Formulazione di problemi con valori assoluti

- funzioni convesse lineari a tratti
- problemi con valori assoluti
- regressione lineare

rif. BT 1.3

Max di funzioni convesse

Teorema
Siano $f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ funzioni convesse. Allora la funzione $f(\mathbf{x}) = \max_{i=1,\ldots,m} f_i(\mathbf{x})$ è convessa

Dimostrazione Siano $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in [0,1]$. Si ha che:

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = \max_{i=1,\ldots,m} f_i(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$$

$$\leq \max_{i=1,\ldots,m} (\lambda f_i(\mathbf{x}) + (1-\lambda) f_i(\mathbf{y}))$$

$$\leq \max_{i=1,\ldots,m} \lambda f_i(\mathbf{x}) + \max_{i=1,\ldots,m} (1-\lambda) f_i(\mathbf{y})$$

$$= \lambda f(\mathbf{x}) + (1 - \lambda f(\mathbf{y}))$$

$\square$

Somma di funzioni convesse

Teorema

Siano $f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ funzioni convesse. Allora la funzione $f(\mathbf{x}) = \sum_{i=1}^m f_i(\mathbf{x})$ è convessa

Dimostrazione Proviamo il risultato per $m = 2$, il caso generale segue per induzione. Siano $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in [0, 1]$. Si ha che:

$$f(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = f_1(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) + f_2(\lambda\mathbf{x} + (1-\lambda)\mathbf{y})$$

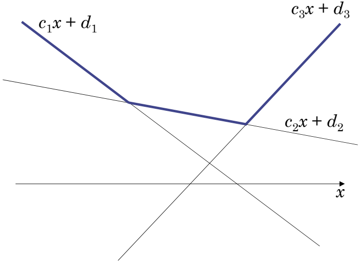
$$\leq \lambda f_1(\mathbf{x}) + (1-\lambda)f_1(\mathbf{y}) + \lambda f_2(\mathbf{x}) + (1-\lambda)f_2(\mathbf{y})$$

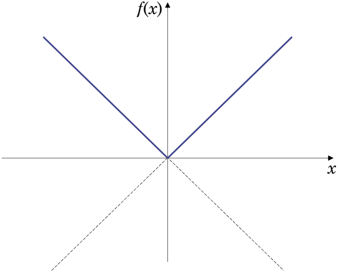
$$= \lambda f(\mathbf{x}) + (1 - \lambda f(\mathbf{y}))$$

$\square$

Funzioni convesse lineari a tratti

Una funzione del tipo $\max_{i=1,\ldots,m}(\mathbf{c}_i^T\mathbf{x} + d_i)$ è detta funzione *convessa lineare a tratti*





$$f(x) = |x| = max\{x, -x\}$$

Generalizzazione del problema di PL

$$\begin{aligned} \min \max_{i=1,\dots,m} & (\mathbf{c}_i^T \mathbf{x} + d_i) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

Si osservi che $\max_{i=1,\dots,m}(\mathbf{c}_i^T \mathbf{x} + d_i)$ è uguale al minimo numero z che soddisfa $z \geq \mathbf{c}_i^T \mathbf{x} + d_i$. Quindi, è equivalente al seguente problema di PL

$$\begin{aligned} & \min z \\ \text{s.t.} \quad & z \geq \mathbf{c}_i^T \mathbf{x} + d_i \\ & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

Allo stesso modo, un vincolo $f(\mathbf{x}) = \max_{i=1,\dots,m}(\mathbf{w}_i^T \mathbf{x} + g_i) \leq h$ può essere sostituito dagli m vincoli

$$\mathbf{w}_i^T \mathbf{x} + g_i \leq h, \qquad , i = 1, \dots, m$$

Problemi con valori assoluti

$$\begin{aligned} \min & \sum_{i=1,\ldots,n} c_i |x_i| \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

in cui assumiamo $c_i \geq 0$.
Osserviamo che $|x_i|$ è pari al più piccolo numero z_i tale che $z_i \geq x_i$ e $z_i \geq -x_i$. Quindi otteniamo il problema di PL equivalente:

$$\begin{aligned} \min & \sum_{i=1,\ldots,n} c_i z_i \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & z_i \geq x_i \\ & z_i \geq -x_i \end{aligned}$$

Esercizio

$$\begin{aligned} & \min 2x_1 + 3|x_2 - 10| \\ \text{s.t.} \quad & |x_1 + 2| + |x_2| \leq 5 \end{aligned}$$

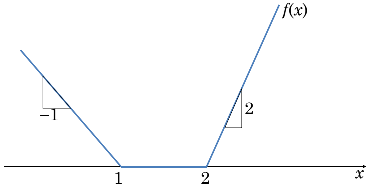
riformulazione lineare:

$$\begin{aligned} & \min 2x_1 + 3z_1 \\ \text{s.t.} \quad & z_2 + z_3 \leq 5 \\ & z_1 \geq x_2 - 10 \\ & z_1 \geq -x_2 + 10 \\ & z_2 \geq x_1 + 2 \\ & z_2 \geq -x_1 - 2 \\ & z_3 \geq x_2 \\ & z_3 \geq -x_2 \end{aligned}$$

Esercizio

Consideriamo un problema della forma:

$$\begin{aligned} \min \mathbf{c}^T\mathbf{x} + f(\mathbf{d}^T\mathbf{x}) \\ \text{s.t.} \qquad \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$



Osserviamo che $f(x) = \max\{1 - x, 0, 2x - 4\}$. Quindi otteniamo la riformulazione lineare:

$$\begin{aligned} \min \mathbf{c}^T\mathbf{x} + z \\ \text{s.t.} \qquad z \geq -\mathbf{d}^T\mathbf{x} + 1 \\ z \geq 0 \\ z \geq 2\mathbf{d}^T\mathbf{x} - 4 \\ \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

Regressione lineare

Nei problemi di classificazione un data-set composto da coppie $(\mathbf{a}_i, b_i), i = 1, \dots, m$, in cui $\mathbf{a}_i \in \mathbb{R}^n$ è un'osservazione e b_i il corrispondente risultato.

Un modello di regressione lineare consiste in un vettore $\mathbf{x}$ (da determinare) per cui il risultato di una generica osservazione $\mathbf{a}$ possa rappresentarsi come $b = \mathbf{a}^T\mathbf{x}$

Tale vettore è scelto in modo da minimizzare l'errore sui campioni del data-set:

- criterio I: minimizzare lo scarto massimo $\max_{i=1,\dots,m} |b_i - \mathbf{a}_i^T\mathbf{x}|$
- criterio II: minimizzare la somma degli scarti $\sum_{i=1}^m |b_i - \mathbf{a}_i^T\mathbf{x}|$

Regressione lineare: formulazione di PL

criterio I:

$$
\begin{aligned}
& \min z \\
\text{s.t.} \quad & z \geq b_i - \mathbf{a}_i^T \mathbf{x}, i = 1, \ldots, m \\
& z \geq -b_i + \mathbf{a}_i^T \mathbf{x}, i = 1, \ldots, m
\end{aligned}
$$

criterio II:

$$
\begin{aligned}
& \min \sum_{i=1,\ldots,m} z_i \\
\text{s.t.} \quad & z_i \geq b_i - \mathbf{a}_i^T \mathbf{x}, i = 1, \ldots, m \\
& z_i \geq -b_i + \mathbf{a}_i^T \mathbf{x}, i = 1, \ldots, m
\end{aligned}
$$