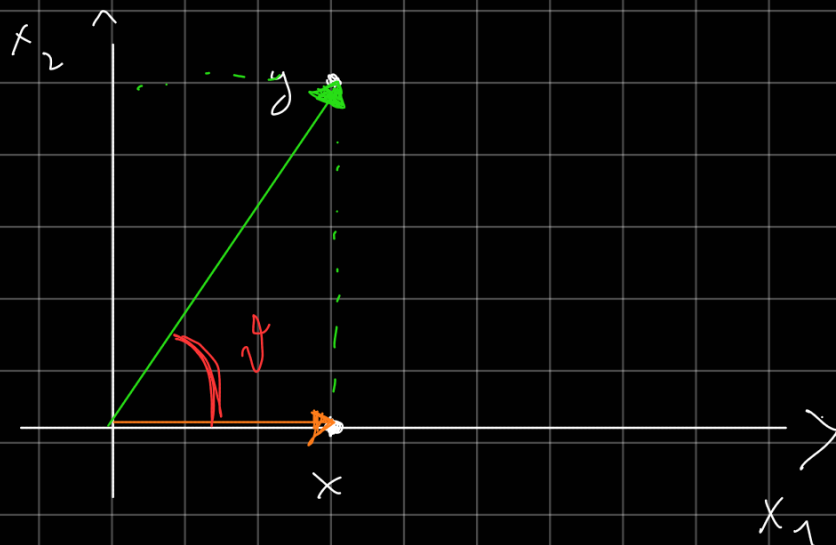
angolo compreso
tra x e y

$$x^T y = \sum_{i=1}^n x_i y_i$$

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$$

(sempre ≥ 0)

Es. sul piano



$$\mathbb{R}^2$$

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$x = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \quad y = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ 1 \end{pmatrix}$$

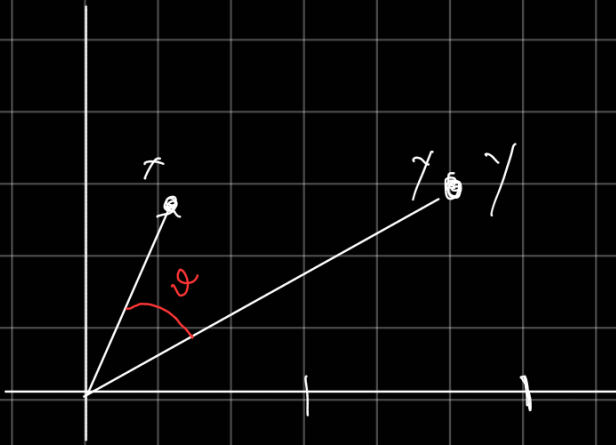
$$\cos \vartheta = \frac{x^T y}{\|x\| \|y\|}$$

$$x^T y = \frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2} = \sqrt{3}$$

$$\|x\| = \sqrt{1 + 3} = 2$$

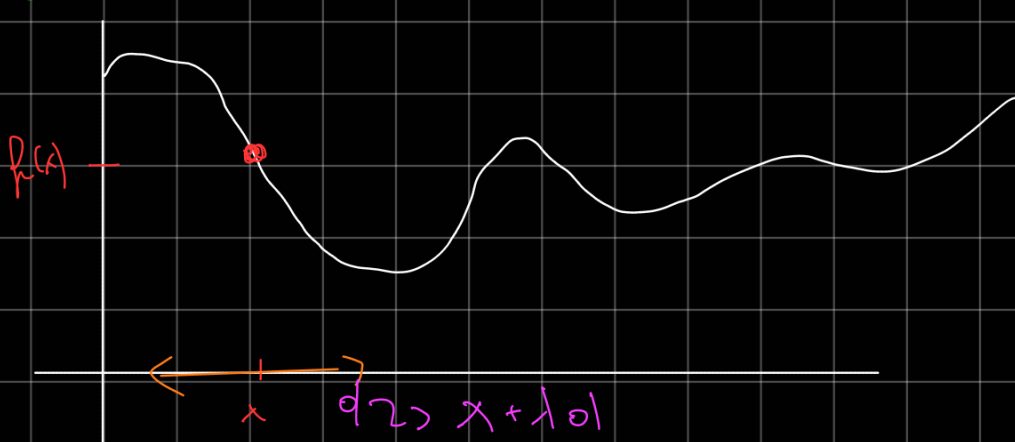
$$\|y\| = \sqrt{3 + 1} = 2$$

$$\cos \gamma = \frac{\sqrt{3}}{2} = \frac{R}{6} = 30^\circ$$



• Direzione DI DISCESA per $f(x)$ in $x \in X$

$\min f(x)$ $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ $X \subseteq \mathbb{R}^n$
 vogliamo vedere dove f decresce



in 2d mi
 porto a dx e sr.
 Se mi sposto
 lungo una direzione
 cosa succede alla
 funzione?

Se mi sposto a dx, $f(x)$ decresce, a sx cresce

Def.

• $\Delta \in \mathbb{R}^n$ è una direzione di discesa di $f(x)$ in $x \in X$ se $\exists \lambda \in \mathbb{R}^+$ tale che

$$f(x + \lambda d) < f(x) \quad \forall 0 < \lambda < 1$$

Come faccio qualitativamente una direzione di discesa? Prodotto sul piano sono

numerumi infinite direzioni.

Se derivata direzionale è negativa lungo una direzione d , allora quella è una direzione di discesa.

Def. Sia $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ con gradiente continuo in $\mathbb{R}^n$. Dati: $x \in \mathbb{R}^n$ e una direzione

$d \in \mathbb{R}^n$, d è una direzione di discesa

Se la derivata direzionale di $f(x)$ in x calcolata rispetto alla direzione d è negativa

Der. sinos.

d direc. di discesa se

$$\lim_{\lambda \rightarrow 0^+} \frac{f(x + \lambda d) - f(x)}{\lambda} = \nabla f(x)^T \cdot d < 0$$

$\lambda > 0$ sempre
 < 0

Se questo rapporto è negativo, significa che il numeratore è negativo, cioè che

Esiste un λ sufficientemente piccolo tale che

$$f(x + \lambda d) - f(x) < 0 \Rightarrow$$

$$\Rightarrow f(x + \lambda d) < f(x)$$

d è discendente. Si discende se

$$\nabla f(x)^T \cdot d < 0$$

angolo tra gradiente
e dir. di
discesa

$$\cos \vartheta = \frac{\nabla f(x)^T \cdot d}{\|\nabla f(x)\| \cdot \|d\|} \Rightarrow$$

norma di un vettore = 0 solo se vettore nullo

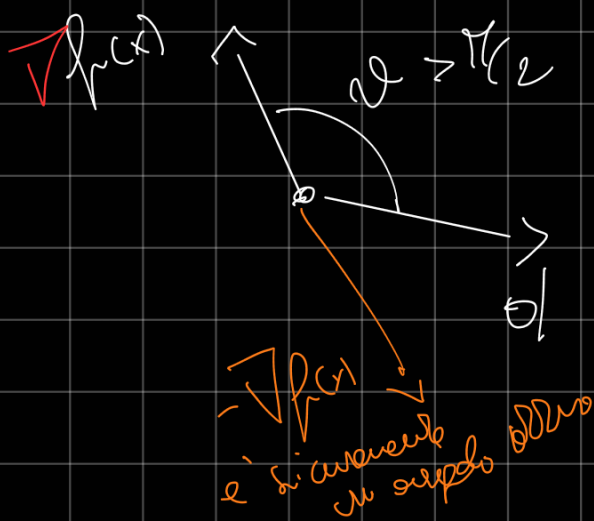
esplicito il prodotto scalare tra due vett.

$$\Rightarrow \nabla f(x)^T d = \|\nabla f(x)\| \cdot \|d\| \cos \vartheta < 0$$

prodotto scalare

negativo per $\vartheta > \pi/2$

• cioè quando ϑ è un angolo ottuso.



Ciò serve a determinare una direzione che forma un angolo ottuso con il gradiente di $f(x)$

- direzione dell'autogradiente

$d = -\nabla f(x)$ direzione dell'autogradiente e

una direzione di discesa per

$f(x)$ in $x \in X$

autogradiente e' sicuramente
una dir. di discesa

Se ^{Spiker} forma invece un angolo acuto
sarà una direzione di salita

• DIREZIONE DI SALITA di $f(x)$ per

$x \in X$

$d \in \mathbb{R}^n$ e' direzione di salita per

$f(x)$ in $x \in X$ se $\exists \bar{\lambda} \in \mathbb{R}^+$ t.c.

$f(x + \lambda d) > f(x)$ per $0 < \lambda < \bar{\lambda}$

• d è una direzione di salita se

$$\nabla f(x)^T \cdot d > 0$$

$$\nabla f(x)^T d = \|\nabla f(x)\| \cdot \|d\| \cos \alpha > 0$$

dove $\cos \alpha > 0$

se $0 < \alpha < \pi/2$



Se $\theta < \pi/2$

d è direzione di salita

- Una direz. ortog. a' la direz. del gradiente

$$d = \nabla f(x) \Rightarrow \theta = 0$$

$\Rightarrow \nabla f(x)$ è una direzione di salita!

Gradiente = direzione e verso lungo cui una funzione cresce più velocemente

- $d \in \mathbb{R}^n$

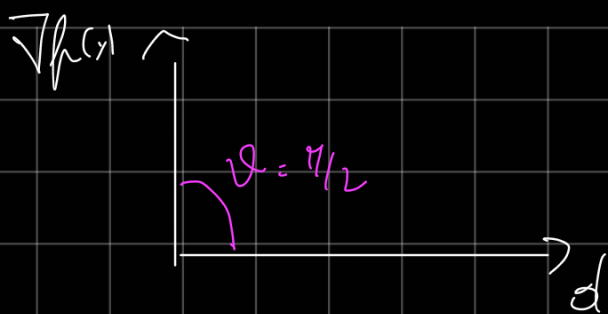
se $\nabla f(x)^T d < 0 \quad (\pi \geq \theta > \pi/2) \Rightarrow d$ è
direz. di
discesa.

Se $\nabla f(x)^T d > 0 \quad (\pi/2 > \theta \geq 0) \Rightarrow d$ è direzione
di salita

Se $\nabla f(x)^T d = 0 \Rightarrow d$ sarà una direzione
ortogonale al gradiente lungo cui la funzione è costante
ne' discende ne' sale.

Supponiamo $\nabla f(x), d \neq 0$, allora

Se i vettori sono ortogonali, f è costante



Siano $x, y \in \mathbb{R}^n$ $x, y \neq 0$

x, y si dicono ortogonali. Se

il loro prodotto scalare è nullo $= 0$

$$x^T y = 0 \Rightarrow \cos \theta = \frac{x^T y}{\|x\| \|y\|} = 0$$

$$\Rightarrow \theta = \pi/2$$

$\|x\| \|y\|$
 $\downarrow \quad \downarrow$
 prodotto scalare

$$x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \quad y = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 2 \end{pmatrix}$$

$$x^T y = 0 \quad \text{non ortogonali}$$

• $x, y \in \mathbb{R}^n$ $x, y \neq 0$

x e y si dicono paralleli. Se

Paralleli: se il loro angolo compreso è 0 o 180°

$$\Rightarrow y = \alpha \cdot x \quad \text{con } \alpha \in \mathbb{R}$$

$$\boxed{\cos \theta} = \frac{x^T y}{\|x\| \|y\|} = \frac{\alpha x^T x}{\|x\| |\alpha| \cdot \|x\|} =$$

$$= \frac{\alpha \|x\|^2}{|\alpha| \|x\|^2}$$

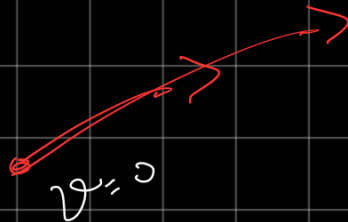
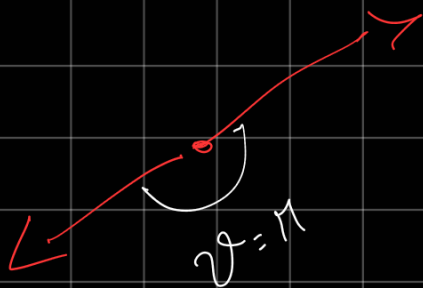
$$= \frac{\alpha}{|\alpha|} = \begin{cases} 1 & \text{se } \alpha > 0 \\ -1 & \text{se } \alpha < 0 \end{cases}$$

$$\cos \theta = 1 \quad \circ \quad \emptyset$$

$\Rightarrow$ Due vettori sono paralleli

se hanno stessa direzione e

verso uguale o opposto (la funzione non varia lungo ϕ)



Esempio

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$f(x) = 2x_1^2 + x_2^2 - 2x_1x_2$$

Considero la Sup. di livello con $\alpha = 2$

$$\mathcal{L}(f(x), \alpha) = \{x \in X : f(x) = \alpha\}$$

$$f(x) = 2x_1^2 + x_2^2 - 2x_1x_2 = 2$$

F quadratica.

è un ellisse

ho 6 punti.

$$A = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$B = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}$$

$$D = \begin{pmatrix} 0 \\ -\sqrt{2} \end{pmatrix}$$

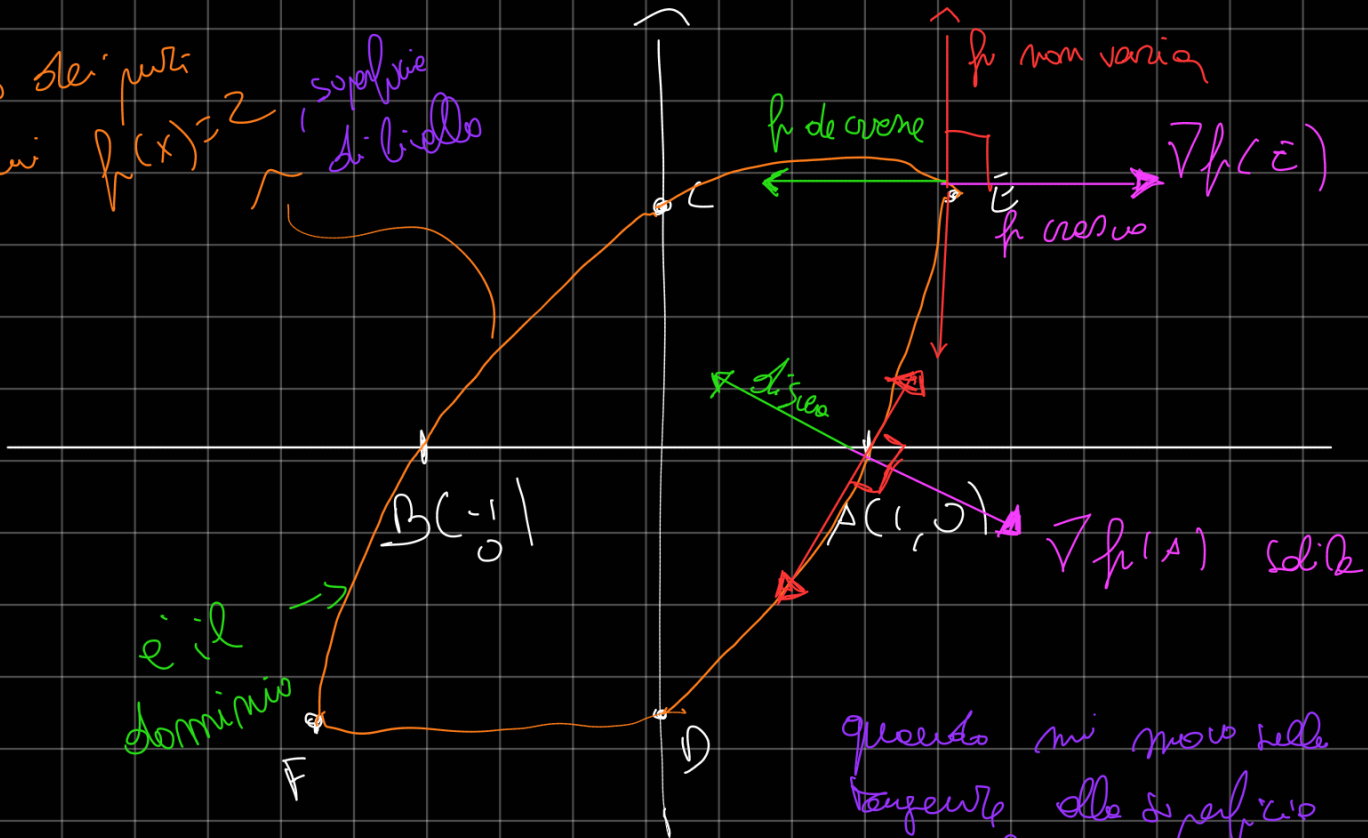
$$E = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \end{pmatrix}$$

$$F = \begin{pmatrix} -\sqrt{2} \\ -\sqrt{2} \end{pmatrix}$$

se mi muovo lungo queste direzioni

lungo dei punti in cui $f(x) = 2$

superficie di livello



quando mi muovo nella tangente alla superficie di livello, se mi muovo

lontano la tangente approssima la curva, la curva di livello.

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

$$= \begin{pmatrix} 4x_1 - 2x_2 \\ 2x_2 - 2x_1 \end{pmatrix}$$

gradiente in A

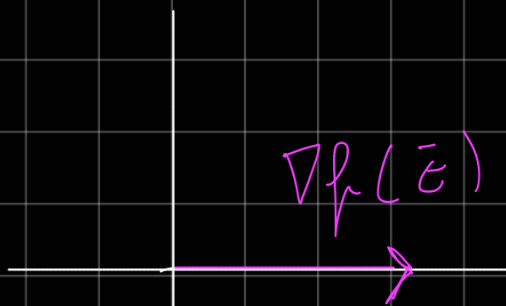
$$\nabla f(A) = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$

$$\nabla f(C) = \begin{pmatrix} -2\sqrt{2} \\ 2\sqrt{2} \end{pmatrix}$$

$$\nabla f(E) = \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$$

$$\nabla f(t) = \begin{pmatrix} 2\sqrt{2} \\ 0 \end{pmatrix}$$

vetture parallele
alle ascisse



Se mi muovo lungo la Tangente, per
distanze piccole, da la tg è uguale alla curva,
quindi velocità non varia

• La tangente alla Superficie σ'

livello in un punto x , definisce
una direzione in cui la $f(x)$

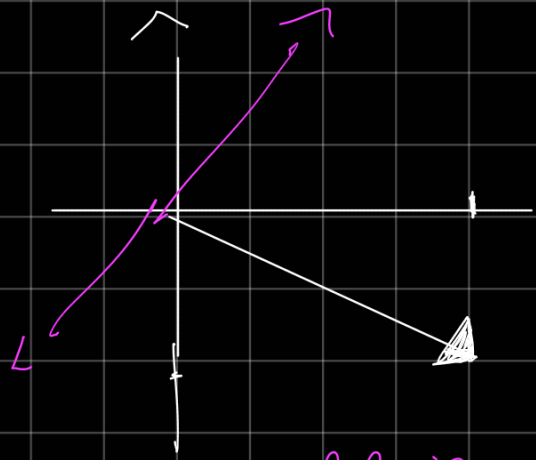
è sempre uguale a α

• La tangente alla Σ_p al livello

in x è ortogonale al gradiente $\nabla f(x)$
calcolato in x . (questo è sempre vero)

(f non varia) (se è ortogonale, $|\nabla f(x)| = 0$)

$$\nabla h(A) = \begin{pmatrix} 4 \\ -2 \end{pmatrix}$$



(la direzione della
tangente, coincide
con la curva nel
punto A, che per def.
è la Superf. di livello)

Calcolo il gradiente,
e mi muovo ortogonale.
ed esso, non mi
muovo

la tg approssima la curva

Problema di ottimizzazione
non vincolato Spazio ammissibile è tutto $\mathbb{R}$

$$X = \mathbb{R}^n$$

$$\bullet \min f(x) \quad x \in \mathbb{R}^n$$

Condizioni di Ottimalità del I° ordine

Sia $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ Una funzione con
con gradiente continuo in $x^* \in \mathbb{R}^n$.

Condizione necessaria affinché x^* sia

Un punto di minimo locale è

$$\nabla f(x^*) = 0$$

~~x^*~~ è solo
un punto stazionario.

D.M.

Punti in cui si annulla il gradiente, sono i punti stazionari.

• H.p. x^* è un minimo locale.

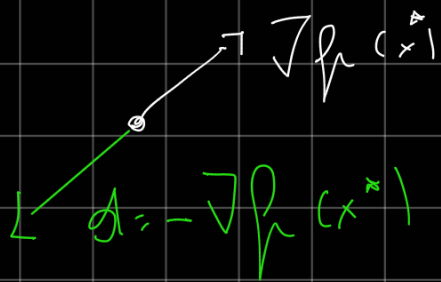
• Th $\nabla f(x^*) = 0$

• Per assurdo supponiamo che

$$\nabla f(x^*) \neq 0$$

Se è $\neq 0$, può essere

una direzione di discesa, perciò direzione anti gradiente



Me. allora $d = -\nabla f(x^*) \neq 0$ è una

direzione di discesa. muovendomi di poco,

la funzione decresce



Se x^* è un punto di minimo locale,

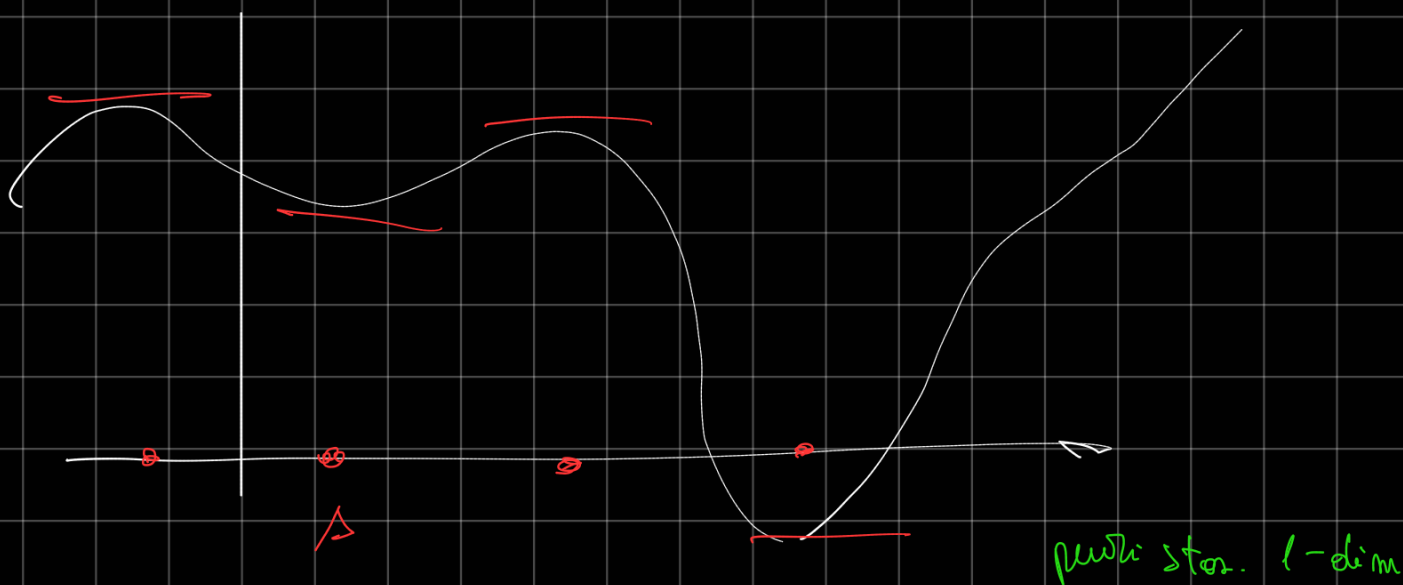
in quell'intorno non può esserci una direzione di discesa

x^* non è un punto di minimo

Assurdo

Def

x^* è un minimo locale se $\exists I(x^*, \varepsilon)$
 tale che $f(x^*) \leq f(x) \forall x \in I(x^*, \varepsilon)$



$$\nabla f(x) \sim \nabla f'(x)$$

$$f'(x) = 0$$

~~alcuni sono minimi, altri massimi. Vediamo
 la derivata seconda per avere più info.~~

Condizione di ottimalità del 1° ordine

• Sia $f_0(x): \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione con
 $\nabla^2 f_0(x)$ continua in un punto $x^* \in \mathbb{R}^n$.

Condizione necessaria affinché x^* sia
minimo locale sia:

$$1) \nabla f(x^*) = 0 \quad x^* \text{ punto stat.}$$

$$2) \nabla^2 f(x^*) \text{ semidefinita positiva}$$

Hessiana semidef. pos. in quel punto.

DIM

• $\forall x^*$ è un minimo locale

• $\forall x^* \nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ semidef. positiva

el. dinos. di spostam. $\in \mathbb{R}^m$
 λ passo $\in \mathbb{R}$

$$x^* \quad x^* + \lambda d \quad d$$

Condiz. 1° ordine $\nabla f(x^*) = 0$

Usa Taylor

$$\nabla = 0$$

$$f(x^* + \lambda d) = f(x^*) + \lambda \nabla f(x^*)^T d + \frac{1}{2} \lambda^2 d^T \nabla^2 f(x^*) d + \beta_2(x^*, \lambda, d)$$

$$f(x^* + \lambda d) - f(x^*) = \frac{1}{2} \lambda^2 d^T \nabla^2 f(x^*) d + \beta_2(x^*, \lambda, d)$$

$\nabla f(x^*) = 0$
 perché $f(x^*)$ min. loc.

t.c. $\lim_{\lambda \rightarrow 0} \frac{\beta_2(x^*, \lambda, d)}{\lambda^2} = 0$

Per λ suff. piccolo, per ipotesi

$$f(x^* + \lambda d) - f(x^*) \geq 0$$

Divido per λ^2

> minimo

$$\frac{f(x^* + \lambda d) - f(x^*)}{\lambda^2} = \frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{\beta_2}{\lambda^2}$$

Faccio il limite $\frac{\beta_2}{\lambda^2} \rightarrow 0$

$$\lim_{\lambda \rightarrow 0} \frac{f(x^* + \lambda d) - f(x^*)}{\lambda^2} = \lim_{\lambda \rightarrow 0} \frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{\beta_2}{\lambda^2}$$

$$\Rightarrow \frac{1}{2} d^T \nabla^2 f(x^*) d \geq 0$$

$\nabla^2 f$ semidef. positiva da ipotesi

$$\forall d \in \mathbb{R}^n$$

non
lo fa

non ipotesi su d
vale su qualsiasi
vettore lineare d

è la defn. di Matrice semidef.
positiva.

$\Rightarrow \nabla^2 f(x^*)$ é semidefinida positiva $\square$

• Condições suficientes DI MINIMO Loc.

Seja $f(x)$ uma função $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$

com $\nabla^2 f(x)$ contínua em um

interior de um ponto $x^* \in \mathbb{R}^n$

Allora x^* é um mínimo local

SUFICIENTE se

$$1) \nabla f(x^*) = 0 \quad 2) \nabla^2 f(x^*) \text{ definida positiva}$$
$$f(x^*) < f(x) \quad x \in I(x^*, \varepsilon)$$

DIH

Hp. $\nabla f(x^*) = 0$; $\nabla^2 f(x^*)$ def. pos.
 ∇f x^* é um mínimo local forte

Sempre Taylor

Se da $f(x^* + \delta) > f(x^*)$

$\Rightarrow$ $f(x^*)$ è un minimo locale
stretto.

Lo stesso ragionamento può essere fatto con i massimi

$$\max_{x \in \mathbb{R}^n} f(x)$$

OTTIMIZZAZIONE
VINCOLATA

Condizione necessaria locale di massimo:

del I° ordine

$$\nabla f(x^*) = 0$$

^{dim}
Se $\nabla f(x^*) \neq 0$
 $d = \nabla f(x^*)$ è lineare.
di solito, x^* non
è un massimo locale

Condizioni necessarie del II° ordine di
max locale.

$$d) \nabla^2 f(x^*) \succeq 0$$

2) $\nabla^2 f(x^*)$ semidefinita negativa $\left(\begin{array}{l} -\nabla^2 f(x^*) \\ \text{semidefinita} \\ \text{POSITIVA} \end{array} \right)$

Condizioni sufficienti di ottim. locale della

1) $\nabla f(x^*) = 0$

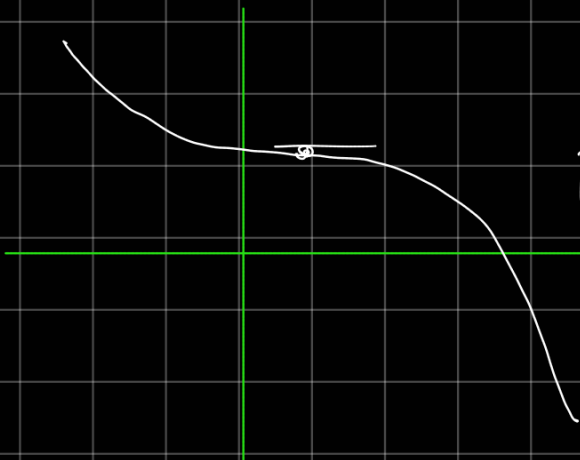
2) $\nabla^2 f(x^*)$ definita negativa
($-\nabla^2 f(x^*)$ è definita positiva)

Supponiamo che $\nabla^2 f(x^*)$ non sia né

(semi)definita positiva, né negativa

e che $\nabla f(x^*) = 0$

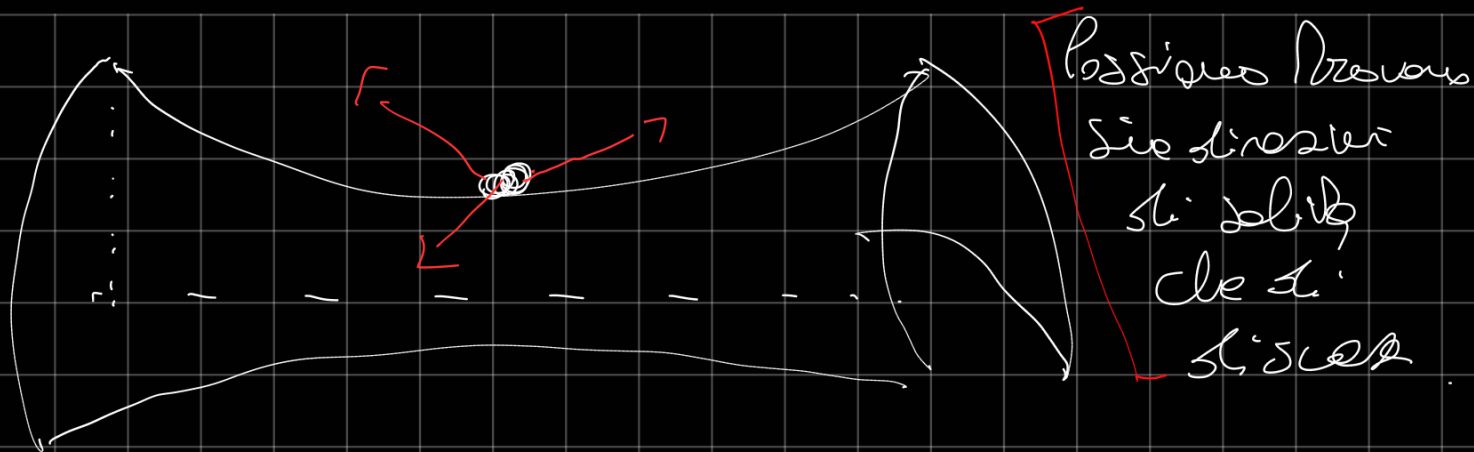
punto che annulla il
gradiente
non è né minimo né massimo



$f'(x) = 0$

ma il pb non è
né min né max loc.

x^* è un punto di SELLA



Es.

Min $f(x) = x_1^2 + x_2^2 + \underbrace{x_1 \cdot x_2^2}_{\text{Terminale}}$

$$f(x): \mathbb{R}^2 \rightarrow \mathbb{R} \quad x \in \mathbb{R}^n$$

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2^2 \\ 2x_2 + 2x_1 x_2 \end{pmatrix}$$

$$x_1 = -\frac{x_2^2}{2}$$

$$2x_2 - x_2^3 = 0$$

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

$$x_2(2 - x_2^2) = 0$$

principio di
Sylvester

$$= \begin{pmatrix} 2 & 2x_2 \\ 2x_2 & 2+2x_1 \end{pmatrix}$$

• Caratteristi a minimo lo calcolano
severi soluzioni

$$\begin{cases} 1) \nabla f(x^*) = 0 \\ 2) \nabla^2 f(x^*) \text{ semi-def. pos.} \end{cases}$$

$$1) \nabla f(x^*) = 0 \quad \begin{cases} 2x_1 + x_2^2 = 0 \\ 2x_2 + 2x_1x_2 = 0 \Rightarrow \end{cases}$$

Condizioni α e β

$$\alpha \begin{cases} 2x_1 + x_2^2 = 0 \\ x_2(2 + 2x_1) = 0 \Rightarrow x_2 = 0 \quad x_1 = -1 \end{cases}$$

$$\alpha \quad x_2 = 0 \Rightarrow x_1 = 0 \Rightarrow A = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\alpha \quad x_1 = -1 \Rightarrow x_2^2 = 2 \Rightarrow x_2 = \pm\sqrt{2}$$

$$= B = \begin{pmatrix} -1 \\ \sqrt{2} \end{pmatrix} \quad C = \begin{pmatrix} -1 \\ -\sqrt{2} \end{pmatrix}$$

Punti che soddisfanno il
problema

Punto A , vediamo la matrice Hess. calcolata nel punto A

$$\nabla^2 h(A) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \det(A) = 4 \quad \lambda_{\min} = 2$$

$\nabla^2 h(A)$ è definita positiva.

Soddisfa le condiz. Suff.

Se fosse stata
Semi-def. pos. /
Soddisfa le condiz.
necessarie, ma non
Suff. - non
potrebbe essere nulla
o zero.

A è un MINIMO LOCALE STRETTO.

$$B = \begin{pmatrix} -1 \\ \sqrt{2} \end{pmatrix}$$

$$\nabla^2 h(B) = \begin{pmatrix} 2 & 2\sqrt{2} \\ 2\sqrt{2} & 0 \end{pmatrix} = \det(B) = -8$$

$\nabla^2 h(B)$ non è Semi-def. pos. /
vedo se è un max

$$\Rightarrow -\nabla^2 h(B) = \begin{pmatrix} -2 & -2\sqrt{2} \\ -2\sqrt{2} & 0 \end{pmatrix} \quad \lambda_{\min} < 0 \quad \det < 0 \quad -8$$

non è Semi-def. pos.

$\nabla^2 h(B)$ non è né semi-def. pos. e neg. /
 B è un punto di sella.

punto C $\nabla^2 h(C) = \begin{pmatrix} 2 & -2\sqrt{2} \\ -2\sqrt{2} & 0 \end{pmatrix}$ ha
det
negativo

$-\nabla^2 h(C) = \begin{pmatrix} -2 & 2\sqrt{2} \\ 2\sqrt{2} & 0 \end{pmatrix}$ non è né definita
positiva né negativa

C è un p.u.b. di sella

• $f(x)$ ammette punti di minimo
globale? No.

$$f(x) = x_1^2 + x_2^2 + x_1 - x_2^2$$

esistono direz. per cui la f .
va a $-\infty$

es. . prendo direzione

$$d = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -2 \\ x_2 \end{pmatrix}$$

$$h(d) = 4 + x_2^2 - 2x_2^2 = 4 - x_2^2$$

Se x_2 va a $\pm\infty$, la f va a $-\infty$

$$\lim_{x_2 \rightarrow \pm\infty} 4 - x_2^2 = -\infty$$

Non ha
minimo locale
globale (!)

