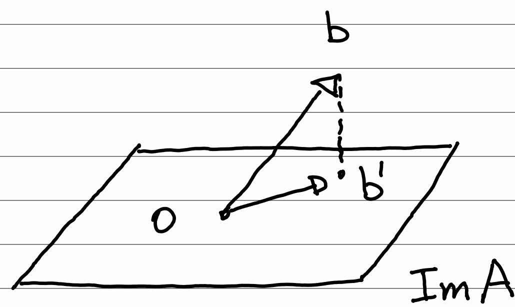
Approssimazione lineare: sistemi non risolubili

Supponiamo di avere un sistema lineare
$$AX = b \qquad (A \in Mat_{m \times n})$$
di m equazioni in n incognite.
Questo sistema è risolubile se e solo se $b \in Col(A) = \mathrm{Im}\,A \subset \mathbb{R}^m$.
Spesso, però, il sistema non è risolubile ma possiamo comunque chiederci: "qual'è il vettore b' più vicino a b per il quale $AX = b'$ è risolubile?"



La risposta è $b' = pr_{\mathrm{Im}A}(b)$.

Vediamo come trovare b':

$b' \in \operatorname{Im} A \Rightarrow b' = AZ$ per qualche $Z \in \mathbb{R}^n$

Quindi la domanda è: quali sono i vettori $Z \in \mathbb{R}^n$ t.c. $AZ = pr_{\operatorname{Im} A}(b)$?

Per definizione

$$b - AZ \in (\operatorname{Im} A)^{\perp} = \operatorname{Ker}(A^t)$$

$$\Rightarrow A^t(b - AZ) = 0_{\mathbb{R}^n}$$

$$\Rightarrow \boxed{A^t A Z = A^t b} \quad (*)$$

sistema di eq. lineari NORMALI.

Le soluzioni Z di questo sistema si chiamano le <u>soluzioni approssimate nel senso dei minimi quadrati</u> (ovvero rispetto al prodotto scalare standard di $\mathbb{R}^m$) del sistema $AX = b$.

<u>OSS</u>: Se $b \in \operatorname{Im} A$, allora $b' = b$ e le soluzioni approssimate sono soluzioni esatte.

· Studiamo quindi sistemi di equazioni normali:

$$\boxed{A^t A Z = A^t b}$$

Intanto, $A^t A \in Mat_{n \times n}$,

$\cdot)$ $\boxed{\operatorname{Ker}(A^t A) = \operatorname{Ker} A}$:

Gio. 13.12

Infatti,

$X \in \operatorname{Ker} A \Rightarrow A^t A X = 0_{\mathbb{R}^n} \Rightarrow X \in \operatorname{Ker} A^t A$,

$X \in \operatorname{Ker} A^t A \Rightarrow A^t A X = 0_{\mathbb{R}^n} \Rightarrow X^t A^t A X = 0_{\mathbb{R}}$

$\Rightarrow (AX) \cdot (AX) = 0 \Rightarrow AX = 0 \Rightarrow X \in \operatorname{Ker} A$

Quindi:

$$\boxed{rg(A^t A) = rg A}$$

In particolare il sistema $AX = b$ ammette un'unica soluzione approssimata se e solo se $rg A = n$.

Polinomi approssimanti

Siano

$$(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$$

$(n+1)$ dati sperimentali aventi ascisse distinte $x_i \neq x_j \quad \forall i \neq j$.

Sappiamo che esiste un unico polinomio di grado n che interpola i dati.

Se n è grande determinare questo polinomio non è molto utile.

I dati possono comunque essere descritti da un polinomio di grado $m < n$ che meglio li approssima nel senso dei minimi quadrati: cerchiamo un polinomio di grado m

$$p(x) = a_0 + a_1 x + \dots + a_m x^m$$

t.c. $p(x_i) \sim y_i$ dove $\sim$ vuol dire vicino nel seguente senso:

ponendo $\bar{y}_i = p(x_i)$

consideriamo i due vettori

$$\bar{Y} = \begin{pmatrix} \bar{y}_0 \\ \vdots \\ \bar{y}_n \end{pmatrix}, \quad Y = \begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^{n+1}$$

vogliamo che $\|Y - \bar{Y}\|$ sia minima:

Posto

$$A = Van_m(x_0, \dots, x_n) = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^m \\ 1 & x_1 & x_1^2 & \cdots & x_1^m \\ \vdots & & & & \\ 1 & x_n & x_n^2 & \cdots & x_n^m \end{pmatrix}$$

vogliamo che

$$\bar{Y} = pr_{\operatorname{Im} A} Y$$

<u>OSS</u>: Per il Teorema degli orlati $rg A = m+1$ quindi $\exists !$ soluzione approssimata di $AX = Y$ e quindi un unico polinomio approssimante di grado m. Si noti che affinché P_m sia unico, è sufficiente che solo $m+1$ ascisse siano distinte!

Sia $\begin{pmatrix} a_0 \\ \vdots \\ a_m \end{pmatrix}$ l' unica soluzione approssimata di $AX=Y$. Allora il polinomio $p_m(x) = a_0 + a_1 x + \dots + a_m x^m$ si chiama il <u>polinomio approssimante di grado m</u> dei dati sperimentali.

Se $m=1$, la retta $y = p_1(x) = a_0 + a_1 x$ si chiama la <u>retta di regressione</u> dei dati.

<u>Es</u>: Trovare i polinomi approssimanti di grado 1 e 2 dei seguenti dati

$(-1,1), (0,0), (2,3), (3,4)$.

<u>Sol.</u>:

$$A = Van_1(-1,0,2,3) = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}, \quad Y = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix}$$

Dobbiamo risolvere

$$A^t A Z = A^t Y$$

$$A^t A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 4 \\ 4 & 14 \end{pmatrix}$$

$$A^tY = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ 17 \end{pmatrix}$$

$$\det A^tA = \det \begin{pmatrix} 4 & 4 \\ 4 & 14 \end{pmatrix} = 4 \det \begin{pmatrix} 1 & 1 \\ 4 & 14 \end{pmatrix}$$

$$= 8 \det \begin{pmatrix} 1 & 1 \\ 2 & 7 \end{pmatrix} = 40$$

$$(A^tA)^{-1} = \frac{1}{40} \begin{pmatrix} 14 & -4 \\ -4 & 4 \end{pmatrix} = \frac{1}{20} \begin{pmatrix} 7 & -2 \\ -2 & 2 \end{pmatrix}$$

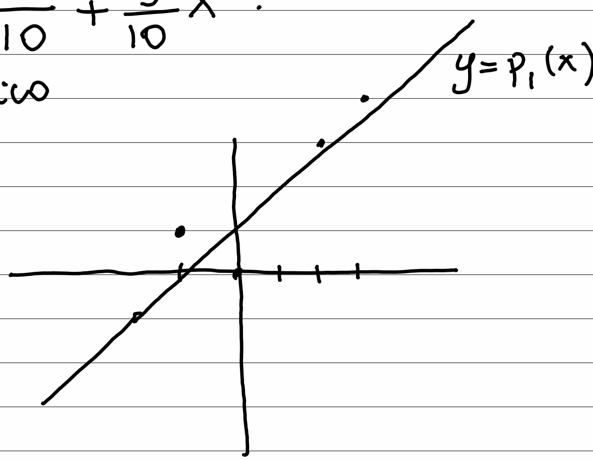
$$(A^tA)^{-1} A^tY = \frac{1}{20} \begin{pmatrix} 7 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 8 \\ 17 \end{pmatrix} =$$

$$= \frac{1}{20} \begin{pmatrix} 22 \\ 18 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 11 \\ 9 \end{pmatrix}$$

Il polinomio approssimante di grado 1 è

$$p_1(x) = \frac{11}{10} + \frac{9}{10}x .$$

Il grafico



Per trovare P_2:

$$A = \mathrm{Van}_2(-1, 0, 2, 3) = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$$

$$B = A^t A = \begin{pmatrix} 4 & 4 & 14 \\ 4 & 14 & 34 \\ 14 & 34 & 98 \end{pmatrix}$$

$$b = A^t Y = \begin{pmatrix} 8 \\ 17 \\ 49 \end{pmatrix}$$

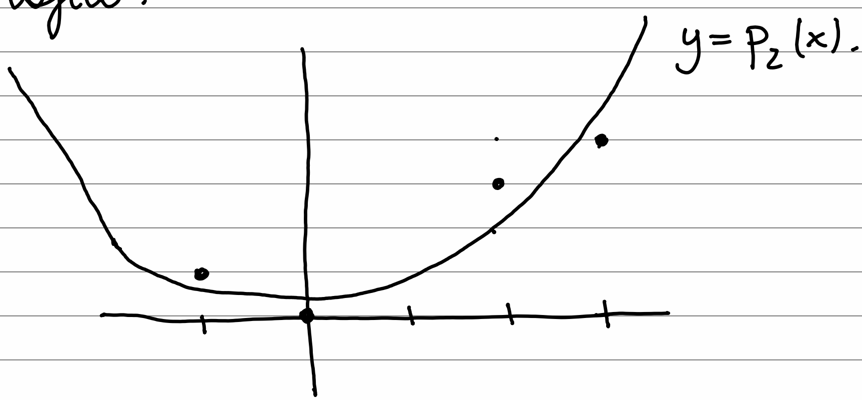
L'unica soluzione di $BZ = b$ è

$$Z = \begin{pmatrix} 3/5 \\ 7/30 \\ 1/3 \end{pmatrix}$$

Il polinomio approssimante di grado 2 è

$$\frac{3}{5} + \frac{7}{30}x + \frac{1}{3}x^2$$

Il grafico:



Es: Trovare le soluzioni approssimate del sistema

$$\begin{cases} x-y=3 \\ 2x+y=-1 \\ x+5y=-4 \end{cases}$$

$$(QR)^t QR z = (QR)^t b$$
$$R^t R z = R^t Q^t b$$

Sol.: $A=\begin{pmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 5 \end{pmatrix}$ $\quad b=\begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix}$

$B=A^tA=\begin{pmatrix} 6 & 6 \\ 6 & 27 \end{pmatrix}$ $\quad \det B = 6\cdot 21 \neq 0$

$A^tb=\begin{pmatrix} -3 \\ -24 \end{pmatrix}$. L'unica soluzione approssimata é

$$Z = B^{-1}\begin{pmatrix} -3 \\ -24 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1 \end{pmatrix}$$

Troviamo la dec. QR di A:

$Q_0=\begin{pmatrix} 1 & -2 \\ 2 & -1 \\ 1 & 4 \end{pmatrix}$ $\quad R_0=\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

$F_1=\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$; $\quad F_2 = A^2 - \frac{A^2\cdot F_1}{F_1\cdot F_1}F_1 = \begin{pmatrix} -1 \\ 1 \\ 5 \end{pmatrix} - \frac{6}{6}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 4 \end{pmatrix}$

$Q=\begin{pmatrix} 1/\sqrt{6} & -2/\sqrt{21} \\ 2/\sqrt{6} & -1/\sqrt{21} \\ 1/\sqrt{6} & 4/\sqrt{21} \end{pmatrix}$ $\quad R=\begin{pmatrix} \sqrt{6} & \sqrt{6} \\ 0 & \sqrt{21} \end{pmatrix}$

$RZ=Q^tb$: $\quad Q^tb=\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{21}} & -\frac{1}{\sqrt{21}} & \frac{4}{\sqrt{21}} \end{pmatrix}\begin{pmatrix} 3 \\ -1 \\ -4 \end{pmatrix} = \begin{pmatrix} -3/\sqrt{6} \\ -21/\sqrt{21} \end{pmatrix}$

$$\left(\begin{array}{cc|c} \sqrt{6} & \sqrt{6} & -3/\sqrt{6} \\ 0 & \sqrt{21} & -21/\sqrt{21} \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 1 & -1/2 \\ 0 & 1 & -1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & 0 & 1/2 \\ 0 & 1 & -1 \end{array}\right) \checkmark$$

Dato uno sp. metrico (V, s) ed un sottospazio vettoriale $U \subseteq V$ possiamo pensare la proiezione ortogonale su U come un endomorfismo

$$pr_U : V = U \oplus U^{\perp} \longrightarrow V$$
$$v = v_1 + v_2 \longmapsto v_1$$

lineare di V. Si noti che

$$\operatorname{Ker} pr_U = U^{\perp}, \quad \operatorname{Im} pr_U = U.$$

Come sono fatte le matrici che rappresentano pr_U ?

Se scegliamo una base ortogonale

$$\mathcal{B} = \{F_1, \dots, F_K, F_{K+1}, \dots, F_n\}$$

di V t.c.

$$U = \langle F_1, \dots, F_K \rangle$$

e quindi

$$U^{\perp} = \langle F_{K+1}, \dots, F_n \rangle$$

come nel Teorema di decomposizione

ortogonale, allora la matrice che rappresenta pr_U in questa base é

$$\begin{pmatrix} \mathbb{1}_K & 0 \\ 0 & 0 \end{pmatrix} \in Mat_{n\times n}$$

Infatti,

$$pr_U(F_i) = F_i \quad \forall i=1,\dots,K$$

$$pr_U(F_j) = 0_V \quad \forall j=K+1,\dots,n.$$

Se non vogliamo ortogonalizzare: sia $\{v_1,\dots,v_K\}$ una base di U, estendiamola ad una base

$$\mathcal{B} = \{v_1,\dots,v_K, v_{K+1},\dots,v_n\}$$

di V. La matrice che rappresenta pr_U nella base $\mathcal{B}$ è della forma

$$\begin{pmatrix} \mathbb{1}_K & B \\ 0 & C \end{pmatrix}$$

In $\mathbb{R}^n$:

$$U = \langle v_1, \dots, v_k \rangle \subseteq \mathbb{R}^n$$

$$A = (v_1 | \dots | v_k) \in Mat_{n \times k}$$

Allora dato $\cancel{x} \in \mathbb{R}^n$

$$pr_U(X) = y_1 v_1 + \dots + y_k v_k = AY$$

chi è Y?

$$X - pr_U(X) \in U^\perp = Col(A)^\perp = \operatorname{Ker} A^t$$

$$\Rightarrow \quad A^t X = A^t pr_U(X)$$

$$\Rightarrow \quad A^t X = A^t A Y$$

$rg\, A = k \Rightarrow A^t A$ è invertibile

$$\Rightarrow \quad Y = (A^t A)^{-1} A^t X.$$

Quindi

$$\boxed{pr_U(X) = A (A^t A)^{-1} A^t X}$$

Se $\{v_1, \dots, v_k\}$ è ortonormale, allora $A^t A = \mathbb{1}_k$ e

$$\boxed{pr_U(X) = A A^t X}$$

.. Quindi, in $\mathbb{R}^n$, la matrice P che rappresenta pr_U nella base canonica è

(*) $P = A(A^tA)^{-1}A^t \in Mat_{n\times n}$

dove $A = (v_1|\dots|v_k)$ e $\{v_1,\dots,v_k\}$ è una base di U.

Se $\{F_1,\dots,F_k\}$ è una base ortonormale di U (in $(\mathbb{R}^n,\cdot)$), allora posto

$$B = (F_1|\cdots|F_k)$$

si ha $B^tB = \mathbb{1}_k$ e quindi

$$P = B(B^tB)^{-1}B^t = BB^t$$

Si noti che:

1) $P^2 = BB^tBB^t = BB^t = P$

[P è <u>idempotente</u>]

2) $P^t = (BB^t)^t = BB^t = P$.

[P è <u>simmetrica</u>]

Vale anche il viceversa:

<u>Prop.</u>: Una matrice $P \in Mat_{n \times n}$ rappresenta una proiezione ortogonale nella base standard di $\mathbb{R}^n$ se e solo se

1) $P^2 = P$
2) $P^t = P$

<u>dim</u>: Se P rappresenta una proiezione ortogonale allora è idempotente e simmetrica per ciò che abbiamo discusso prima.

Supponiamo che P sia idempotente e simmetrica. Dimostriamo che P è una proiezione ortogonale su $\mathrm{Im}P$.

Consideriamo la dec. ortogonale di $(\mathbb{R}^n, \cdot)$:

$$\mathbb{R}^n = \mathrm{Im}\,P \oplus \mathrm{Ker}\,P^t$$

$$P^t = P \to \; = \mathrm{Im}\,P \oplus \mathrm{Ker}\,P$$

Sia $X \in \mathbb{R}^n$, allora

$$X = Y + Z$$

dove $Y = pr_{\mathrm{Im}P}(X)$ e $Z = pr_{\mathrm{Ker}P}(X)$.

Quindi, $Y = PT$. Chi è T?

$$PX = PY + PZ = P^2T \overset{P^2=P}{=} PT = Y$$

Ne segue che

$$pr_{\mathrm{Im}P}(X) = PX$$

$\blacksquare$

Es: Sia

$$U : \begin{cases} 2x+y-z=0 \\ x-y+z=0 \end{cases}$$

Trovare la matrice P che rappresenta $P_U : \mathbb{R}^3 \to \mathbb{R}^3$ nella base standard

Sol.: Troviamo una base di U, ovvero troviamo le soluzioni base del sistema:

$$U = \operatorname{Ker}\begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} = \operatorname{Ker}\begin{pmatrix} 1 & -1 & 1 \\ 2 & 1 & -1 \end{pmatrix} =$$

$$= \operatorname{Ker}\begin{pmatrix} 1 & -1 & 1 \\ 0 & 3 & -3 \end{pmatrix} = \operatorname{Ker}\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \operatorname{Ker}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

$$= \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle$$

Una base ortonormale di U è

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

poniamo

$$B = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

allora

$$P = BB^t = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

$\boxtimes$