Risoluzione ai minimi quadrati

Sia $V_n^{\cdot}(\mathbb{R})$ uno spazio vettoriale Euclideo, cioè dotato di un prodotto scalare definito positivo. Sia inoltre per ogni $\mathbf{x} \in V_n$, $||\mathbf{x}|| := ||\mathbf{x}||_2 = \sqrt{\mathbf{x} \cdot \mathbf{x}}$.

Lemma 1. *Sia $\mathbf{v} \in V_n$ e si consideri un sottospazio W. Allora esiste un'unica coppia di vettori $(\mathbf{v}_\|, \mathbf{v}_\perp) \in W \times W^\perp$ tale che $\mathbf{v} = \mathbf{v}_\| + \mathbf{v}_\perp$.*

Dimostrazione. Basta osservare che $\mathbf{v} \in W \oplus W^\perp$. □

Definizione 1. *Dato un vettore $\mathbf{v} \in V_n$ ed un sottospazio W di V_n si dice* proiezione ortogonale *di $\mathbf{v}$ su W il vettore $\mathbf{v}_\| = \mathbf{v} - \mathbf{v}_\perp$ ove $\mathbf{v}_\|$ e $\mathbf{v}_\perp$ sono definiti come sopra. Poniamo $\pi_W(\mathbf{v}) := \mathbf{v}_\|$.*

Teorema 1. *Vale la seguente proprietà:*

$$\pi(\mathbf{v}) = \operatorname{argmin}_{\mathbf{w} \in W} ||\mathbf{v} - \mathbf{w}||.$$

Dimostrazione. Sia $(\mathbf{e_1}, \dots, \mathbf{e_k})$ una base ortogonale di W. Possiamo scrivere

$$\mathbf{v}_\| := \sum_{i=1}^{k} \frac{\mathbf{v} \cdot \mathbf{e_i}}{\mathbf{e_i} \cdot \mathbf{e_i}} \mathbf{e_i}, \qquad \mathbf{v}_\perp := \mathbf{v} - \mathbf{v}_\|;$$

infatti $\mathbf{v}_\| \in W$. Inoltre per ogni $\mathbf{w} = \sum_{i=1}^{k} w_k \mathbf{e_k} \in W$ si ha

$$\mathbf{v}_\perp \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} - \mathbf{v}_\| \cdot \mathbf{w} = \sum_{i=1}^{k} w_k (\mathbf{v} \cdot \mathbf{e_i}) - \sum_{i=1}^{k} \sum_{j=1}^{k} w_k \frac{\mathbf{v} \cdot \mathbf{e_i}}{\mathbf{e_i} \cdot \mathbf{e_i}} (\mathbf{e_i} \cdot \mathbf{e_j}) =$$

$$\sum_{i=1}^{k} w_k (\mathbf{v} \cdot \mathbf{e_i}) - \sum_{i=1}^{k} w_k (\mathbf{v} \cdot \mathbf{e_i}) = 0.$$

Pertanto, $\mathbf{v}_\| = \pi_W(\mathbf{v})$.

Prendiamo ora $\mathbf{w} \in W$. Scriviamo $\mathbf{w} = \mathbf{v}_\| + \mathbf{u}$ con $\mathbf{u} \in W$. Allora,

$$||\mathbf{v} - \mathbf{w}||^2 = ||\mathbf{v}_\perp + \mathbf{v}_\| - \mathbf{w}||^2 = ||\mathbf{v}_\perp + \mathbf{v}_\| - \mathbf{v}_\| + \mathbf{u}||^2 =$$

$$||\mathbf{v}_\perp + \mathbf{u}||^2 = (\mathbf{v}_\perp + \mathbf{u}) \cdot (\mathbf{v}_\perp + \mathbf{u}) = ||\mathbf{v}_\perp||^2 + ||\mathbf{u}||^2.$$

Chiaramente questa quantità è minima se e solamente se $||\mathbf{u}|| = 0$, da cui $\mathbf{u} = \mathbf{0}$. In particolare $\pi_W(\mathbf{v}) = \mathbf{v}_\|$. □

1 Minimi quadrati

Consideriamo un sistema lineare $AX = B$ ed indichiamo con $\mathscr{C}$ la copertura lineare dello spazio delle colonne di A. Per il teorema di Rouché–Capelli, il sistema è compatibile se e solamente se $B \in \mathscr{C}$. Qualora sia dato un sistema non compatibile

$$AX = B \tag{1}$$

possiamo sempre considerare il sistema compatibile *approssimato*

$$AX = \pi_{\mathscr{C}}(B). \tag{2}$$

Vediamo ora di ricavare equazioni standard per (2). Innanzi tutto, supponiamo di avere una soluzione Y per (2). Allora $AY = \pi_{\mathscr{C}}(B)$ e dunque $(AY - B) \in \mathscr{C}^\perp$. In particolare, per ogni vettore $C \in \mathscr{C}$ dobbiamo avere $C \cdot (AY - B) = 0$ cioè (in forma matriciale) ${}^{T}C(AY - B) = 0$. Estraendo un sistema di generatori dalle colonne di A questo implica ${}^{T}A(AY - B) = 0$, ovvero Y deve soddisfare

$$({}^{T}AA)X = {}^{T}AB. \tag{3}$$

Viceversa: supponiamo che Y soddisfi (3). Allora il vettore $AY - B$ è necessariamente ortogonale a tutti gli elementi di $\mathscr{C}$. In particolare, poiché $B = (B - AY) + AY$ vediamo che AY deve essere l'unico vettore B' di $\mathscr{C}$ tale che $B = B' + B''$ con $B'' \in \mathscr{C}^\perp$. Ne segue $B' = \pi_{\mathscr{C}}(B)$ e dunque Y è soluzione di (2).

Sulle norme

Definizione 2. *Una funzione* $||\cdot|| : V(\mathbb{K}) \to \mathbb{R}$ *è detta* norma *se soddisfa le seguenti proprietà:*

1. $\forall \mathbf{x} \in V : ||\mathbf{x}|| \geq 0$;
2. $||\mathbf{x}|| = 0 \Rightarrow \mathbf{x} = \mathbf{0}$.
3. $\forall \mathbf{x} \in V, \alpha \in \mathbb{R} : ||\alpha \mathbf{x}|| = |\alpha| \cdot ||\mathbf{x}||$;
4. $\forall \mathbf{x}, \mathbf{y} \in V : ||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$.

Il seguente teorema segue immediatamente dalle proprietà del prodotto euclideo.

Teorema 2. *Sia* $\cdot : V(\mathbb{R}) \times V(\mathbb{R}) \to \mathbb{R}$ *un prodotto scalare definito positivo (euclideo). Allora la funzione*

$$||\mathbf{x}|| := \sqrt{\mathbf{x} \cdot \mathbf{x}}$$

è una norma.

Noi abbiamo considerato il problema dei minimi quadrati sostituendo al vettore B nel caso di un sistema lineare non compatibile $AX = B$ il vettore $B' = \text{argmin}_{U \in \mathscr{C}} ||B - U||$ rispetto la norma euclidea.

Consideriamo ora altre possibili norme. Sia $\mathbf{x} = (x_1, \ldots, x_n)$. Poiniamo:

$$||\mathbf{x}||_1 := \sum_{i=1}^{n} |x_i|;$$

$$||\mathbf{x}||_2 := \left(\sum_{i=1}^{n} {x_i}^2 \right)^{\frac{1}{2}};$$

$$||\mathbf{x}||_\infty := \max \{ |x_i| : i = 1, \ldots, n \}.$$

Chiaramente $||\mathbf{x}||_2$ è la norma euclidea. Consideriamo ora il sistema

$$\begin{cases} x = b_1 \\ x = b_2 \\ x = b_3 \end{cases} \tag{4}$$

con $b_1 \geq b_2 \geq b_3$ ed approssimiamolo nelle divese norme introdotte prima.

1. rispetto $||\mathbf{x}||_1$: dobbiamo minimizzare la quantità $|x - b_1| + |x - b_2| + |x - b_3| = 0$. Tale minimo si ha per $x = b_2$, quindi il sistema (4) si approssima con

   $$x = b_2,$$

   ovvero l'equazione "centrale".

2. rispetto $||\mathbf{x}||_2$: la funzione da minimizzare in questo caso è $(x - b_1)^2 + (x - b_2)^2 + (x - b_3)^2 = 3x - 2(b_1 + b_2 + b_3)x + (b_1^2 + b_2^2 + b_3^2)$. Considerazioni geometriche elementari (o il calcolo della derivata) mostrano che il minimo si ha per

   $$x = \frac{1}{3}(b_1 + b_2 + b_3),$$

   ovver la "media aritmetica" dei valori dati.

3. rispetto $||\mathbf{x}||_\infty$: la funzione da minimizzare è ora $\max\{|x - b_1|, |x - b_2|, |x - b_3|\}$. Poiché b_2 è compreso fra b_3 e b_1 questo è la stessa cosa di $\max\{|x - b_1|, |x - b_3|\}$. Da considerazioni elementari segue

   $$x = \frac{1}{2}(b_1 + b_3).$$