

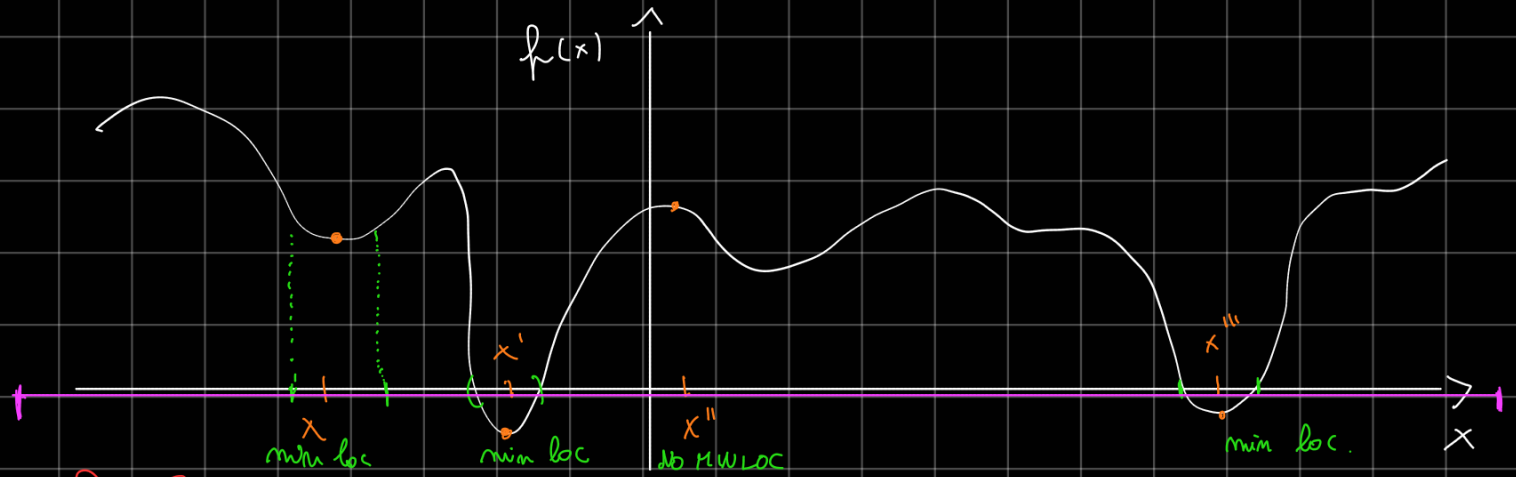
Lez. 2

$$\min f(x)$$

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

$$x \in X$$

f generica



Def PUNTI DI MINIMO DI UNA FUNZIONE

- x^* si dice punto di minimo globale

se $\forall x \in X$ si ha $f(x^*) \leq f(x)$

se $f(x^*) < f(x)$ x^* è un

MINIMO GLOBALE STRETTO

- $x^* \in X$ si dice PUNTO DI MINIMO LOCALE

se $\exists I(x^*, \varepsilon)$ tale che $f(x^*) \leq f(x)$

$\forall x \in I(x^*, \varepsilon)$

ovunque un min globale è un minimo locale
ma non il viceversa.

Se $f(x^*) < f(x)$ allora x^* è un
minimo locale stretto.

ci concentriamo nello studio dei minimi locali
e globali

Derivata di una funzione

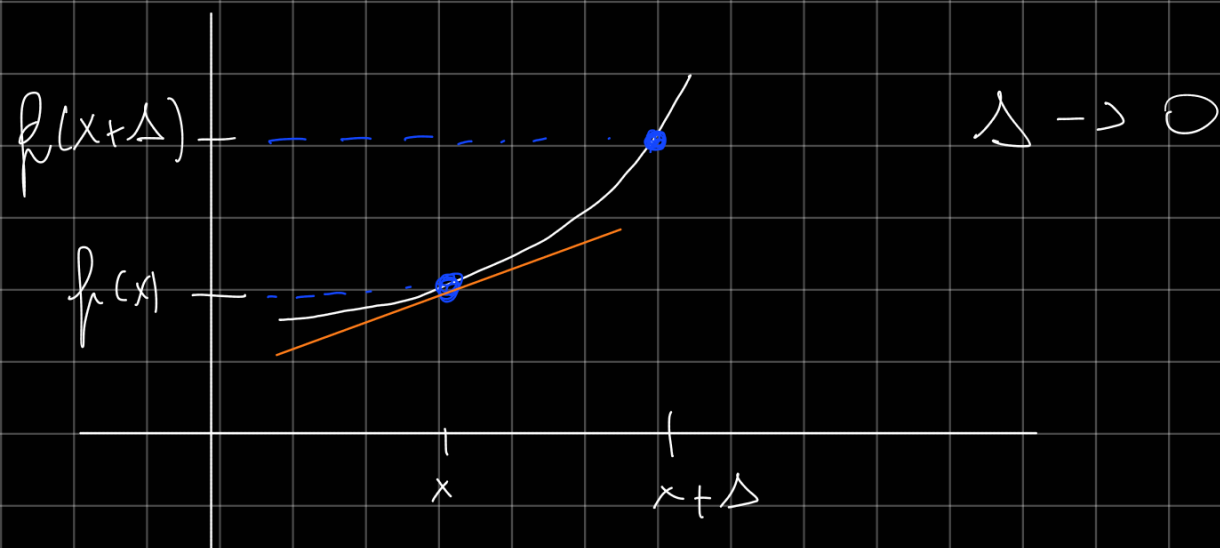
• $f: \mathbb{R} \rightarrow \mathbb{R}$

derivata della $f(x)$ in un punto x è:

$$f'(x) = \lim_{\Delta \rightarrow 0} \frac{f(x+\Delta) - f(x)}{\Delta}$$

limite rapporto
incrementale

variazione della funzione x quando mi sposto
in modo infinitesimo lungo x , mi dice come
varia la funzione



$$f = mx + q$$

$$\underline{m = f'(x)}$$

derivata = inclinazione della retta tangente, coefficiente

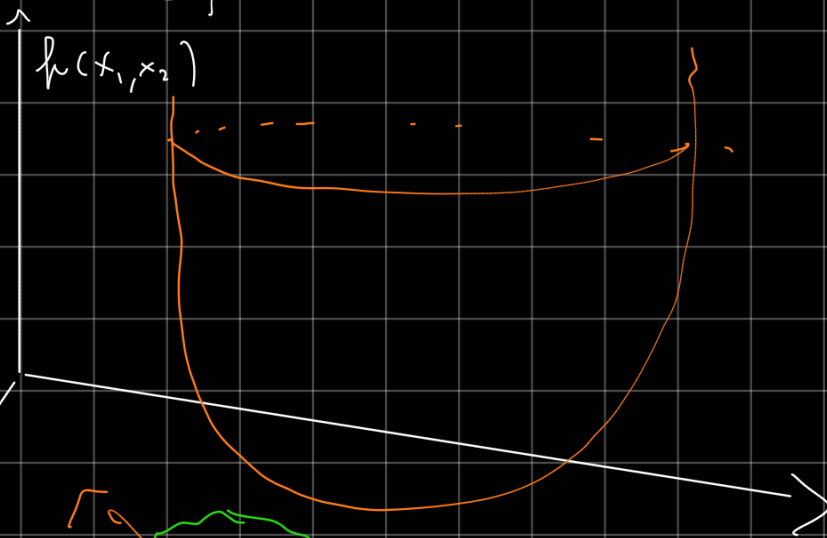
Se ho una funzione in più variabili?

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

in più variabili lo spostamento lungo x
non è solo in avanti o dietro...

Es. $f(x_1, x_2): \mathbb{R}^2 \rightarrow \mathbb{R}$

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$$



ora mi posso
spostare secondo
diverse direzioni

scelgo di
spostarmi lungo
questa direzione

ho infinite direzioni
lungo cui posso spostarmi

• Derivata direzionale di $f(x): \mathbb{R}^m \rightarrow \mathbb{R}$

$x \rightsquigarrow x + \lambda d$
minimo lungo la direzione $d \in \mathbb{R}^m$

$d \in \mathbb{R}^m$ direzione, vettore
 $\lambda \in \mathbb{R}$ spostamento, scalare
 $x \in \mathbb{R}^m$

La derivata direzionale di

$f(x): \mathbb{R}^m \rightarrow \mathbb{R}$ calcolata nel punto

Secondo la definizione di ∇ è:

$$\lim_{\lambda \rightarrow 0^+} \frac{f(\vec{x} + \lambda \vec{d}) - f(\vec{x})}{\lambda}$$

λ è positivo
e tende a 0

C'è una derivata direzionale particolare

$$d = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

è

componente i -esima di $\vec{d}$

Se muoviamo lungo un vettore

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}$$

$\Rightarrow$ la derivata direzionale di $f(x)$
muoviamo lungo la i -esima componente

secondo d è $\frac{\partial f(x)}{\partial x_i}$

derivata parziale
di $f(x)$ rispetto
ad x_i

Derivata parziale = caso particolare di derivata direzionale

GRADIENTE di $f(x)$ calcolato
in x

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

Vettore che ha
come componenti
le derivate
parziali rispetto
ogni componente
cioè derivata direzionale

Derivata seconda

1-dim

• $f(x): \mathbb{R} \rightarrow \mathbb{R}$

$f''(x)$ Derivata seconda

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

$f \in C^2$
e $f(x)$ 2 volte
continuamente differenziabile

x_i e x_j prendo una coppia

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

= derivata parziale seconda,
derivata prima rispetto ad
 x_i e poi rispetto a x_j

ho n^2 possibili di "combinazioni"

Teo di SCHWARTZ

posso invertire ordine di deriv.

le derivate seconde sono simmetriche rispetto
alle due variabili.

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

$$= \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$$

Se $x_j = x_i$

$$\Rightarrow \frac{\partial^2 f(x)}{\partial x_i^2}$$

Introduco un oggetto che ha
tutte le derivate seconde

MATRICE HESSIANA DI $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$

$$\nabla^2 f(x) \in \mathbb{R}^{n \times n}$$

Matrice ovale per contenere
tutte le derivate seconde

$\nabla^2 f(x) =$

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

Sviluppo di Taylor multidim

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

$$x \in \mathbb{R}$$

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

$$\begin{array}{c} | \quad | \quad \rightarrow \\ x \quad x+h \end{array}$$

qual è il valore della funzione x quando mi avvicino ad $x+h$? lo stimo con il polinomio di Taylor 1 dim

$$f(x+h) = f(x) + f'(x)h + \beta_1(x, h)$$

$O(x, h)$

β_1 è un infinitesimo di ordine superiore

$$\lim_{h \rightarrow 0} \frac{\beta_1(x, h)}{h} = 0$$

β_1 tende a 0 molto più velocemente rispetto ad h

Formula di Taylor del primo ordine

2° ordine

$$f(x+h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2} + \beta_2(x, h)$$

$O(x, h^2)$ Taylor "arrestato" al 2° ordine

$$\lim_{h \rightarrow 0} \frac{\beta_2(x, h)}{h^2} = 0$$

Taylor caso multidimensionale

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

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$x \rightsquigarrow x+h$$

$$h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}$$

Taylor "arrestato" al 1° ordine

$$f(x+h) = f(x) + \nabla f(x)^T h + \beta_1(x, h)$$

β_1 infinitesimo di ordine superiore rispetto alla norma

$$\|h\| = \sqrt{h_1^2 + h_2^2 + \dots + h_n^2} \quad \leftarrow \text{NORMA DI UN VETTORE}$$

$$\lim_{\|h\| \rightarrow 0} \frac{\beta_1(x, h)}{\|h\|} = 0$$

$$\boxed{\nabla f(x)^T \cdot h = \begin{pmatrix} \frac{\partial f(x)}{\partial h(x_1)} & \frac{\partial f(x)}{\partial h(x_2)} & \dots & \frac{\partial f(x)}{\partial h(x_n)} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}}$$

Taylor medesima 1° ordine

$$f(x+h) = f(x) + \nabla f(x)^T h + \frac{h^T \nabla^2 f(x) h}{2} + \beta_2(x, h)$$

Matrice Hesseiana
e simmetrica

love $\lim_{h \rightarrow 0} \frac{\beta_2(x, h)}{\|h\|^2}$

$$\begin{pmatrix} h_1 & h_2 & \dots & h_n \end{pmatrix} \begin{pmatrix} \nabla^2 f(x) \\ \underline{n \times n} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}$$

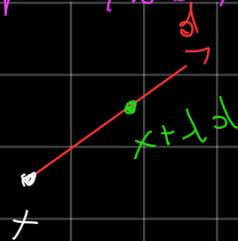
• Derivata Direzionale di $f(x)$

in x rispetto alla direzione d

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

Come veder la f se siamo da x lungo questa direzione?

$$\lim_{\lambda \rightarrow 0^+} \frac{f(x + \lambda d) - f(x)}{\lambda}$$



Come farlo in pratica a calcolo?

uso Taylor ---

$$x + h = x + \lambda d$$

$$h = \lambda d$$

$$f(x + \lambda d) = f(x) + \lambda \nabla^T f(x) d + \beta_1(x, \lambda, d)$$

$$f(x + \lambda d) - f(x) = \lambda \nabla^T f(x) d + \beta_1(x, \lambda, d)$$

divisto tutto per λ
e ora passo al limite

$$\lim_{\lambda \rightarrow 0^+} \frac{f(x + \lambda d) - f(x)}{\lambda} = \lim_{\lambda \rightarrow 0^+} \left(\nabla^T f(x) d + \frac{\beta_1(x, \lambda, d)}{\lambda} \right)$$

Def. di derivata
lineare

non scappa
da λ

ma β è
un infinitesimo
di ordine
superiore!
si annulla..

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

si $f(x)$

Definizione: Derivata direzionale = prodotto
scalare tra gradiente e direzione

Combinazione lineare di punti

$$x^{(1)}, x^{(2)}, \dots, x^{(k)} \in \mathbb{R}^n \quad x^{(i)} = \begin{pmatrix} x_1^{(i)} \\ x_2^{(i)} \\ \vdots \\ x_n^{(i)} \end{pmatrix}$$

$$\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$$

$$\lambda_1 x^{(1)} + \lambda_2 x^{(2)} + \dots + \lambda_k x^{(k)} = \sum_{i=1}^k \lambda_i x^{(i)}$$

vettore (i)esimo

è chiamata **COMBINAZIONE LINEARE**

1) se $\lambda_i \geq 0 \quad i=1, \dots, k \Rightarrow \sum_{i=1}^k \lambda_i x^{(i)}$ **COMBINAZIONE CONCA**

2) se $\sum_{i=1}^k \lambda_i = 1 \Rightarrow \sum_{i=1}^k \lambda_i x^{(i)}$ **COMBINAZIONE AFFINE**

3) $\sum_{i=1}^k \lambda_i = 1 \quad \lambda_i \geq 0 \quad i=1, \dots, k \Rightarrow \sum_{i=1}^k \lambda_i x^{(i)}$ **COMBINAZIONE CONVESSA**

Es. $x, y \in \mathbb{R}^2 \Rightarrow \lambda_1 x + \lambda_2 y$
 $\lambda_1, \lambda_2 \in \mathbb{R}$



• **COMBINAZ. LINEARE**, se λ_1, λ_2 prendo

tutti i valori possibili, la comb. lin mi descrive
tutto IL PIANO ($\mathbb{R}^2$)

• **COMBINAZ. CONICA** Attenzione tutti i punti nel
cono generati da x e y
 $\lambda_1, \lambda_2 \geq 0$

COMBINAZ. AFFINE Attenzione nona parte per
 x e y
 $\lambda_1 + \lambda_2 = 1$

COMBINAZ. CONVESSA $\Rightarrow$ Segmento che
unisce x e y
 $\lambda_1, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1$

$\rightarrow$ posso anche scriverla come

$$\hookrightarrow x, y \in \mathbb{R}^n \quad \lambda_1, \lambda_2 \in \mathbb{R} \quad \lambda_1, \lambda_2 \geq 0 \quad \text{e} \quad \lambda_1 + \lambda_2 = 1$$

$$\Rightarrow \lambda_2 = 1 - \lambda_1$$

COMB. LINEARE

$$\lambda \in \mathbb{R}, \lambda \in [0, 1] \Rightarrow \lambda x + (1 - \lambda)y$$

al variare di λ mi sposto dal punto
 y al punto x



se $\lambda \rightarrow 0$, $\lambda x + (1-\lambda)y \rightarrow y$

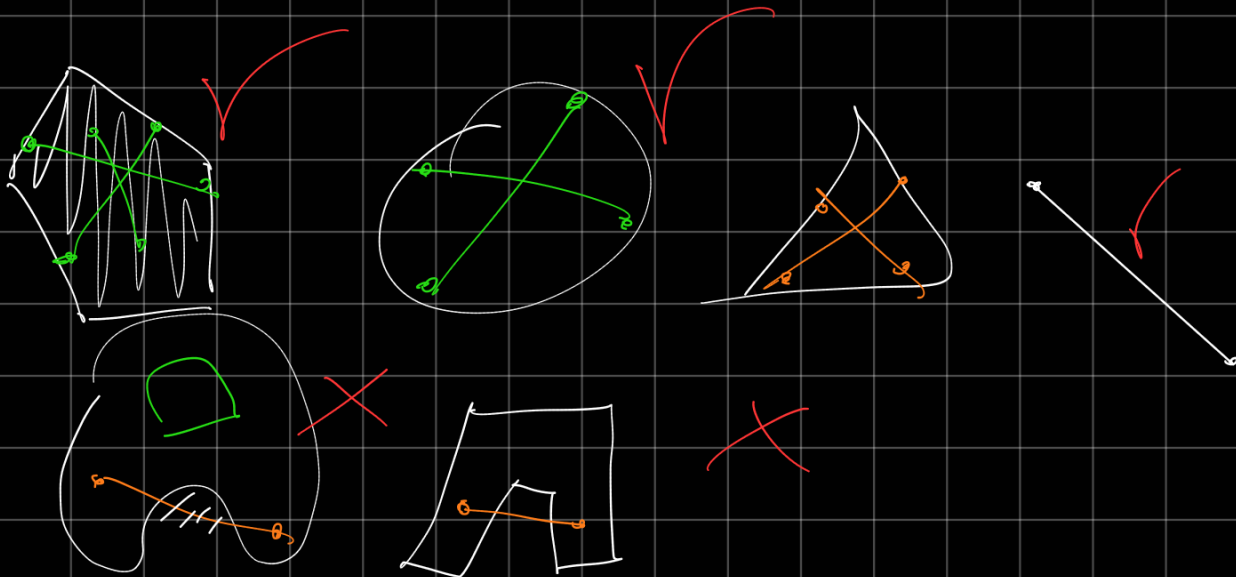
se $\lambda \rightarrow 1$, $\lambda x + (1-\lambda)y \rightarrow x$

INSIEME CONVESSO

$X \subseteq \mathbb{R}^n$ è un INSIEME CONVESSO se

$\forall x, y \in X$ il segmento che unisce x e y è tutto contenuto in X

es.



Def. FUNZIONE CONVESSA

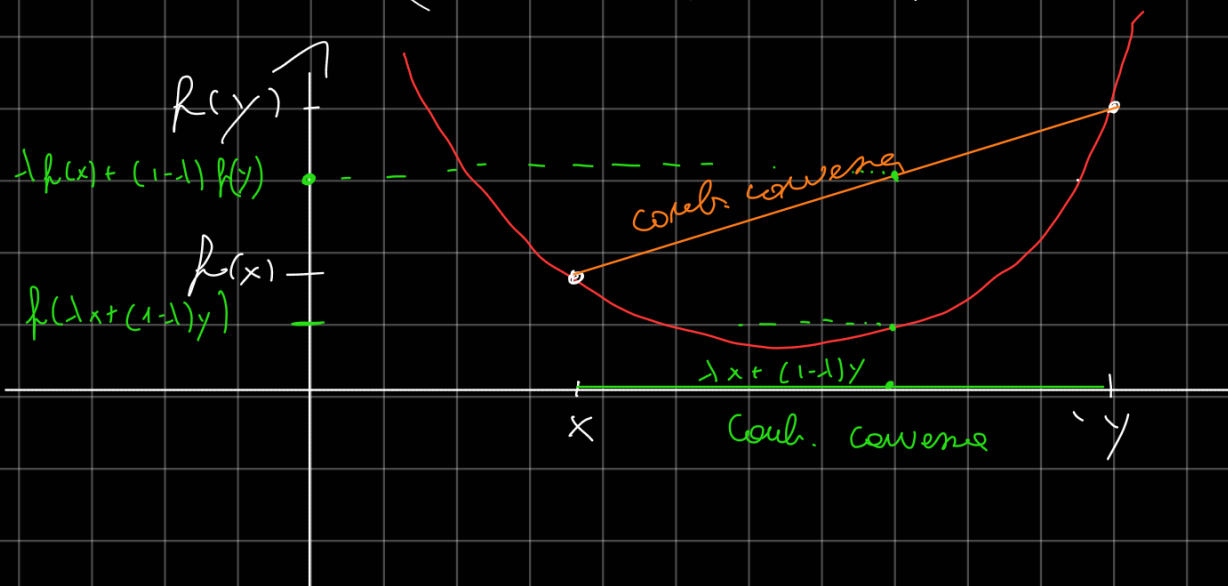
$f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ $x \in \mathbb{R}^n$,

X insieme convesso

$f(x)$ è una funzione convessa in X se

$\forall x, y \in X$, e $\forall \lambda \in [0, 1]$, si ha

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y)$$



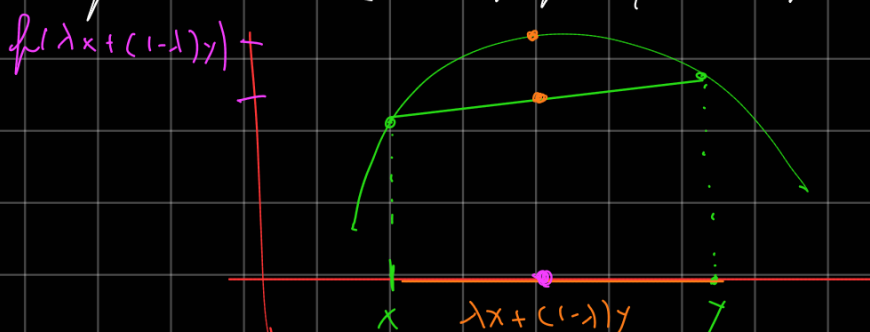
" Se la f è convessa, qualsiasi due punti sulla curva, il segmento è sempre AL DI SOPRA della curva

- Se $f(x)$ è convessa $\Rightarrow$
 $\Rightarrow -f(x)$ su' su' è **CONCAVA**

• $f(x)$ è una funzione concava in X

se $\forall x, y \in X$ e $\lambda \in [0, 1]$ si ha

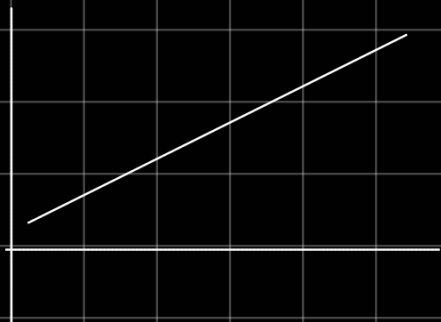
$$\lambda f(x) + (1-\lambda)f(y) \leq f(\lambda x + (1-\lambda)y)$$



segmento che collega due punti dalla curva è sotto la curva

Se $f(x)$ è concava $\Rightarrow -f(x)$ è convessa

Una $f(x)$ lineare è sia concava
che convessa



Se f è differenziabile, $F: \Omega \rightarrow \mathbb{R}^m$
in x se $F(x+h) = F(x) + J_x \cdot h + o(\|h\|)$

Sia $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x)$ convessa
e continuamente differenziabile
in $\mathbb{R}^n$, allora

$\Rightarrow \forall x, y \in \mathbb{R}^n$ si ha che

vale anche
il viceversa

$$f(y) - f(x) \geq \nabla f(x)^T (y - x)$$

Cioè se vale questo,
da f è convessa

D.M

$f(x)$ è convessa e cont. diff.

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

$$\lambda f(x) + (1-\lambda)f(y) \geq f(\lambda x + (1-\lambda)y)$$

$$\varepsilon = 1-\lambda \Rightarrow \lambda = 1-\varepsilon$$

$$\Rightarrow (1-\varepsilon)f(x) + \varepsilon f(y) \geq f((1-\varepsilon)x + \varepsilon y)$$

$$\Rightarrow (1-\varepsilon)x + \varepsilon y = \overset{x - \varepsilon x + \varepsilon y}{x} + \underbrace{\varepsilon(y-x)}_{\Delta d} \quad \times \text{ scegliendo di un certo valore}$$

$$\Rightarrow x + \alpha d \quad \alpha = \varepsilon \quad y - x = d$$

allora riscriviamo la formula di Taylor così

$$\rightarrow f(x + \varepsilon(y-x)) = f(x) + \varepsilon \nabla f(x)^T (y-x) + \beta_1(\varepsilon, x, y)$$

$$\rightarrow \underline{(1-\varepsilon)f(x) + \varepsilon f(y)} \geq \underbrace{f(x) + \varepsilon \nabla f(x)^T (y-x)}_{= f(\lambda x + (1-\lambda)y)} + \beta_1(\varepsilon, x, y)$$

$$\cancel{f(x) + \varepsilon(f(y) - f(x))} \geq \cancel{f(x) + \varepsilon \nabla f(x)^T (y-x)} + \frac{\beta_1}{\varepsilon} \quad \text{divido per } \varepsilon$$

$$\bullet f(y) - f(x) \geq \nabla f(x)^T (y-x) + \frac{\beta_1}{\varepsilon}$$

Passo al limite

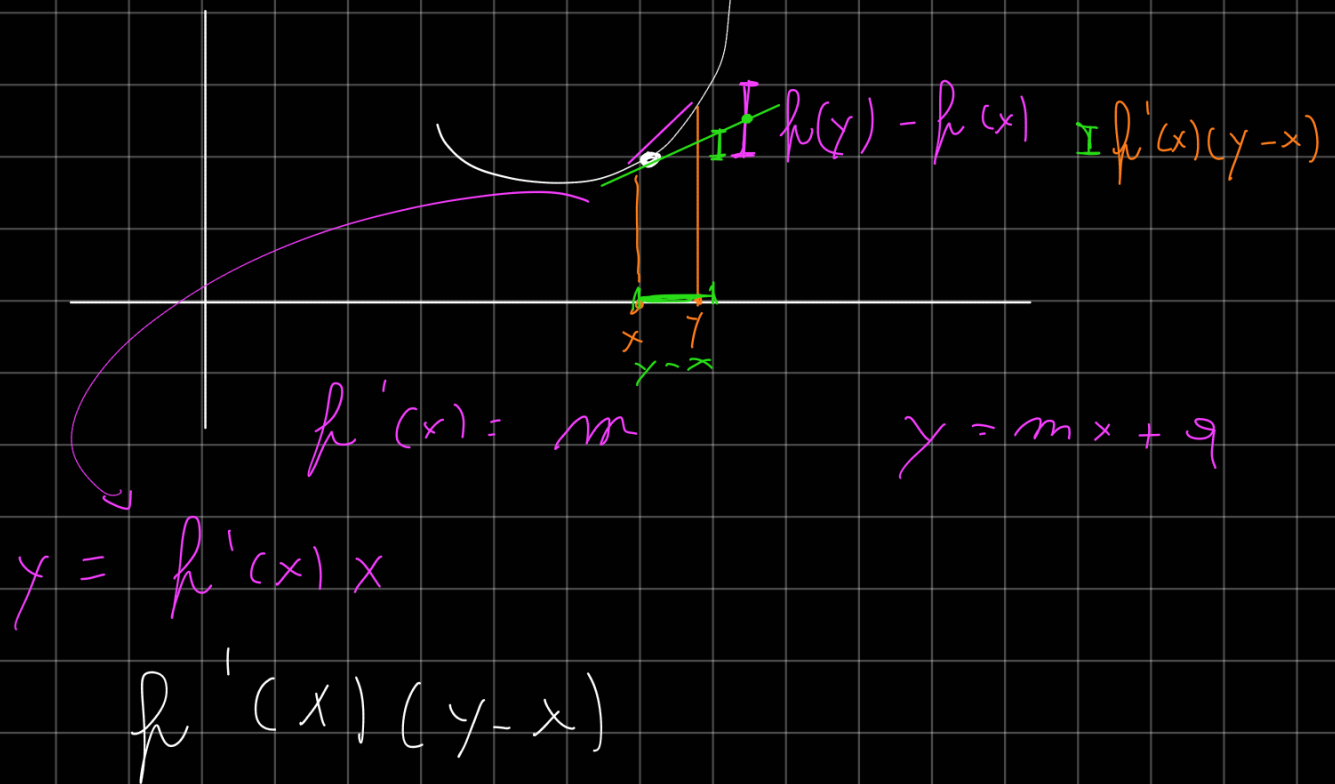
$$\lim_{\varepsilon \rightarrow 0} \frac{\beta_1}{\varepsilon} \text{ es. nulla}$$

Tesi dimostrata



in una dimensione
 $f(x): \mathbb{R} \rightarrow \mathbb{R}$ $f(x)$ convessa e diff.

$$\forall x, y \in \mathbb{R} \Rightarrow f(y) - f(x) \geq f'(x)(y - x)$$



Se f è convessa, la retta tangente in x è sempre al di sotto della curva (la derivata è sempre sotto il grafico)