$f : \mathbb{R}^n \to \mathbb{R}$, $\quad D \subseteq \mathbb{R}^n$

"Trovare $\bar{x} \in \mathbb{R}^n$ tale che
- $\bar{x} \in D$ (ammissibilità)
- $f(\bar{x}) \le f(x) \quad \forall x \in D$ (ottimalità)"

$\ge$

(P) $\min \{ f(x) : x \in D \}$

max

## Classificazione

$D = \mathbb{R}^n$ ottimizzazione (programmazione) non vincolata

$D \subsetneq \mathbb{R}^n$ ottimizzazione vincolata $\longrightarrow$ D finito/numerabile $\to$ altrimenti "ott. continua"

$\downarrow$

ottimizzazione discreta $\begin{cases} D \subseteq \{0,1\}^n & \text{combinatoria} \\ D \subseteq \mathbb{Z}^n & \text{intera} \end{cases}$

fz. obiettivo $\begin{cases} f & \text{lineare} \\ f & \text{nonlineare} \end{cases}$

MNO: ottimizzazione non lineare non vincolata e vincolata (caso continuo)

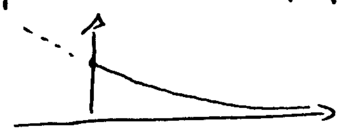
## Definizioni di ottimalità

**Def** $\bar{x} \in \mathbb{R}$ si dice [PUNTO DI] MINIMO GLOBALE (max) di (P) se
- $\bar{x} \in D$
- $f(\bar{x}) \le f(x) \quad \forall x \in D$ $\quad (\ge)$

In tal caso $f(\bar{x})$ si dice valore ottimo di (P)

Valore ottimo esiste ma non esiste alcun punto di minimo : $f(x) = e^{-x}$ $\quad D = \mathbb{R}_+$ $(D = \mathbb{R})$

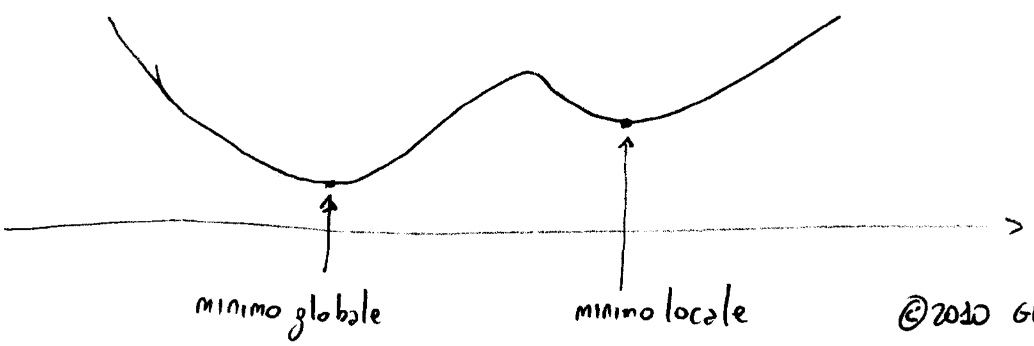
(inteso come $\inf \{ f(x) : x \in D \}$) $\quad \begin{cases} \bar{\ell} \le f(x) \quad \forall x \in D \\ \ell \le f(x) \ \forall x \in D \Rightarrow \ell \le \bar{\ell} \end{cases}$



Valore ottimo non esiste : $f(x) = -x^2$, $D = \mathbb{R} \to f(x) \longrightarrow -\infty$ per $x \to \pm\infty$

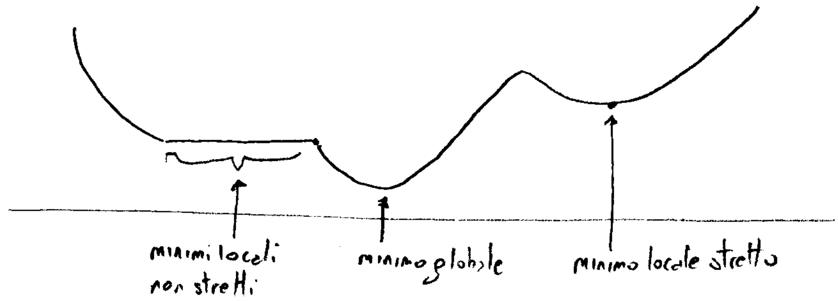
Weierstrass : f continua, D compatto $\Rightarrow$ minimo [e massimo] esistono

**Def** $\bar{x} \in \mathbb{R}^n$ si dice [PUNTO DI] MINIMO LOCALE di (P) se
- $\bar{x} \in D$
- $\exists\, \varepsilon > 0 : f(\bar{x}) \le f(x) \quad \forall x \in D \cap B(\bar{x}, \varepsilon)$



**Def** $\bar{x} \in \mathbb{R}^n$ si dice [PUNTO DI] MINIMO LOCALE STRETTO di (P) se

- $\bar{x} \in D$
- $\exists \varepsilon > 0 \; : \; f(\bar{x}) < f(x) \quad \forall x \in D \cap B(\bar{x}, \varepsilon),\ x \neq \bar{x}$.

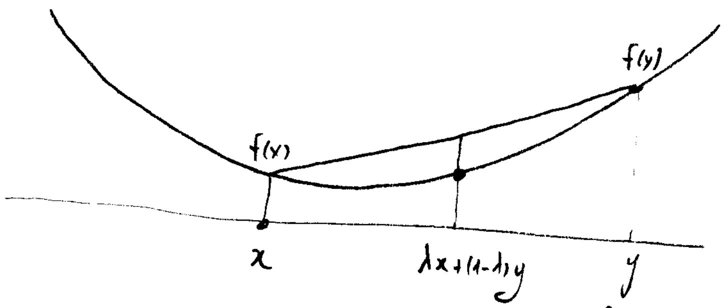



Ulteriore distinzione: ottimizzazione globale vs ottimizzazione locale

**Def**: $f : \mathbb{R}^n \to \mathbb{R}$ si dice CONVESSA se  (f concava se $(-f)$ convessa)

$$f(\underbrace{\lambda x + (1-\lambda) y}_{\text{segmento } [x,y] \subseteq \mathbb{R}^n}) \leq \underbrace{\lambda f(x) + (1-\lambda) f(y)}_{\text{segmento } [f(x), f(y)] \subseteq \mathbb{R}}, \quad \forall x, y \in \mathbb{R}^n \ \forall \lambda \in [0,1] \qquad (*)$$

f si dice strettamente convessa se (*) vale con $<$ al posto di $\leq$




**Oss** f è convessa

$\Updownarrow$

$$f\left(\sum_{i=1}^{K} \lambda_i x_i\right) \leq \sum_{i=1}^{K} \lambda_i f(x_i)$$

$\forall x_1, \ldots, x_K \in \mathbb{R}^n,\ \forall \lambda_i \geq 0$ con $\sum_{i=1}^{K} \lambda_i = 1$

**Teorema** Sia $f : \mathbb{R}^n \to \mathbb{R}$ convessa, $D \subseteq \mathbb{R}^n$ convesso. Allora ogni punto di minimo locale di (P) è un punto di minimo globale (LOCALE ≡ GLOBALE)

(vedi pg ⑥)

**dim** Sia $\bar{x} \in D$ minimo locale, supponiamo che non sia minimo globale:

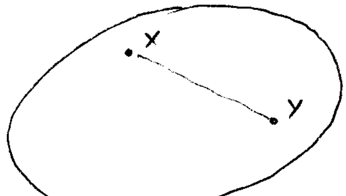
$\exists x \in D \; : \; f(x) < f(\bar{x})$. Sia $x_\lambda = \lambda x + (1-\lambda)\bar{x}$ con $\lambda \in [0,1]$: $x_\lambda \in D$

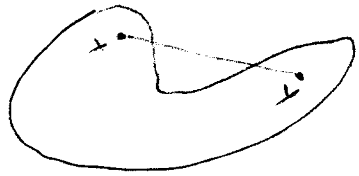
$$f(x_\lambda) \underset{\text{convessità}}{\leq} \lambda f(x) + (1-\lambda) f(\bar{x}) < \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) = f(\bar{x})$$

$x_\lambda \xrightarrow[\lambda \to 0]{} \bar{x}$ : $\forall \varepsilon > 0 \ \exists \bar{\lambda} \in [0,1]$ t.c. $x_{\bar{\lambda}} \in B(\bar{x}, \varepsilon)$ ed inoltre $f(x_{\bar{\lambda}}) < f(\bar{x})$

Quindi $\bar{x}$ non è un minimo locale (contraddizione) ∎

**Def** $D \subseteq \mathbb{R}^n$ si dice CONVESSO se $\forall x, y \in D \ \forall \lambda \in [0,1]$ : $\underbrace{\lambda x + (1-\lambda) y}_{\text{segmento } [x,y] \subseteq \mathbb{R}^n} \in D$




**Oss** $\mathbb{R}^n$ è convesso così come $\emptyset$

**Teorema** Siano $f : \mathbb{R}^n \to \mathbb{R}$ strettamente convessa, $D \subseteq \mathbb{R}^n$ convesso. Se (P) ammette un punto di minimo, allora è l'unico punto di minimo

**dim** Sia $\bar{x} \in D$ minimo (globale ≡ locale) di (P) e supponiamo che esista $\hat{x} \in D$ $\hat{x} \neq \bar{x}$ tale che $f(\hat{x}) = f(\bar{x})$. Sia $\lambda \in [0,1]$ : $\lambda \hat{x} + (1-\lambda)\bar{x} \in D$ (convessità di D)

e $f(\lambda \hat{x} + (1-\lambda)\bar{x}) < \lambda f(\hat{x}) + (1-\lambda) f(\bar{x}) = \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) = f(\bar{x})$

Quindi $\bar{x}$ non è un minimo (globale ≡ locale) ← contraddizione ∎

[Ulteriore] distinzione : ottimizzazione convessa vs ottimizzazione non convessa

(D) **Prop 1** $f$ convessa $\Rightarrow$ $f$ continua (su $\mathbb{R}^n$)

(D) **Prop 2** Sia $f : \mathbb{R}^n \to \mathbb{R}$ convessa. Allora l'insieme di sottolivello

$$C_\alpha = \{ x \in \mathbb{R}^n \mid f(x) \leq \alpha \}$$

è convesso per ogni $\alpha \in \mathbb{R}$.

**dim** Sia $C_\alpha \neq \emptyset$ e siano $x, y \in C_\alpha$. Allora per ogni $\lambda \in [0,1]$ risulta

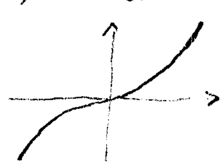
$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y) \leq \lambda \alpha + (1-\lambda)\alpha = \alpha$$

da cui $\lambda x + (1-\lambda) y \in C_\alpha$. ∎

**Nota** : l'implicazione opposta non vale

$n=1$ $f(x) = x^3$, $C_\alpha = \{ x \in \mathbb{R} \mid x^3 \leq \alpha \} = (-\infty, \sqrt[3]{\alpha}\,]$ ← convesso

non convessa :



$x = -1$
$y : 1/2$
$\lambda = 1/3$

$\lambda x + (1-\lambda) y = -\frac{1}{3} + \frac{2}{3}\left(\frac{1}{2}\right) = 0$

$f(0) = 0 > \frac{1}{3} f(-1) + \frac{2}{3} f(1/2) = -\frac{1}{3} + \frac{2}{3}\frac{1}{8} = -\frac{1}{3} + \frac{1}{12} = -\frac{1}{4}$

**Prop 3** Siano $f_i : \mathbb{R}^n \to \mathbb{R}$ convesse

(i) $\sum_{i=1}^{k} f_i$ è convessa (ii) $(\sup_{i \in I} f_i)(x) := \sup_{i \in I} f_i(x)$ è convessa (quale che sia $I$)

**dim** (i) ovvio (applicare la definizione ad ogni $f_i$ e sommare membro a membro)

(ii) (caso $I$ finito: $\exists k$ opportuno) $(\sup f_i)(\lambda x + (1-\lambda) y) = f_k(\lambda x + (1-\lambda) y) \leq \lambda f_k(x) + (1-\lambda) f_k(y) \leq \lambda (\sup f_i)(x) + (1-\lambda)(\sup f_i)(y)$ ∎

**Teo 1** Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile (su $\mathbb{R}^n$). Allora

$$f \text{ è convessa } \iff f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y \in \mathbb{R}^n$$

("il grafico di $f$ sta sopra il piano tangente a $f$ in $(x, f(x))$")

**dim** $\Rightarrow$) Sia $\lambda \in [0,1]$ qualsiasi, $x, y \in \mathbb{R}^n$

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

$$\to f(y) - f(x) \geq \frac{f(x + \lambda (y-x)) - f(x)}{\lambda} \xrightarrow[\lambda \to 0]{} \nabla f(x)^T (y-x)$$

$\Leftarrow$) Siano $\lambda \in [0,1]$, $x, y \in \mathbb{R}^n$:

① $f(x) - f(\lambda y + (1-\lambda) x) \geq \nabla f(\lambda y + (1-\lambda) x)^T [\lambda (x-y)]$

② $f(y) - f(\lambda y + (1-\lambda) x) \geq \nabla f(\lambda y + (1-\lambda) x)^T [(\lambda - 1)(x-y)]$

$(1-\lambda)$ ① $+ \lambda$ ② $\to (1-\lambda) f(x) + \lambda f(y) \geq (1-\lambda) f(\lambda y + (1-\lambda) x) + \lambda f(\lambda y + (1-\lambda) x) = f(\lambda y + (1-\lambda) x)$ ∎

**Oss** $f$ differenziabile in $\bar{x} \in \mathbb{R}^n$ soltanto $+$ $f$ convessa $\Rightarrow f(y) \geq f(\bar{x}) + \nabla f(\bar{x})^T (y - \bar{x})$

**Teo 2** Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile 2 volte (su $\mathbb{R}^n$). Allora

$f$ è convessa $\iff \nabla^2 f(x)$ è semidefinita positiva $\forall x \in \mathbb{R}^n$, ovvero $y^T \nabla^2 f(x) y \geq 0 \ \forall y \in \mathbb{R}^n$

**dim** $\Rightarrow$) Siano $x, y \in \mathbb{R}^n$ con $\|y\|_2 = 1$, $\lambda \in \mathbb{R}$:

$$f(x + \lambda y) - f(x) - \lambda \nabla f(x)^T y \geq 0 \quad (\text{teo } 1)$$

$$\text{Taylor II}^\circ \longrightarrow \ \| \qquad \tfrac{1}{2} \lambda^2 y^T \nabla^2 f(x) y + r(\lambda y)$$

Quindi $\frac{1}{2} y^T \nabla^2 f(x) y + r(\lambda y)/\lambda^2 \geq 0$ $\quad \to \|\lambda y\|_2^2 = \lambda^2 \|y\|_2^2 = \lambda^2$

$\downarrow \lambda \to 0$

$\frac{1}{2} y^T \nabla^2 f(x) y$ $\qquad$ da cui $\quad y^T \nabla^2 f(x) y \geq 0$

$\Leftarrow$) Siano $x, y \in \mathbb{R}^n$, [redacted] :

$t \in (0,1)$ opportuno

$$f(y) - f(x) - \nabla f(x)^T (y-x) = \frac{1}{2} (y-x)^T \nabla^2 f(x + t(y-x)) (y-x) \overset{\text{ipotesi}}{\geq} 0$$

resto Taylor I°

Dal <u>teo 1</u> segue che $f$ è convessa ∎

<u>Oss</u> $f$ differenziabile 2 volte in $\bar{x} \in \mathbb{R}^n$ soltanto + $f$ convessa $\Rightarrow \nabla^2 f(\bar{x})$ semidefinita positiva

$f$ concava : Teo 1 con $f(y) \leq f(x) + \nabla f(x)^T (y-x)$

Teo 2 con $\nabla^2 f(x)$ semidefinita negativa $(y^T \nabla^2 f(x) y \leq 0)$

$f$ strettamente convessa : Teo 1 con $f(y) > f(x) + \nabla f(x)^T (y-x) \quad (y \neq x)$

Teo 2 $\to$ vale la parte necessaria con $\nabla^2 f(\bar{x})$ semidefinita positiva [e non definita positiva]

$\hookrightarrow \nabla^2 f(x)$ definita positiva $\Rightarrow f$ strettamente convessa $\qquad (y^T \nabla^2 f(x) y > 0 \ \forall y \neq 0)$

(Esempio: $n=1$, $f(x) = x^4$ è strettamente convessa, ma $\nabla^2 f(0) = 0$ $[\nabla^2 f(x) = 12x^2]$ )

Ⓐ <u>Prop 4</u> Sia $f: \mathbb{R}^n \to \mathbb{R}$. Allora

$$f \text{ convessa} \iff \overbrace{\text{epi}(f) := \{(x,t) \mid x \in \mathbb{R}^n, t \geq f(x)\}}^{\text{EPIGRAFICO di } f} \subseteq \mathbb{R}^{n+1} \text{ è convesso}$$

<u>dim</u> $\Rightarrow$) Siano $(x,t), (y,\tau) \in \text{epi}(f)$, $\lambda \in [0,1]$ :

$$\lambda t + (1-\lambda)\tau \geq \lambda f(x) + (1-\lambda) f(y) \overset{\text{convessità}}{\geq} f(\lambda x + (1-\lambda) y),$$ da cui $(\lambda x + (1-\lambda) y, \lambda t + (1-\lambda)\tau) \in \text{epi}(f)$

$\Leftarrow$) Siano $x, y \in \mathbb{R}^n$, $\lambda \in [0,1]$ : $\quad (x, f(x)), (y, f(y)) \in \text{epi}(f)$

epi$(f)$ convesso $\Rightarrow (\lambda x + (1-\lambda) y, \lambda f(x) + (1-\lambda) f(y)) \in \text{epi}(f)$

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

**Prop** Siano $f: \mathbb{R}^n \to \mathbb{R}$ convessa, $D \subseteq \mathbb{R}^n$ convesso. Allora, l'insieme dei punti di minimo di (P) è convesso

**dim** Siano $\bar{x}, \hat{x} \in D$ punti di minimo di (P) : $f(\bar{x}) = f(\hat{x}) \leq f(x) \ \forall x \in D$.

Sia $x_\lambda = \lambda \hat{x} + (1-\lambda)\bar{x}$ con $\lambda \in [0,1]$ : $x_\lambda \in D$

$$f(x_\lambda) \leq \lambda f(\hat{x}) + (1-\lambda) f(\bar{x}) = \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) = f(\bar{x}) \leq f(x) \ \forall x \in D \qquad \blacksquare$$

---

**Funzioni quadratiche e convessità**

$Q \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$, $Q$ simmetrica, $c \in \mathbb{R}$

$f(x) = \frac{1}{2} x^T Q x + b^T x + c \quad \longrightarrow$ f diff. 2 volte: $\nabla^2 f(x) \equiv Q \ \ \forall x$.

Dal teo 2 : f è convessa $\Leftrightarrow$ $Q$ è semidefinita positiva (<-> gli autovalori di $Q$ sono tutti $\geq 0$ )

$\hookrightarrow$ $Q$ definita positiva $\Rightarrow$ f è strettamente convessa