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

$x, y \in D$, $\lambda \in [0,1]$ $\quad f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$ $\quad$ (f concava $\overset{def}{\Leftrightarrow}$ $(-f)$ convessa)

OPERAZIONI SULLE FUNZIONI CONVESSE

- SOMMA: f_i convesse, $i \in I$, $|I| < +\infty \Rightarrow f(x) = \sum_{i \in I} f_i(x)$ convessa
- MOLT. SCALARE: f convessa, $\mu \geq 0 \Rightarrow (\mu f)(x) = \mu f(x)$ è convessa
- COMPOSIZIONE

  Affine $\quad A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, f convessa $\Rightarrow x \mapsto f(Ax+b)$ convessa

  Convessa $\quad h: \mathbb{R} \to \mathbb{R}$

  - f convessa, h convessa non decrescente
  - f concava, h convessa non crescente

  $\Big\} \Rightarrow (h \circ f)(x) = h(f(x))$ convessa

  (possibile considerare anche composizione tra fz. vettoriali e comp. convesse)

- ESTREMO SUPERIORE PUNTUALE

  f_i convesse, $i \in I \Rightarrow (\sup_{i \in I} f_i)(x) = \sup_{i \in I} f_i(x)$ convessa (dim attraverso epi)

  Nota: con $|I| = +\infty$ può darsi che $\sup_{i \in I} f_i(x) = +\infty$ per qualche x (→ fz. convesse $\mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$)

  necessità di introdurre dominio $\operatorname{dom} f = \{x \mid f(x) < +\infty\}$

  fz poliedrali (convesse) $\equiv$ sup # finito di fz. affini

PROPRIETÀ E CARATTERIZZAZIONI

- Epigrafico: $(\operatorname{epi} f) \cap [D \times \mathbb{R}] = \{(x,t) \in D \times \mathbb{R} \mid t \geq f(x)\}$

  convesso $\Leftrightarrow$ f convessa su D

- Sottolivelli: f convessa su $D \Rightarrow \underbrace{\{x \in \mathbb{R}^n : f(x) \leq \alpha\}}_{S_f(\alpha)} \cap D$ convesso $\forall \alpha \in \mathbb{R}$

  (quasiconvessità) → $\not\Leftarrow$

FZ CONVESSE "SMOOTH"

- Caratterizzazione "piano tangente al grafico"
- Caratterizzazioni del secondo ordine.

- monotonia del gradiente

Prop Sia f differenziabile su D. Allora

i) f convessa su D $\Longleftrightarrow$ $(\nabla f(y) - \nabla f(x))^T (y-x) \geq 0 \quad \forall x,y \in D$

ii) f strett.nte convessa su D $\Longleftrightarrow$ $(\nabla f(y) - \nabla f(x))^T (y-x) > 0 \quad \forall x,y \in D, \ y \neq x$

iii) f fort.nte convessa su D di modulo μ $\Longleftrightarrow$ $(\nabla f(y) - \nabla f(x))^T (y-x) \geq \mu \|y-x\|^2 \quad \forall x,y \in D$

dim i) $\Rightarrow$) la convessità garantisce

$$f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle$$
$$f(x) \geq f(y) + \langle \nabla f(y), x-y \rangle$$

la tesi segue sommando membro a membro

$\Leftarrow$) Sia $g(t) = f(x+t(y-x))$, $x,y \in D$

f differenziabile $\Rightarrow$ g derivabile e $g'(t) = \nabla f(x+t(y-x))^T (y-x)$

L'ipotesi può essere riscritta come $g'(\lambda) \geq g'(0) \quad \forall \lambda \in [0,1]$

$$f(y) = g(1) = \text{[redacted]}\ g(0) + \int_0^1 g'(\lambda)\, d\lambda \geq g(0) + g'(0) = f(x) + \nabla f(x)^T (y-x)$$

Dall'arbitrarietà di x e y segue la convessità di f su D.

ii) analoga

iii) Applicare i) a $f(x) - \frac{\mu}{2}\|x\|^2$

<u>CONTINUITÀ (LIPSCHITZIANA) delle FZ. CONVESSE</u> (Roberts-Varberg, The American Math. Monthly 81 (1974) 1014-1016)

Teo Siano $D \subseteq \mathbb{R}^n$ convesso con $\text{int}\, D \neq \emptyset$ [redacted] e f convessa su D. Allora

i) f è localmente lipschitziana vicino ad ogni $\bar{x} \in \text{int}\, D$ [redacted]

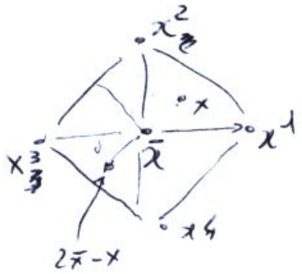
[ii) f è globalmente lipschitziana su ogni $D' \subseteq D$ compatto]

Cor f convessa su D $\Rightarrow$ f continua in ogni $x \in \text{int}\, D$ (Nota: f può essere discontinua su ∂D)

Per dimostrare il teorema, si usa il seguente

Lemma f convessa su D $\Rightarrow$ f limitata in un intorno di ogni $\bar{x} \in \text{int}\, D$

dim $\left.\begin{array}{l} x^i = \bar{x} + \delta e_i \\ x^{n+i} = \bar{x} - \delta e_i \end{array}\right\}$ con $\delta > 0$ t.c. $x^i \in D$ $\quad i = 1 \ldots n$



$$T = \left\{ \sum_{i=1}^{2n} \lambda_i x^i \;\middle|\; \lambda_i \geq 0 \quad \sum_{i=1}^{2n} \lambda_i = 1 \right\}$$

$x \in T$, $\text{int}\, T \neq \emptyset$ $\left(B(x, \tfrac{\sqrt{2}}{2}\delta) \subseteq T \right)$

$\hookrightarrow \quad x = \sum_{i=1}^{2n} \lambda_i x^i$ [redacted] $f(x) \leq \sum_{i=1}^{2n} \lambda_i f(x^i) \leq \max_{i=1\ldots 2n} f(x^i) =: M$

[redacted]

$\bar{x} = x/2 + (2\bar{x} - x)/2 \qquad \hat{x} = 2\bar{x} - x \rightarrow \hat{x} \in T$

$f(\bar{x}) \leq \frac{1}{2} f(x) + \frac{1}{2} f(\hat{x}) \quad \rightarrow \quad f(x) \geq 2f(\bar{x}) - f(\hat{x}) \geq 2f(\bar{x}) - M$

Quindi f è limitata su T ∎

dim teo i) Grazie al lemma, esistono $M = M(\bar{x})$, $\varepsilon = \varepsilon(\bar{x}) > 0$ t.c.

$$|f(x)| \leq M \quad \forall x \in B(\bar{x}, 2\varepsilon)$$

Siano $x, y \in B(\bar{x}, \varepsilon)$ con $x \neq y$, e si ponga $\alpha = \|x - y\|$. Allora

$z = x + \frac{\varepsilon}{\alpha}(x - y) \in B(\bar{x}, 2\varepsilon)$, [redacted]

$\quad \| $

$\frac{\alpha+\varepsilon}{\alpha} x - \frac{\varepsilon}{\alpha} y$ da cui $\quad x = \frac{\alpha}{\alpha+\varepsilon} z + \frac{\varepsilon}{\alpha+\varepsilon} y$

e $f(x) \leq \frac{\alpha}{\alpha+\varepsilon} f(z) + \frac{\varepsilon}{\alpha+\varepsilon} f(y)$. Sottraendo f(y) ad ambo membri:

$$f(x) - f(y) \leq \frac{\alpha}{\alpha+\varepsilon} [f(z) - f(y)] \leq \frac{\alpha}{\varepsilon} |f(z) - f(y)| \leq \frac{2M}{\varepsilon} \alpha = \frac{2M}{\varepsilon} \|x - y\|$$

Invertendo i ruoli di x e y si ottiene anche $f(y) - f(x) \leq 2M/\varepsilon \|x - y\|$, e quindi

$$|f(x) - f(y)| \leq \frac{2M}{\varepsilon} \|x - y\| \qquad \forall x, y \in B(\bar{x}, \varepsilon)$$ ∎

Oss $f:\mathbb{R}^n \to \mathbb{R}$ convessa su $\mathbb{R}^n \Rightarrow \text{epi} f,\ S_f(\alpha)$ sono chiusi $(\alpha \in \mathbb{R})$

Nota Anche in assenza di convessità vale:

f inf.nte semicontinua $\Leftrightarrow$ epi f chiuso $\Leftrightarrow S_f(\alpha)$ chiuso $\forall \alpha \in \mathbb{R}$

(le f inf.nte semicontinue vengono anche chiamate chiuse)

ESISTENZA DERIVATA DIREZIONALE

$$f'(\bar{x};d) = \lim_{t \downarrow 0} [f(\bar{x}+td) - f(\bar{x})]/t \qquad (\text{se } \exists \lim)$$

Prop Una fz. convessa su $D \subseteq \mathbb{R}^n$ ammette derivata direzionale $f'(\bar{x};d)$ in ogni direzione $d \in \mathbb{R}^n$ in ogni punto $\bar{x} \in \text{int}\, D$.

dim Sia $g(t) = [f(\bar{x}+td) - f(\bar{x})]/t$ (se t suff.nte piccolo $\bar{x}+td \in D$)

f loc.nte lips. $\Rightarrow |g(t)| = |f(\bar{x}+td) - f(\bar{x})| \frac{1}{|t|} \leq M\|d\|$ se t suff.nte piccolo $(t>0)$

Siano $0 < t_2 < t_1$ t.c. $\bar{x}+t_1 d \in D$:

$$\bar{x}+t_2 d = (1 - t_2/t_1)\bar{x} + t_2/t_1 (\bar{x}+t_1 d)$$

$$f(\bar{x}+t_2 d) \leq (1 - t_2/t_1) f(\bar{x}) + t_2/t_1 f(\bar{x}+t_1 d)$$

$$f(\bar{x}+t_2 d) - f(\bar{x}) \leq t_2/t_1 [f(\bar{x}+t_1 d) - f(\bar{x})] \to g(t_2) < g(t_1)$$

g limitata e crescente su $[0, t_1] \Rightarrow \exists f'(\bar{x};d) = \lim_{t \downarrow 0} g(t) = \inf_{t>0} g(t)$ ∎

Nota: $|f'(\bar{x};d)| \leq M\|d\|$

SOTTOGRADIENTI E SOTTODIFFERENZIALE

$f:\mathbb{R}^n \to \mathbb{R}$ convessa su $D \subseteq \mathbb{R}^n$ (D convesso)

Def $s \in \mathbb{R}^n$ si dice SOTTOGRADIENTE di f in $\bar{x} \in D$ se

$$f(x) \geq f(\bar{x}) + \langle s, x - \bar{x} \rangle \qquad \forall x \in D$$

Def $\partial f(\bar{x}) = \{ s \in \mathbb{R}^n \mid s \text{ è sottogradiente di } f \text{ in } \bar{x} \}$ si dice SOTTODIFFERENZIALE di f in $\bar{x} \in D$

Prop Sia f differenziabile in $\bar{x} \in \operatorname{int} D$. Allora $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$

dim Siano $s \in \partial f(\bar{x})$, $d = s - \nabla f(\bar{x})$ e $t > 0$ t.c. $\bar{x} + td \in D$.

$s \in \partial f(\bar{x}) \Rightarrow f(\bar{x}+td) \geq f(\bar{x}) + t \langle s, d \rangle$

f diff. $\Rightarrow f(\bar{x}+td) = f(\bar{x}) + t \langle \nabla f(\bar{x}), d \rangle + r(td)$

da cui $0 \geq t \langle s - \nabla f(\bar{x}), d \rangle + r(td)$ e $0 \geq \langle s - \nabla f(\bar{x}), d \rangle + r(td)/t$

Considerando $t \downarrow 0$ si ottiene $0 \geq \langle s - \nabla f(\bar{x}), d \rangle = \| s - \nabla f(\bar{x}) \|^2$

ovvero $s = \nabla f(\bar{x})$ ■

Teo $\partial f(\bar{x})$ è non vuoto, convesso e compatto per ogni $\bar{x} \in \operatorname{int} D$.

dim $s_1, s_2 \in \partial f(\bar{x})$, $\lambda \in [0,1]$

$$f(\bar{x}) + \langle \lambda s_1 + (1-\lambda) s_2, x - \bar{x} \rangle = \lambda [f(\bar{x}) + \langle s_1, x - \bar{x} \rangle] + (1-\lambda) [f(\bar{x}) + \langle s_2, x - \bar{x} \rangle \leq$$
$$\leq \lambda f(x) + (1-\lambda) f(x) = f(x)$$

CONVESSITÀ

$\partial f(\bar{x})$ è chiuso per la continuità di $s \mapsto \langle s, x - \bar{x} \rangle$

$\partial f(\bar{x})$ è limitato poiché f è localmente lipschitziana vicino a $\bar{x}$:

$\langle s, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \leq L \|x - \bar{x}\|$ se $x \in B(\bar{x}, \varepsilon)$ per $\varepsilon > 0$ opportuno

$x = \bar{x} + \frac{\varepsilon}{\|s\|} s$: $\frac{\varepsilon}{\|s\|} \langle s, s \rangle \leq \frac{\varepsilon}{\|s\|} L \|s\| \rightarrow \|s\| \leq L$

$\partial f(\bar{x}) \neq \emptyset$ richiede il seguente corollario del teorema di separazione stretta:

Cor $A \subseteq \mathbb{R}^m$ convesso, $\bar{z} \in \partial A \Rightarrow \exists z^* \in \mathbb{R}^m, z^* \neq 0$ t.c. $\langle z^*, z \rangle \leq \langle z^*, \bar{z} \rangle \; \forall z \in A$.

$(\bar{x}, f(\bar{x})) \in \partial(\operatorname{epi} f)$, quindi $\exists (s^*, \mu^*) \in \mathbb{R}^n \times \mathbb{R}$ $(s^*, \mu^*) \neq (0,0)$ t.c.

$$\langle s^*, x \rangle + \mu^* t \leq \langle s^*, \bar{x} \rangle + \mu^* f(\bar{x}) \qquad \forall x \in D \; \forall t \geq f(x)$$

$$\langle s^*, x - \bar{x} \rangle + \mu^* (t - f(\bar{x})) \leq 0 \qquad (*)$$

$\mu^* \leq 0$, altrimenti $t \uparrow +\infty$ contraddice $(*)$.

Se fosse $\mu^* = 0$, risulterebbe $\langle s^*, x - \bar{x} \rangle \leq 0 \; \forall x \in D$ e quindi $s^* = 0$ poiché $\bar{x} \in \operatorname{int} D$

Quindi $\mu^* < 0$ e possiamo supporre $\mu^* = -1$: $\langle s^*, x - x^* \rangle + f(\bar{x}) - t \leq 0$

da cui per $t = f(x)$ si ottiene $f(x) \geq f(\bar{x}) + \langle s^*, x - x^* \rangle$, ovvero $s^* \in \partial f(\bar{x})$ ∎

Prop. Siano f strett.nte convessa su D e $s^* \in \partial f(\bar{x})$ per $\bar{x} \in \operatorname{int} D$. Allora

$$f(x) > f(\bar{x}) + \langle s^*, x - \bar{x} \rangle \quad \forall x \in D,\ x \neq \bar{x}.$$

dim Per def. $f(x) \geq f(\bar{x}) + \langle s^*, x - \bar{x} \rangle \quad \forall x \in D$ ⊞

Supponiamo esista $\hat{x} \in D$ $(\hat{x} \neq \bar{x})$ t.c. $f(\hat{x}) = f(\bar{x}) + \langle s^*, \hat{x} - \bar{x} \rangle$. Sia $\lambda \in]0,1[$:

$$f(\lambda \hat{x} + (1-\lambda)\bar{x}) < \lambda f(\hat{x}) + (1-\lambda) f(\bar{x}) = f(\bar{x}) + \lambda \langle s^*, \hat{x} - \bar{x} \rangle$$

in contraddizione con ⊞ per $x = \lambda \hat{x} + (1-\lambda)\bar{x}$. ∎

Analogamente al caso differenziabile si dimostra:

Prop. Se per ogni $\bar{x} \in \operatorname{int} D$, $\partial f(\bar{x}) \neq \emptyset$, allora f è convessa su $\operatorname{int} D$.

Per semplicità d'ora in avanti supponiamo $f: \mathbb{R}^n \to \mathbb{R}$ convessa su $\mathbb{R}^n$

Prop. $\partial f(\bar{x}) = \{ s \in \mathbb{R}^n \mid f'(\bar{x}; d) \geq \langle s, d \rangle \ \forall d \in \mathbb{R}^n \}$

dim $\subseteq)$ $s \in \partial f(\bar{x}) \to f(\bar{x} + td) \geq f(\bar{x}) + t \langle s, d \rangle \to \dfrac{f(\bar{x}+td) - f(\bar{x})}{t} \geq \langle s, d \rangle$

$t \downarrow 0 \to f'(\bar{x}; d) \geq \langle s, d \rangle$

$\supseteq)$ Sia $s \in \mathbb{R}^n$ t.c. $f'(\bar{x}; d) \geq \langle s, d \rangle \ \forall d \in \mathbb{R}^n$.

$$\langle s, d \rangle \leq f'(\bar{x}; d) = \inf_{t > 0} \frac{f(\bar{x}+td) - f(\bar{x})}{t} \underset{t=1}{\leq} f(\bar{x}+d) - f(\bar{x})$$

da cui [redacted] con $d = y - x$: $f(y) \geq f(\bar{x}) + \langle s, y - x \rangle \quad \forall y \in \mathbb{R}^n$

ovvero $s \in \partial f(\bar{x})$ ∎

Lemma $d \mapsto f'(\bar{x}; d)$ è sublineare per ogni $\bar{x} \in \mathbb{R}^n$

dim Sia $\gamma > 0$: $f'(\bar{x}; \gamma d) = \lim_{t \downarrow 0} \gamma \dfrac{f(\bar{x} + \gamma t d) - f(\bar{x})}{\gamma t} = \gamma \lim_{\tau \downarrow 0} \dfrac{f(\bar{x} + \tau d) - f(\bar{x})}{\tau} = \gamma f'(\bar{x}; d)$

Siano $d_1, d_2 \in \mathbb{R}^n$:

$$\bar{x} + t(d_1 + d_2) = \tfrac{1}{2}(\bar{x} + 2td_1) + \tfrac{1}{2}(\bar{x} + 2td_2)$$

$$f(\bar{x}+t(d_1+d_2)) \le \tfrac{1}{2} f(\bar{x}+2td_1)+\tfrac{1}{2} f(\bar{x}+2td_2)$$ da cui

$$\frac{f(\bar{x}+t(d_1+d_2))-f(\bar{x})}{t} \le \frac{f(\bar{x}+2td_1)-f(\bar{x})}{2t}+\frac{f(\bar{x}+2td_2)-f(\bar{x})}{2t}$$ da cui $t \downarrow 0$:

$$f'(\bar{x};d_1+d_2) \le f'(\bar{x};d_1)+f'(\bar{x};d_2)$$ ∎

Teo $f'(\bar{x},\cdot)$ è la funzione di supporto di $\partial f(\bar{x})$ per ogni $\bar{x} \in \mathbb{R}^n$, ovvero

$$f'(\bar{x};d) = \max\{\langle s,d\rangle \mid s \in \partial f(\bar{x})\} \quad \forall d \in \mathbb{R}^n.$$

dim Siano $s \in \partial f(\bar{x})$, $d \in \mathbb{R}^n$

$$f'(\bar{x};d) = \inf_{t>0} \frac{f(\bar{x}+td)-f(\bar{x})}{t} \ge \inf_{t>0} \frac{t\langle s,d\rangle}{t} = \langle s,d\rangle$$

Dall'arbitrarietà di s segue $f'(\bar{x};d) \ge \max\{\langle s,d\rangle \mid s \in \partial f(\bar{x})\}$.

Supponiamo che esiste $\bar{d} \in \mathbb{R}^n$ t.c.

$$f'(\bar{x};\bar{d}) > \max\{\langle s,\bar{d}\rangle \mid s \in \partial f(\bar{x})\} = \langle \bar{s},\bar{d}\rangle$$

per l'opportuno $\bar{s} \in \partial f(\bar{x})$. Senza perdere di generalità possiamo supporre $\|\bar{d}\|=1$ (per la positiva omogeneità di $f(\bar{x};\cdot)$ e $\langle s,\cdot\rangle$)

Siano $\gamma = f'(\bar{x};\bar{d}) - \langle \bar{s},\bar{d}\rangle > 0$ e $\hat{s} = \bar{s}+\gamma\bar{d}$

$$\langle \hat{s},\bar{d}\rangle = \langle \bar{s},\bar{d}\rangle + \gamma\langle \bar{d},\bar{d}\rangle = \langle \bar{s},\bar{d}\rangle + \gamma = f'(\bar{x};\bar{d})$$

$$\langle \hat{s},-\bar{d}\rangle = -\langle \hat{s},\bar{d}\rangle = -f'(\bar{x};\bar{d}) \le f'(\bar{x},-\bar{d})$$

$$(0 = f'(\bar{x};0) = f'(\bar{x};\bar{d}-\bar{d}) \le f'(\bar{x};\bar{d}) + f'(\bar{x};-\bar{d}))$$

Quindi $\langle \hat{s},\gamma\bar{d}\rangle \le f'(\bar{x};\gamma d) \quad \forall \gamma \in \mathbb{R}$

Dal teo (Hahn-Banach) segue che esiste $s \in \mathbb{R}^n$ t.c. $\begin{cases} \langle s,\bar{d}\rangle = \langle \hat{s},\bar{d}\rangle \\ f'(\bar{x};d) \ge \langle s,d\rangle \quad \forall d \in \mathbb{R}^n \end{cases}$

Quindi $s \in \partial f(\bar{x})$, ma $\langle s,\bar{d}\rangle = \langle \hat{s},\bar{d}\rangle > \langle \bar{s},\bar{d}\rangle$ in contraddizione con la scelta di $\bar{s}$ ∎

Prop Siano $t>0$, $f_i : \mathbb{R}^n \to \mathbb{R}$ convessa su $\mathbb{R}^n$ $i=1,2$. Allora per ogni $\bar{x} \in \mathbb{R}^n$ valgono:

i) $\partial(t f_1)(\bar{x}) = t\, \partial f_1(\bar{x})$

ii) $\partial(f_1+f_2)(\bar{x}) = \partial f_1(\bar{x}) + \partial f_2(\bar{x})$

dim i) Immediata conseguenza della definizione o equivalentemente della positiva omogeneità di $f'(\bar{x};\cdot)$.

ii) I due insiemi hanno la stessa funzione di supporto:

$$\sigma_{\partial(f_1+f_2)(\bar{x})}(d) = \max\{\langle s,d\rangle \mid s \in \partial(f_1+f_2)(\bar{x})\} = (f_1+f_2)'(\bar{x};d) =$$

$$= f_1'(\bar{x};d) + f_2'(\bar{x},d) = \max\{\langle s,d\rangle \mid s\in\partial f_1(\bar{x})\} + \max\{\langle s,d\rangle \mid s\in\partial f_2(\bar{x})\}$$

$$= \max\{\langle s,d\rangle \mid s\in\partial f_1(\bar{x}) + \partial f_2(\bar{x})\} = \sigma_{\partial f_1(\bar{x})+\partial f_2(\bar{x})}(d) \quad \forall d\in\mathbb{R}^n.$$

$\partial f_i(\bar{x})$ compatto, convesso $\Rightarrow$ $\partial f_1(\bar{x}) + \partial f_2(\bar{x})$ compatto, convesso.

Due insiemi convessi e compatti hanno la stessa funzione di supporto se e solo se sono uguali *, pertanto $\partial(f_1+f_2)(\bar{x}) = \partial f_1(\bar{x}) + \partial f_2(\bar{x})$ ∎

*__Prop__ Siano $A, B \subseteq \mathbb{R}^n$ convessi e chiusi. Allora

$$\underbrace{\max\{\langle s,d\rangle \mid s\in A\}}_{\sigma_A(d)} \le \underbrace{\max\{\langle s,d\rangle \mid s\in B\}}_{\sigma_B(d)} \quad \forall d\in\mathbb{R}^n \iff A\subseteq B$$

dim $\Leftarrow$) ovvio

$\Rightarrow$) Supponiamo che esista $\bar{s}\in A$ t.c. $\bar{s}\notin B$. Per il teo (separazione stretta) esistono $\bar{d}\in\mathbb{R}^n$, $\gamma\in\mathbb{R}$ tali che

$$\langle \bar{s},\bar{d}\rangle \ge \gamma > \langle s,\bar{d}\rangle \quad \forall s\in B$$

da cui la contraddizione

$$\sigma_A(\bar{d}) \ge \langle \bar{s},\bar{d}\rangle > \gamma \ge \max\{\langle s,\bar{d}\rangle \mid s\in B\} = \sigma_B(\bar{d}).$$

[illegible]

∎

Prop** Siano $S_1, \dots, S_K \subseteq \mathbb{R}^n$ convessi e compatti e sia $S = \bigcup_{i=1}^{K} S_i$. Allora (nota: S è compatto)

$$\max_{i=1\dots K} \sigma_{S_i}(d) = \sigma_{\text{conv}\, S}(d) \quad \forall d \in \mathbb{R}^n$$

dove $\text{conv}\, S = \{ \sum_{i=1}^{n+1} \lambda_i s^i \mid s^i \in S, \lambda_i \geq 0 \ \sum_{i=1}^{n+1} \lambda_i = 1 \}$.

dim $S_i \subseteq \text{conv}\, S \Rightarrow \sigma_{S_i}(d) = \max\{ \langle s, d \rangle \mid s \in S_i \} \leq \max\{ \langle s, d \rangle \mid s \in \text{conv}\, S \} = \sigma_{\text{conv}\, S}(d)$

$s \in \text{conv}\, S \Rightarrow s = \sum_{i=1}^{n+1} \lambda_i s^i$: $\langle s, d \rangle = \sum_{i=1}^{n+1} \lambda_i \langle s^i, d \rangle \leq (\sum_{i=1}^{n+1} \lambda_i) \max_{i=1\dots n+1} \langle s^i, d \rangle =$

$= \max_i \langle s^i, d \rangle \leq \sigma_{S_{\bar{k}}}(d) \leq \max_i \sigma_{S_i}(d)$ dove $\bar{k}$ t.c. $\max_i \langle s^i, d \rangle = \langle s^{\bar{i}}, d \rangle$ e $s^{\bar{i}} \in S_{\bar{k}}$. ∎

Prop Siano $f_i : \mathbb{R}^n \to \mathbb{R}$ convessa su $\mathbb{R}^n$ $i = 1, \dots, m$ e $f(x) = \max\{ f_1(x), f_2(x), \dots, f_m(x) \}$. Allora per ogni $\bar{x} \in \mathbb{R}^n$ risulta

$$\partial f(\bar{x}) = \text{conv}\{ \bigcup_{i \in I_f(\bar{x})} \partial f_i(\bar{x}) \} \quad \text{dove } I_f(\bar{x}) = \{ i \mid f(\bar{x}) = f_i(\bar{x}) \}$$

dim Siano $s \in \partial f_i(\bar{x})$ con $i \in I_f(\bar{x})$ e $y \in \mathbb{R}^n$:

$$f(y) \geq f_i(y) \geq f_i(\bar{x}) + \langle s, y - \bar{x} \rangle = f(\bar{x}) + \langle s, y - \bar{x} \rangle$$

da cui $s \in \partial f(\bar{x})$. Poiché $\partial f(\bar{x})$ è convesso e $\partial f_i(\bar{x}) \subseteq \partial f(\bar{x})$ per $i \in I_f(\bar{x})$, segue che $\text{conv}\{ \bigcup_{i \in I_f(\bar{x})} \partial f_i(\bar{x}) \} \subseteq \partial f(\bar{x})$.

Per le due precedenti proposizioni l'inclusione opposta vale se

$$f'(\bar{x}; d) \leq \max\{ f_i'(\bar{x}; d) \mid i \in I_f(\bar{x}) \} \quad \forall d \in \mathbb{R}^n.$$

Siano $d \in \mathbb{R}^n$ e $t_k \downarrow 0$: sia $I_k = I_f(\bar{x} + t_k d)$; per costruzione $I_k \neq \emptyset$ per ogni k e la continuità delle f.z. f_i garantisce che $I_k \subseteq I_f(\bar{x})$ se k è suff.nte grande.

Per la finitezza del numero di f.z. esistono $j \in I_f(\bar{x})$ e una sottosuccessione k_ℓ t.c. $j \in I_{k_\ell}$ per ogni ℓ. Quindi:

$$f'(\bar{x}; d) = \lim_{\ell \to +\infty} [f(\bar{x} + t_{k_\ell} d) - f(\bar{x})] / t_{k_\ell} = \lim_{\ell \to +\infty} [f_j(\bar{x} + t_{k_\ell} d) - f_j(\bar{x})] / t_{k_\ell} = f_j'(\bar{x}; d) \leq$$

$$\leq \max\{ f_i'(\bar{x}; d) \mid i \in I_f(\bar{x}) \}$$ ∎

Cor Se f_i sono inoltre differenziabili in $\bar{x}$, allora $\partial f(\bar{x}) = \text{conv}\{ \nabla f_i(\bar{x}) \mid i \in I_f(\bar{x}) \}$.

Prop Sia $f : \mathbb{R}^n \to \mathbb{R}$ convessa su $\mathbb{R}^n$. Allora per ogni $\bar{x} \in \mathbb{R}^n$ risulta

$$\partial f(\bar{x}) = \{ s \in \mathbb{R}^n \mid (s, -1) \in N(\text{epi} f, (\bar{x}, f(\bar{x}))) \}$$

dim $(s,-1) \in N(\text{epi} f, (\bar{x}, f(\bar{x}))) \iff \forall y \in \mathbb{R}^n \ \forall r \geq f(y) : \langle s, y - \bar{x} \rangle + (-1)(r - f(\bar{x})) \leq 0$

$\iff \forall y \in \mathbb{R}^n \ \forall r \geq f(y) : \langle s, y - \bar{x} \rangle + f(\bar{x}) \leq r$

$\Rightarrow$ si prende $r = f(y)$

$\iff s \in \partial f(\bar{x})$ ∎

TEOREMI DI SEPARAZIONE

Mangasarian, Nonlinear programming, SIAM 1994 (McGraw-Hill 1969)
Chp 3 (richiede Chp 2 [teoremi di alternativa])

Bazaraa - Sherali - Shetty, Nonlinear programming: theory & algorithms, Wiley 1993
Section 2.4

Teo (separazione stretta)

$A \subseteq \mathbb{R}^m$ convesso, chiuso, $B \subseteq \mathbb{R}^m$ compatto.

$A \cap B = \emptyset \Rightarrow \exists\, z^* \in \mathbb{R}^m, \gamma \in \mathbb{R}$ t.c. $\langle z^*, z \rangle < \gamma < \langle z^*, w \rangle \quad \forall z \in A \; \forall w \in B$

Cor $A \subseteq \mathbb{R}^m$ convesso, $\bar{z} \in \partial A \Rightarrow \exists\, z^* \in \mathbb{R}^m, z^* \neq 0$ t.c. $\langle z^*, z - \bar{z} \rangle \leq 0 \quad \forall z \in A$

dim $\bar{z} \in \partial A \Rightarrow \exists\, z_k \to \bar{z}$ con $z_k \notin \mathrm{cl}\, A$

Si applica il teo (separazione stretta) per $\mathrm{cl}\, A$ e $\{z_k\}$ e si normalizza z_k^* e si passa al limite $k \to +\infty$.

Esempi di fz. convesse

1) $x \mapsto \|x\|$ [non differenziabile in $\bar{x} = 0$]

2) $x \mapsto \|x\|^2$ [differenziabile]

la convessità segue immediatamente dall'uguaglianza

$$\|\lambda x + (1-\lambda) y\|^2 = \lambda \|x\|^2 + (1-\lambda)\|y\|^2 - \lambda(1-\lambda)\|x-y\|^2$$

$$\Big(\lambda <x,x> + (1-\lambda) <y,y> - \lambda(1-\lambda) <x-y, x-y> =$$
$$= \lambda <x,x> + (1-\lambda) <y,y> - \lambda(1-\lambda) [<x,x> + <y,y> - 2<x,y>] =$$
$$= \lambda^2 <x,x> + (1-\lambda)^2 <y,y> + 2\lambda(1-\lambda) <x,y> = \|\lambda x + (1-\lambda) y\|^2 \Big)$$

3) $x \mapsto \frac{1}{2} x^T Q x + b^T x$ con $Q \in \mathbb{R}^{n \times n}$, $Q = Q^T$ semidefinita positiva

4) $x \mapsto \overbrace{\inf \{\|x-y\|^2 : y \in C\}}^{f(x)}$ con C convesso

Nota $y \mapsto \|x-y\|^2$ è fortmente convessa, quindi $\inf = \min > -\infty$.

Siano $\varepsilon > 0$, $y_1, y_2 \in C$ t.c. $\|x_i - y_i\|^2 \le f(x_i) + \varepsilon$ $i=1,2$, $\lambda \in [0,1]$

$$\begin{aligned} f(\lambda x_1 + (1-\lambda) x_2) &\le \|\lambda x_1 + (1-\lambda) x_2 - \lambda y_1 - (1-\lambda) y_2\|^2 = \\ &= \|\lambda (x_1 - y_1) + (1-\lambda)(x_2 - y_2)\|^2 \le \\ &\le \lambda \|x_1 - y_1\|^2 + (1-\lambda)\|x_2 - y_2\|^2 \le \\ &\le \lambda f(x_1) + \lambda \varepsilon + (1-\lambda) f(x_2) + (1-\lambda)\varepsilon \\ &\le \lambda f(x_1) + (1-\lambda) f(x_2) + \varepsilon \end{aligned}$$

Dall'arbitrarietà di $\varepsilon > 0$ segue la convessità di f

5) $x \mapsto \sup \{<x,y> : y \in C\}$ funzione di supporto di $C \subseteq \mathbb{R}^n$

6) $Q = Q^T \in \mathbb{R}^{n \times n}$ semidefinita positiva

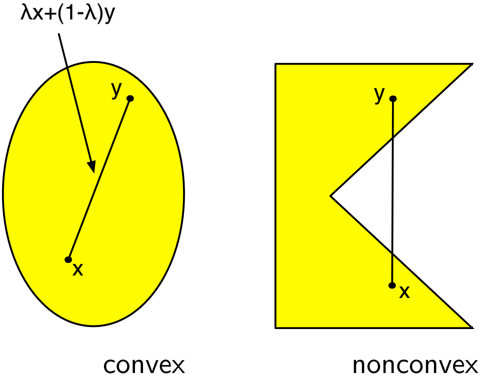
$f(Q) = \sup \{d^T Q d : \|d\| = 1\}$ massimo autovalore di Q.

(f definita sull'insieme delle matrici semidef. pos. $\subseteq \mathbb{R}^{2n}$)

Convexity for sets

$C \subseteq \mathbb{R}^n$ is a convex set if

$$x, y \in C,\ \lambda \in [0,1] \implies \lambda x + (1-\lambda) y \in C$$



Operations preserving convexity

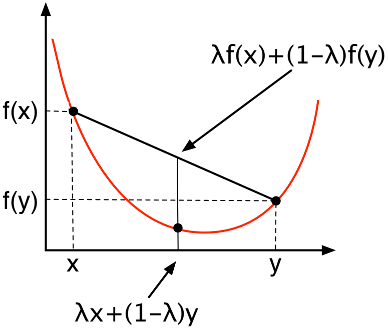
- Let $C \subseteq \mathbb{R}^n$ be convex.
  - The closure and the (relative) interior of C are convex
- Let $A \in \mathbb{R}^{m \times n}$.
  (i) If $C \subseteq \mathbb{R}^n$ is convex, then $A(C) = \{ Ax : x \in C \}$ is convex
  (ii) If $D \subseteq \mathbb{R}^m$ is convex, then $A^{-1}(D) = \{ x : Ax \in D \}$ is convex
- Let $C_i \subseteq \mathbb{R}^{n_i}$ be convex and $\mu_i \in \mathbb{R}$ for all $i \in I$.
  (i) $\prod_{i \in I} C_i$ is convex
  (ii) If $n_i = n$, then $\bigcap_{i \in I} C_i$ is convex
  (iii) If I is finite and $n_i = n$, then $\sum_{i \in I} \mu_i C_i$ is convex

Convexity for functions

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function on C if

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

holds for all $x, y \in C$, $\lambda \in [0,1]$.

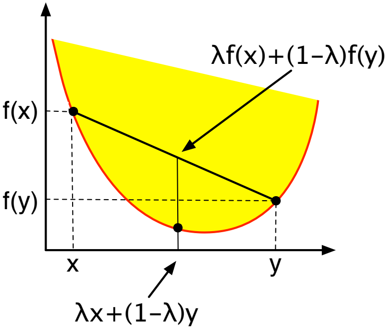


Convexity for functions

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function on C if

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

holds for all $x, y \in C$, $\lambda \in [0,1]$.



Convexity for functions

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function on C if

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

holds for all $x, y \in C$, $\lambda \in [0,1]$.

Proposition

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is convex on C if and only if (the restriction of) its epigraph (to C), namely

$$epi(f) \cap (C \times \mathbb{R}) = \{(x,t) \in C \times \mathbb{R} \ : \ t \geq f(x)\},$$

is a convex set in $\mathbb{R}^{n+1}$.

Convexity for functions

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function on C if

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

holds for all $x, y \in C$, $\lambda \in [0,1]$.

Proposition

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is convex on C if and only if (the restriction of) its epigraph (to C), namely

$$epi(f) \cap (C \times \mathbb{R}) = \{(x,t) \in C \times \mathbb{R} \ : \ t \geq f(x)\},$$

is a convex set in $\mathbb{R}^{n+1}$. Morever, if f is convex on C, then the (the restriction of) its α-sublevel set (to C), namely

$$S_\alpha(f) \cap C = \{x \in C \ : \ f(x) \leq \alpha\},$$

is a convex set in $\mathbb{R}^n$ for any $\alpha \in \mathbb{R}$.

Let $C \subseteq \mathbb{R}^n$ be convex. $f : \mathbb{R}^n \to \mathbb{R}$ is

- convex on C if

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

  holds for all $x, y \in S$, $\lambda \in [0,1]$

- strictly convex on C if

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

  holds for all $x, y \in C$ with $x \neq y$, $\lambda \in]0,1[$

- strongly convex on C with modulus μ if

  $$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\mu}{2}\lambda(1-\lambda)\|x-y\|_2$$

  holds for all $x, y \in C$, $\lambda \in [0,1]$

f strongly convex on C if and only if $f - \frac{\mu}{2}\|\cdot\|_2^2$ is convex on C.

Operations preserving convexity

Let $C \subseteq \mathbb{R}^n$ be convex.

- Let $f_i : \mathbb{R}^n \to \mathbb{R}$ be convex on C and $\mu_i > 0$ for all $i \in I$.
  - (i) If I is finite, then $\left(\sum_{i \in I} \mu_i f_i\right)(x) = \sum_{i \in I} \mu_i f_i(x)$ is convex on C
  - (ii) If $\left(\sup_{i \in I} f_i\right)(x) = \sup_{i \in I} f_i(x) < +\infty$ for all $x \in C$, then the pointwise supremum function $\left(\sup_{i \in I} f_i\right)$ is convex on C
- Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex on $\mathbb{R}^n$.
  - (i) $g(x) = f(Ax + b)$ is convex on $\mathbb{R}^m$ for any $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$
  - (ii) $g(x) = h(f(x))$ is convex on $\mathbb{R}^n$ if $h : \mathbb{R} \to \mathbb{R}$ is convex and non-decreasing
  - (iii) $g(x) = h(-f(x))$ is convex on $\mathbb{R}^n$ if $h : \mathbb{R} \to \mathbb{R}$ is convex and non-increasing

Examples of convex functions

- $f(x) = \|x\|_2$ (nonsmooth at $x = 0$)
- $f(x) = \|x\|_2^2$ (smooth)
- $f(x) = \|x\|_1$ (nonsmooth along the coordinate axes)
- $f(x) = \frac{1}{2} x^T Q x + b^T x + c$ with $Q = Q^T \in \mathbb{R}^{n \times n}$ positive semidefinite, $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ (strongly convex if Q is positive definite)
- Let $C \subseteq \mathbb{R}^n$ be convex.
  (i) $d_C(x) = \min\{\, \|y - x\|_2 \;:\; y \in C\}$ (nonsmooth on the boundary)
  (ii) $d_C^2(x) = \min\{\, \|y - x\|_2^2 \;:\; y \in C\}$ (smooth)
  (iii) $\sigma_C(x) = \sup\{\, y^T x \;:\; y \in C\}$ [support function of C]

Convexity and optimization

Let $C \subseteq \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$.

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

Suppose C is convex. Then,

- If f is convex on C, then any local minimum point of (P) is also a global minimum point. Moreocer, the set of all the minima is a convex set.
- If f is strictly convex on C, there exists at most one minimum point of (P).
- If f is strongly convex on C, there exists exactly one minimum point of (P).

Differentiable convex functions

Let $C \subseteq \mathbb{R}^n$ be convex and $f : \mathbb{R}^n \to \mathbb{R}$ differentiable on C.

(i) f is (strictly) convex on C if and only if

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

or equivalently

$$\big(\nabla f(x) - \nabla f(y)\big)^T (x - y) \geq 0 \qquad (>)$$

holds for all $x, y \in C$ (with $x \neq y$).

Differentiable convex functions

Let $C \subseteq \mathbb{R}^n$ be convex and $f : \mathbb{R}^n \to \mathbb{R}$ differentiable on C.

(i) f is (strictly) convex on C if and only if

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

or equivalently

$$\big(\nabla f(x) - \nabla f(y)\big)^T (x - y) \geq 0 \qquad (>)$$

holds for all $x, y \in C$ (with $x \neq y$).

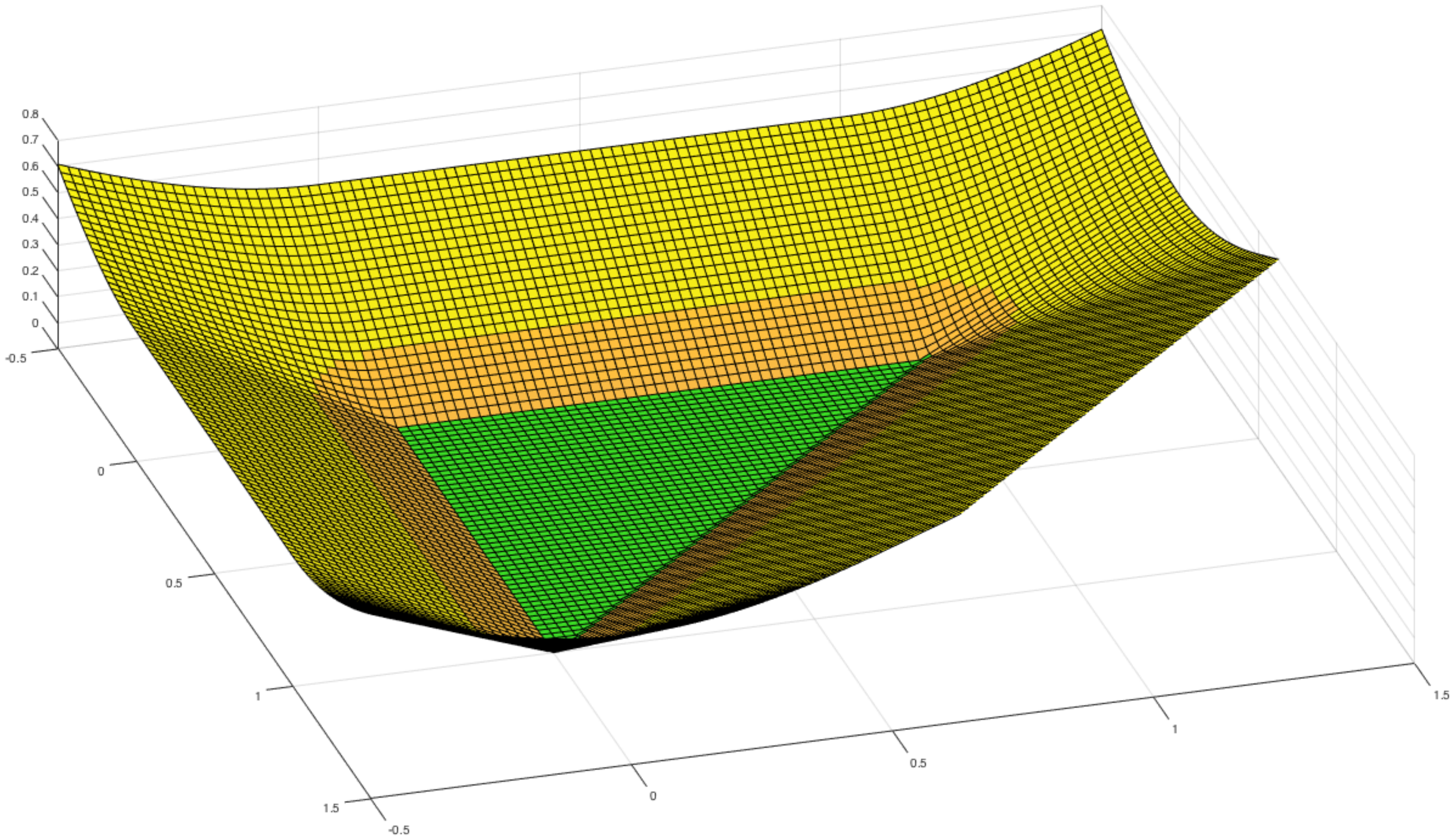
(ii) f is strongly convex on C with modulus $\mu > 0$ if and only if

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{\mu}{2} \|y - x\|_2^2$$

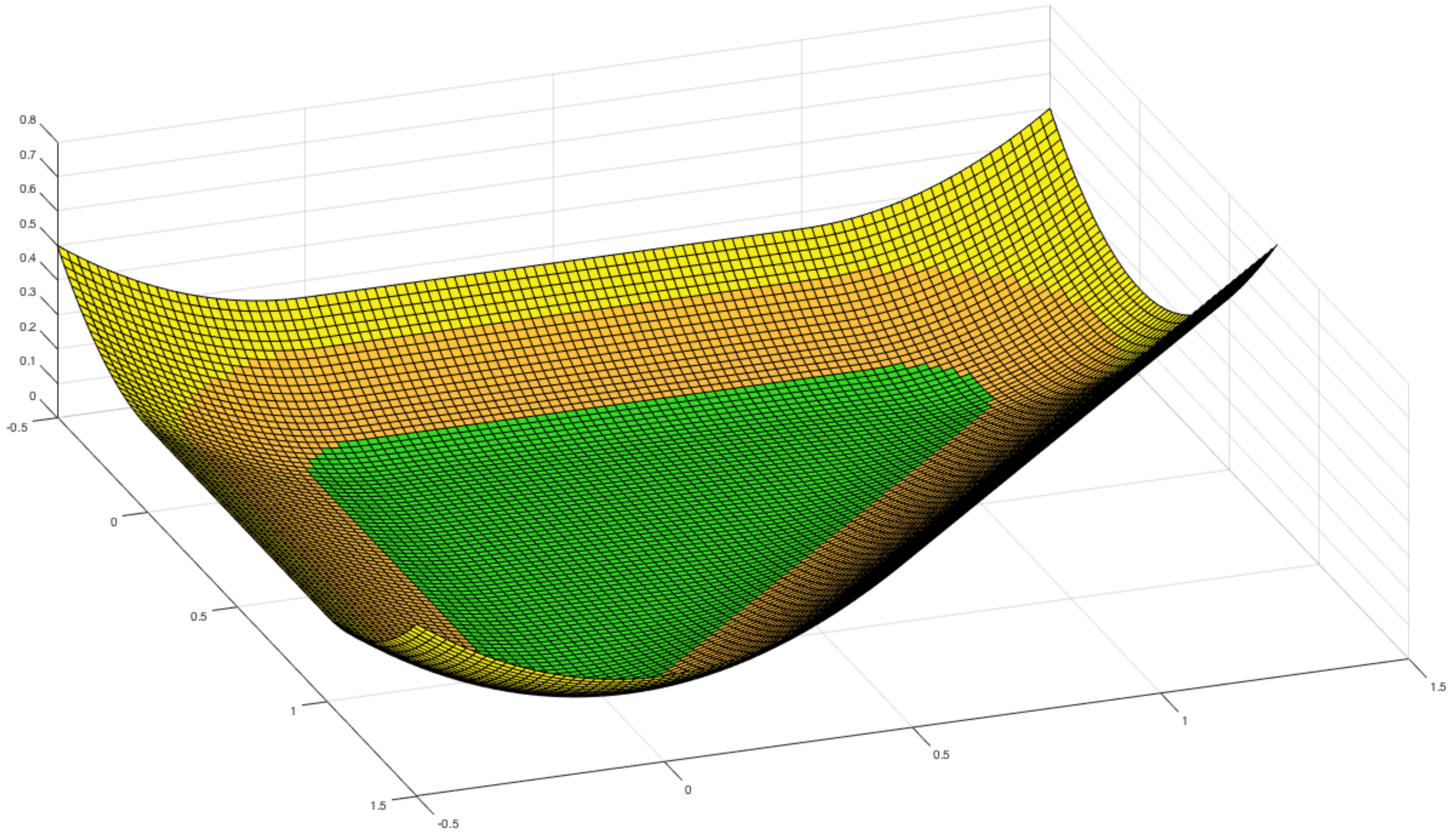
or equivalently

$$\big(\nabla f(x) - \nabla f(y)\big)^T (x - y) \geq \mu \|y - x\|_2^2$$

holds for all $x, y \in C$.



0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.5
0
0.5
1
1.5
-0.5
0
0.5
1
1.5


0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.5
0
0.5
1
1.5
-0.5
0
0.5
1
1.5

Mean value theorem

Let f be convex on $\mathbb{R}^n$. Given any $x, y \in \mathbb{R}^n$ with $x \neq y$, then

- there exist $\lambda \in]0,1[$ and $s_\lambda \in \partial f(\lambda x + (1-\lambda)y)$ such that
$$f(y) - f(x) = s_\lambda^T (y - x)$$

- $f(y) - f(x) = \int_0^1 s_\lambda^T (y-x)\, d\lambda$

  for any choice of the subgradients $s_\lambda \in \partial f(\lambda x + (1-\lambda)y)$

Calculus rules

Let $f, f_i : \mathbb{R}^n \to \mathbb{R}$ be convex on $\mathbb{R}^n$ for all $i \in I$ [$\leftarrow$ finite index set]

- $\partial f(x) = \alpha \partial f(x)$ for all $\alpha > 0$
- $\partial \Big(\sum_{i \in I} f_i \Big)(x) = \sum_{i \in I} \partial f_i(x)$
- Let $I(x) = \{ j \in I \ : \ f_j(x) = \max_{i \in I} f_i(x) \}$

$$\partial \Big(\max_{i \in I} f_i \Big)(x) = \mathrm{conv} \bigcup_{i \in I(x)} \partial f_i(x)$$

- Let $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, $g(x) = f(Ax + b)$

$$\partial g(x) = A^T \partial f(Ax + b)$$

- Let $h : \mathbb{R} \to \mathbb{R}$ be convex, non-decreasing and differentiable

$$\partial (h \circ f)(x) = h'(f(x))\, \partial f(x)$$