Lezione 26 (ore 52,53) - 25 novembre 2025, 16:30 - 18:30 F3

(3.16) Definizione (soluzione nel senso dei minimi quadrati di un sistema).

Siano $A \in R^{r \times c}$ con $r > c$, e $b \in R^r$. Un elemento $x^* \in R^c$ si chiama *soluzione del sistema* $A\,x = b$ *nel senso dei minimi quadrati* se x^* è punto di minimo assoluto della funzione $SQ: R^c \to R$ definita da:

$$SQ(x) = \| A\,x - b \|^2$$

Si osservi che: se $y \in R^c$ è una soluzione di $A\,x = b$ allora y è *anche* una soluzione di $A\,x = b$ nel senso dei minimi quadrati (come mai?) ma, salvo casi particolari, una soluzione di $A\,x = b$ nel senso dei minimi quadrati *non è* una soluzione di $A\,x = b$.

Vediamo *come si determinano* le soluzioni di $A\,x = b$ nel senso dei minimi quadrati.

(3.17) Osservazione (scomposizione ortogonale di un vettore).

Siano $A = (a_1,\dots,a_c) \in R^{r \times c}$ con $r > c$, e $b \in R^r$. Detta b_* la *proiezione ortogonale*[1] di b su $\mathrm{span}\{ a_1,\dots,a_c \} = C(A)$[2], e posto $b_\perp = b - b_*$ si ottiene la scomposizione ortogonale:

$$b = b_* + b_\perp$$

Si osservi che:

(1) Poiché $b_* \in C(A)$, esiste $y \in R^c$ tale che $b_* = A\,y$;

(2) Per definizione di proiezione ortogonale, la colonna $b_\perp = b - b_*$ è ortogonale a tutti gli elementi di $C(A)$.

(3.18) Osservazione.

Per determinare le soluzioni di $A\,x = b$ nel senso dei minimi quadrati si osservi che, utilizzando la scomposizione ortogonale di b introdotta nell'osservazione precedente, per ogni $x \in R^c$ si ha:

$$SQ(x) = \| A\,x - b \|^2 = \| A\,x - b_* + b_\perp \|^2 = \| A\,x - A\,y + b_\perp \|^2 = \| A\,(x - y) + b_\perp \|^2$$

Poiché $A\,(x - y) \in C(A)$ e $b_\perp$ è ortogonale a tutti gli elementi di $C(A)$, per il Teorema di Pitagora[3] si ha:

$$\| A\,(x - y) + b_\perp \|^2 = \| A\,(x - y) \|^2 + \| b_\perp \|^2$$

Allora:

- Per ogni $x \in R^c$ si ha: $SQ(x) = \| A\,(x - y) \|^2 + \| b_\perp \|^2 \geqslant \| b_\perp \|^2$

---

1 La proiezione ortogonale di $v \in R^n$ su un sottospazio $W \subset R^n$ è l'*unico* elemento $v_* \in W$ tale che la differenza $v - v_*$ è ortogonale a tutti gli elementi di W.

2 $C(A)$ si chiama anche *spazio delle colonne di* A e coincide con l'immagine dell'applicazione lineare da R^c in R^r definita da $x \to A\,x$.

3 Siano a,b elementi di R^n, e sia $\langle a,b\rangle = b^t a$ il prodotto scalare di a e b. Se a e b sono ortogonali (ovvero, se $\langle a,b\rangle = 0$) allora si ha:

$$\| a + b \|^2 = \langle a + b,a + b\rangle = \langle a,a\rangle + 2\,\langle a,b\rangle + \langle b,b\rangle = \langle a,a\rangle + \langle b,b\rangle = \| a \|^2 + \| b \|^2$$

- $SQ(x) = \| b_\perp \|^2 \Leftrightarrow \| A\,(x - y) \|^2 = 0 \Leftrightarrow A\,(x - y) = 0$, ovvero $x - y \in \ker A$.[4]

Dunque, l'insieme $S_{MQ}(A,b)$ delle soluzioni di $A\,x = b$ nel senso dei minimi quadrati è dato da:

$$S_{MQ}(A,b) = y + \ker A$$

(3.19) Osservazione (equazioni normali).

Per determinare tutte le colonne $y \in R^c$ tali che $b_* = A\,y$ si osservi che, per definizione di proiezione ortogonale su $C(A)$:

$y \in R^c$ è tale che $A\,y = b_* \Leftrightarrow b - b_* = b - A\,y$ è ortogonale a *tutti* gli elementi di $C(A)$

Ma: perché una colonna $v \in R^r$ sia ortogonale a tutti gli elementi di $C(A)$ è *necessario e sufficiente* che v sia ortogonale alle colonne di A (dimostrarlo!). Dunque: v ortogonale a tutti gli elementi di $C(A) \Leftrightarrow \langle v,a_1\rangle = a_1^t\,v = 0,\ldots,\langle v,a_c\rangle = a_c^t\,v = 0 \Leftrightarrow A^t\,v = 0$. Allora:

$$y \in R^c \text{ è tale che } A\,y = b_* \quad\Leftrightarrow\quad A^t\,(b - A\,y) = 0 \quad\Leftrightarrow\quad A^t A\,y = A^t\,b$$

Il sistema $A^t A\,x = A^t\,b$ si chiama *sistema delle equazioni normali* associato al sistema $A\,x = b$.

Si osservi che: $\ker A = \ker A^t A$ (infatti: $x \in \ker A \Rightarrow A\,x = 0 \Rightarrow A^t\,(A\,x) = 0 \Rightarrow A^t A\,x = 0 \Rightarrow x \in \ker A^t A$; viceversa: $x \in \ker A^t A \Rightarrow A^t A\,x = 0 \Rightarrow x^t\,(A^t A\,x) = 0 \Rightarrow (x^t A^t)(A\,x) = 0 \Rightarrow (A\,x)^t (A\,x) = 0 \Rightarrow \| A\,x \|^2 = 0 \Rightarrow A\,x = 0 \Rightarrow x \in \ker A$). Allora:

$$\boxed{S_{MQ}(A,b) = y + \ker A = y + \ker A^t A = \{\text{ soluzioni del sistema delle equazioni normali}\}}$$

Inoltre:

- La matrice $A^t A \in R^{c\times c}$ è *simmetrica e semidefinita positiva* (infatti, per ogni colonna $x \neq 0$ di R^c si ha: $x^t\,(A^t A)\,x = (x^t A^t)(A\,x) = (A\,x)^t\,(A\,x) = \| A\,x \|^2 \geqslant 0$) ed è *definita positiva* se e solo se le colonne di A sono *linearmente indipendenti* (dimostrarlo!).
- Le colonne di A sono *linearmente indipendenti* $\Leftrightarrow \ker A = \ker A^t A = \{0\} \Leftrightarrow$ il sistema delle equazioni normali ha *una sola soluzione* $\Leftrightarrow$ la matrice $A^t A$ *è invertibile*.
- Le colonne di A sono *linearmente dipendenti* $\Leftrightarrow \dim \ker A = \dim \ker A^t A > 0 \Leftrightarrow$ il sistema delle equazioni normali ha *infinite soluzioni* $\Leftrightarrow$ la matrice $A^t A$ *non è invertibile*.

4 Se $A \in R^{r\times c}$, si indica con $\ker A$ il sottospazio vettoriale di R^c delle soluzioni del sistema omogeneo $A\,z = 0$.