

RISOLUZIONE "AI MINIMI QUADRATI" DI UN SISTEMA LINEARE INCOMPATIBILE

(1) $AX=B$ sistema lineare di m equazioni in n incognite

$$A \text{ m. } m \times n \quad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$A_1, \dots, A_n$: le colonne di A $W := \text{span}(A_1, \dots, A_n) \subset \mathbb{R}^m$ sottosp.

$B \in \mathbb{R}^m$. Allora (1) si può interpretare come

(2) $x_1 A_1 + \dots + x_n A_n = B$ (a primo membro: una comb. lineare delle A_i)

(1) è compatibile $\Leftrightarrow B \in W$.

Viste le cose da questi punti di vista, è lecito assumere che $A_1, \dots, A_n$ siano linearmente indep. Dunque

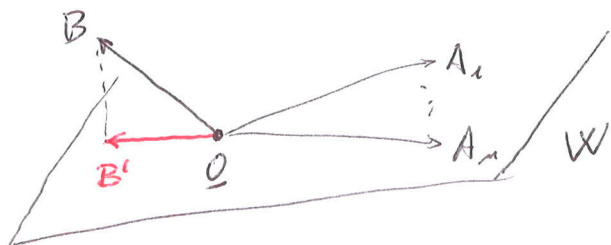
(3) $\text{rg}(A) = n \leq m$

SE $\text{rg}(A) = n = m$, allora A è invertibile $\Rightarrow$ (1) è compatibile.

SIA $\text{rg}(A) = n < m$ $\nexists$ sia anche $\text{rg}[A|B] = n+1$

Dunque (1) non è compatibile.

Come "produrre" qualcosa di sensato? Con gli occhi:



$$B \notin W$$

B' sia la proiezione ortogonale di B su W , cioè

$\vec{BB'} = B' - B \in W^\perp$ questo significa che B' è il punto di W più vicino possibile a B

Il sistema lineare (4) $AX=B'$ è sicuramente compatibile. La sua soluzione sarà unica perché $A_1, \dots, A_n$ è una base di W .

Il vettore $\vec{BB'}$ si chiama vettore d'errore ("errore" che si commette sostituendo $AX=B$ con $AX=B'$).

La scelta fatta per B' è quella che rende la norma $\|\vec{BB'}\|$ la minima possibile, al variare di B' in W . Cioè

$$\|\vec{BB'}\| = \min \{ \|\vec{BC}\| \mid C \in W \}$$

Riassumendo: possiamo risolvere il S.L. $AX=B'$, la soluzione sarà unica, e sarà tanto di meglio si riesce a trovare come "sostituto" di una soluzione di (1) (che non esiste). È un' approssimazione.

Da un punto di vista computazionale sarebbe bello riuscire a trovare la soluzione di $AX=B'$ senza aver effettivamente calcolato B' .

Questo è effettivamente possibile; vediamo come.

Sia $\bar{X} \in \mathbb{R}^n$ la soluzione di (4) $AX=B'$.

Dunque $B' = A\bar{X}$, quindi $\vec{BB'} = B' - B = A\bar{X} - B \in W^\perp$

Concretamente questo significa che, presa comunque una combinazione lineare delle colonne $A_1, \dots, A_n$ di A :

$\lambda_1 A_1 + \dots + \lambda_n A_n \quad \lambda_i \in \mathbb{R} \quad \forall i$, si deve avere

$$\langle \vec{BB'}, \underbrace{\lambda_1 A_1 + \dots + \lambda_n A_n}_{\text{è un unico vettore di } \mathbb{R}^m} \rangle_{st} = 0$$

Se poniamo $C = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}$ ("C" = "coefficienti"), allora

$\lambda_1 A_1 + \dots + \lambda_n A_n = AC$ e la relazione sopra si può riscrivere

$$0 = \langle AC, \vec{BB'} \rangle_{st} = \langle AC, A\bar{X} - B \rangle_{st} = {}^t(AC)(A\bar{X} - B) = \underbrace{{}^t C}_{1 \times n} \underbrace{[{}^t A A \bar{X} - {}^t A B]}_{n \times 1}$$

e questo deve valere $\forall C \in \mathbb{R}^n$. In particolare, prendendo

$C = {}^tAA\bar{X} - {}^tAB$, essendo $\langle \cdot, \cdot \rangle_{st}$ def. positivo si conclude che deve essere

${}^tAA\bar{X} - {}^tAB = 0$ cioè $\bar{X}$ è soluz. di "EQUAZ. NORMALI" ${}^tAA\bar{X} = {}^tAB$ (5)

Mostriamo tra un momento che (5) è compatibile, e che ha un'unica soluzione.

Se chiamiamo $\bar{X}$ una sol. di (5), ripercorrendo a ritroso i passaggi fatti sopra, si vede che $A\bar{X} - B \in W^\perp$, cioè che $A\bar{X} = B$.

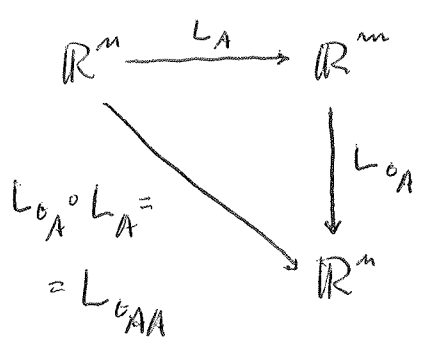
Quindi quello che dobbiamo fare è risolvere il S.L. (5).
A partire da (1) si scrive subit (5) COMMENTARE

Proveremo ora:

$rg(A) = n \implies rg({}^tAA) = n \xrightarrow{{}^tA \text{ è } n \times m} {}^tAA \text{ è invertibile} \implies$

(5) è compatibile, ed ha un'unica soluzione.

Consideriamo le applicazioni lineari



$U \in \text{Ker}(L_A) \iff L_A(U) = 0 \implies L_{tA}(L_A(U)) = 0$

cioè $L_{tAA}(U) = 0 \implies U \in \text{Ker}(L_{tAA})$

$\text{Ker}(L_A) \subset \text{Ker}(L_{tAA})$

Viceversa, sia $U' \in \text{Ker}(L_{tAA})$ cioè $\underbrace{{}^tAA}_{n \times 1} U' = 0 \in \mathbb{R}^n$

Quindi, per ogni $Y \in \mathbb{R}^n$ si ha ${}^tY {}^tAAU' = 0 \in \mathbb{R}$
ovvero $(AY)(AU') = 0, \forall Y \in \mathbb{R}^n$. Ovvero $\langle AY, AU' \rangle_{st} = 0$

$\forall Y \in \mathbb{R}^n$ (NB " $\langle -, - \rangle_{st}$ " è il pr. scalare standard in $\mathbb{R}^m$)

In particolare, prendendo $Y = U^1$ si ha $\langle AU^1, AU^1 \rangle_{st} = 0$
da cui $AU^1 = 0$ essendo $\langle -, - \rangle_{st}$ def. positivo.

Si conclude: $U^1 \in \text{Ker}(L_A)$.

Abbiamo così verificato che $\text{Ker}(L_A) = \text{Ker}(L_{AA})$.

Ne segue

$$\begin{aligned} n = \text{rg}(A) &= \dim(\text{Im}(L_A)) = n - \dim(\text{Ker}(L_A)) = n - \dim(\text{Ker}(L_{AA})) = \\ &= \dim(\text{Im}(L_{AA})) = \text{rg}(A^t A) \quad \underline{\text{rg}(A^t A) = n} \end{aligned}$$

Perché il nome "AI MINIMI QUADRATI"?

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n \text{ qualsiasi} \quad B^u = AX \in W$$

$$\|\vec{BB}^u\|^2 = (a_{11}x_1 + \dots + a_{1n}x_n - b_1)^2 + \dots + (a_{m1}x_1 + \dots + a_{mn}x_n - b_m)^2$$

Si tratta di trovare le "x" in modo che questa somma di quadrati sia la minima possibile.

È possibile trovare il SL (5) anche mediante l'Analisi.

Ecco come.

Penso il secondo membro dell'espressione qui sopra come una funzione nelle variabili $x_1, \dots, x_n$. Allora

$$\frac{\partial \|\vec{BB}^u\|^2}{\partial x_j} = \sum_{i=1}^m \uparrow (a_{i1}x_1 + \dots + a_{in}x_n - b_i) a_{ij} \uparrow = 0 \quad \forall j = 1, \dots, n$$

(WIPONKO)

Cioè

$$\underbrace{{}^t(AX - B)}_{1 \times m} \underbrace{A}_{m \times n} = 0 \Rightarrow [{}^t(AX) - {}^tB]A = 0 \Rightarrow {}^t(AX)A = {}^tBA$$

$$\Rightarrow \text{transpongo} \quad {}^tAAX = {}^tAB \quad \text{è la (5)!}$$

ISOMETRIE

($B = (v_1, \dots, v_m)$ base ON fix 30/11/16

(5)

$\mathbb{R}^m$; $\langle -, - \rangle$ prod. scalare in $\mathbb{R}^m$, fissato; $\forall \mathbb{E}^m$

$A, B \in \mathbb{R}^m$ $d(A, B) \stackrel{\text{def}}{=} \|\vec{AB}\| = \|B - A\|$ distanza tra A e B

1) $d(A, B) \geq 0$ per ogni $A, B \in \mathbb{E}^m$

$$d(A, B) = 0 \iff A = B$$

2) $d(A, B) = d(B, A)$

$$d(B, A) = \|A - B\| = \|(-1)(B - A)\| = |-1| \cdot \|B - A\| = 1 \cdot d(A, B)$$

3) DISUGUAGLIANZA TRIANGOLARE

$$d(A, B) \leq d(A, C) + d(C, B) \quad \text{per ogni } A, B, C \in \mathbb{E}^m.$$

Posso pensare d come un'applicazione $d: \mathbb{E}^m \times \mathbb{E}^m \rightarrow \mathbb{R}$

Si dice che $(\mathbb{E}^m, d)$ è uno "spazio metrico".

In questa situazione è importante considerare le applicazioni $f: \mathbb{E}^m \rightarrow \mathbb{E}^m$ che "conservano le distanze":

$$d(f(A), f(B)) = d(A, B) \quad \text{per ogni } A, B \in \mathbb{E}^m$$

ISOMETRIE

Ogni isometria è iniettiva

Dim. Siano $A, B \in \mathbb{E}^m$ con $A \neq B$. Allora $d(A, B) > 0$. Ma

$$d(f(A), f(B)) = d(A, B) > 0. \quad \text{Dunque } f(A) \neq f(B) \quad \blacksquare$$

Se $f, g: \mathbb{E}^m \rightarrow \mathbb{E}^m$ sono isometrie, allora anche $g \circ f$ è un'isometria $\mathbb{E}^m \rightarrow \mathbb{E}^m$.

~ ~ ~

Per capire come sono fatte le isometrie abbiamo bisogno di una premessa.

Sia $Q \in \mathbb{E}^m$ un punto fissato. Lo posso pensare anche come vettore.

$$Q = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix} \quad (\text{"vettore", } Q(q_1, \dots, q_m) \text{ "punto"})$$

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} \in \mathbb{R}^m \text{ vettore} \quad Q+u = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} = \begin{pmatrix} q_1+u_1 \\ \vdots \\ q_m+u_m \end{pmatrix} \quad \text{lo penso come punto.}$$

Alora: preso comunque il punto $P(p_1, \dots, p_m) \in \mathbb{E}^m$ (ed essendo fissato Q come sopra) esiste uno ed un solo vettore $u \in \mathbb{R}^m$ tale che $Q+u=P$

Basta prendere $u = P - Q = \begin{pmatrix} p_1 - q_1 \\ \vdots \\ p_m - q_m \end{pmatrix}$ scriviamo anche $u = \overrightarrow{QP}$

Si può dimostrare il

la notazione suggerisce che f_* dipende da f

TEOREMA

Sia $f: \mathbb{E}^m \rightarrow \mathbb{E}^m$ un'isometria. Allora esiste un' applicazione lineare $f_*: \mathbb{R}^m \rightarrow \mathbb{R}^m$ tale che, fissato comunque $Q \in \mathbb{E}^m$ si ha

$$\boxed{f(P) = f(Q) + f_*(\overrightarrow{QP})} \quad \forall P \in \mathbb{E}^m$$

È questo è un fatto qualitativo importantissimo.

Verifichiamo che conseguenza ha per l'applicazione lineare f_* il fatto che f sia un'isometria.

Sia $u \in \mathbb{R}^m$ arbitrario; poniamo $P = Q + u$. Allora

$$\langle u, u \rangle = \langle P - Q, P - Q \rangle = \|P - Q\|^2 = \|f(P) - f(Q)\|^2 = \langle f_*(u), f_*(u) \rangle$$

(f isometria)

$$\langle f_*(u), f_*(u) \rangle = \langle u, u \rangle \quad \forall u \in \mathbb{R}^m$$

Si può fare ancora meglio usando la formula di polarizzazione:

$$\langle u+v, u+v \rangle = \langle u, u \rangle + 2\langle u, v \rangle + \langle v, v \rangle \Rightarrow$$

$$\boxed{\langle u, v \rangle = \frac{1}{2} \left(\underbrace{\langle u+v, u+v \rangle - \langle u, u \rangle - \langle v, v \rangle}_u \right)} \quad \text{f.o.b.p.}$$

$$\langle f_*(u), f_*(v) \rangle = \frac{1}{2} \left(\langle f_*(u) + f_*(v), f_*(u) + f_*(v) \rangle - \langle f_*(u), f_*(u) \rangle - \langle f_*(v), f_*(v) \rangle \right)$$

Riassumendo

$$\boxed{\forall u, v \in \mathbb{R}^n \quad \langle f_*(u), f_*(v) \rangle = \langle u, v \rangle} \quad \text{f}_* \text{ conserva il prod. scalare}$$

Che cosa significa questo in termini della matrice

$$A = M_B^B(f_*) \quad \text{Ricordiamo che } B = (v_1, \dots, v_n) \text{ è base ON di } \mathbb{R}^n \text{ risp. } \langle \cdot, \cdot \rangle.$$

$$A = \begin{pmatrix} | & & | \\ f(v_1) & & f(v_n) \\ | & & | \end{pmatrix}$$

$$f(v_1) \quad f(v_2) \quad \dots \quad f(v_n)$$

Le colonne di A sono ancora versori, a due a due $\perp$ tra loro:

le colonne di A formano una base ON risp. $\langle \cdot, \cdot \rangle$.

$$\text{Cioè } \|f(v_i)\| = 1 \quad \forall i = 1, \dots, n \quad \text{e} \quad \langle f(v_i), f(v_j) \rangle = 0 \quad \text{se } i \neq j$$

Tutti questi prodotti scalari si possono scrivere in un modo estremamente condensato e utile così:

$$\boxed{A^t A = I_n} \Rightarrow A^t A A = I_n \Rightarrow \text{anche le righe di } A \text{ formano una base ON di } \mathbb{R}^n.$$

$$\Rightarrow A \text{ è invertibile e } \boxed{A^{-1} = A^t}$$