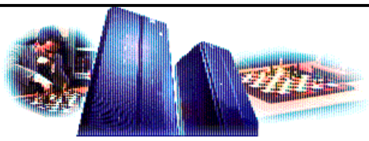
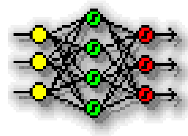




Sistemi Intelligenti
Stimatori e identificazione - II

Alberto Borghese

Università degli Studi di Milano
Laboratory of Applied Intelligent Systems (AIS-Lab)
Dipartimento di Informatica
borghese@di.unimi.it

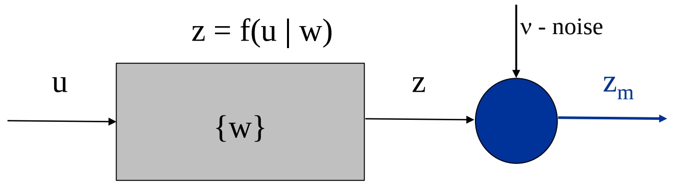
Overview

Modelli

Sistemi lineari

Soluzione ai minimi quadrati

Modello



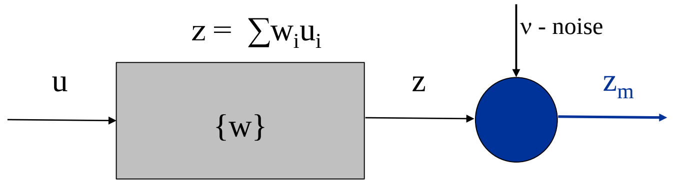
u – causa => z_m – effetto (misurato on errore)

Control / Classification / Prediction: determine {z} from {u},{w}

Inverse problem: determine cause {u} from {z_m},{w}

Inverse problem: Identification: determine {w} from {u},{z_m} - *Learning*

Modello lineare



u – causa => z_m – effetto (misurato on errore)

Control / Classification / Prediction: determine {z} from {u},{w}

Inverse problem: determine cause {u} from {z_m},{w}

Inverse problem: Identification: determine {w} from {u},{z_m} - *Learning*

Overview

Modelli

Sistemi lineari

Soluzione ai minimi quadrati

Sistema lineare

$a_{11}x_1 + a_{12} x_2 + ... \quad a_{1N} x_N = b_1$

$a_{21}x_1 + a_{22} x_2 + ... \quad a_{2N} x_N = b_2$

....

$a_{M1}x_1 + a_{M2} x_2 + ... \quad a_{MN} x_N = b_M$

$\{a_{ij}\}$ – coefficienti in numero N x M

$\{x_j\}$ – incognite, N

$\{b_j\}$ – termini noti, M

I sistemi lineari sono interessanti perchè sono manipolabili con operazioni semplici (algebra delle matrici)

NB le x qui sono i parametri w del modello, i coefficienti a_{ij} i valori di ingresso e b_i i valori di uscita.

Esempio:

$3x_1 + 2 x_2 + ... \quad 4x_N = 5$

$4x_1 -2 x_2 + ... \quad 0.5 x_N = 3$

....

$2x_1 + 3 x_2 + ... \quad -3 x_N = -1$

Matrici

$$A = [a_{i,j}] \qquad A^T = [a_{j,i}]$$

$$\alpha A = [\alpha a_{i,j}] \qquad C = A + B = [a_{i,j} + b_{i,j}]$$

$$C = AB = [c_{i,j}] \ \text{ dove } \ [c_{i,j}] = \sum_{k=1}^{n} a_{i,k} b_{k,j}$$

Prodotto degli elementi di una riga per gli elementi di una colonna.

Se A (n x m) ➔ B (m x p) ➔ C (n x p)

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & -4 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -1 \\ 1 & 3 \\ 2 & 0 \end{bmatrix} \quad ==> \quad C = \begin{bmatrix} 7 & 7 \\ -3 & -13 \end{bmatrix}$$

Se il numero di righe = numero di colonne, matrice quadrata

Matrici (Proprietà)

La somma è associativa e commutativa (A + B) + C = A + (B + C).

Il prodotto è associativo rispetto alla somma ma non gode della proprietà commutativa:

(A+B)C = AC + BC.

AB ≠ BA

$$I = [a_{i,j}] = \begin{cases} 1 & \text{per } i = j \\ 0 & \text{altrimenti} \end{cases} \quad \text{matrice identità}$$

$$AI = A = IA$$

$$\text{vettore come matrice colonna}: \bar{u}^T = \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix}$$

$$\text{prodotto vettore matrice}: \bar{v} = \bar{u}^T M$$

Minore complementare

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

A_{ij}* minore complementare di a_{ij} = determinante della matrice ottenuta eliminando la riga i e la colonna j di A.

$$A^*_{21} = \det \begin{bmatrix} 3 & -2 \\ 1 & 2 \end{bmatrix} = 3*2 - (-2*1) = +8$$

Determinante di una matrice Quadrata

$$\boxed{\det(A) = \sum_i (-1)^{(i+j)} a_{ij} A^*_{ij} = \sum_j (-1)^{(i+j)} a_{ij} A^*_{ij}}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \det(A) = a_{11}\, a_{22} - a_{12}\, a_{21}$$

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \longleftarrow$$ Elementi sulla riga

$\det(A) = (-1)^{(2+1)}$ (2) $[(3 * 2) - (-2 * 1)] + (-1)^{(2+2)}$ (0) $[(1*2) - (-2*1)] +$
$(-1)^{(2+3)}$ (1) $[(1*1)-(3*1)] = -16 + 2 = -14$

Calcolo della matrice inversa

Matrice dei complementi algebrici

$$A^{-1} = [1/\det(A)]\, A+$$

$$A^+_{ij} = (-1)^{i+j} \det (A^*_{ji})$$

Minore complementare della matrice A trasposta

Esempio di matrice Inversa

A = $[a_{i,j}]$, **matrice quadrata.**

$A+_{ij}$ matrice dei complementi algebrici = minori complementari moltiplicati $(-1)^{i+j}$

$$A^{-1} = 1/\det(A) \begin{bmatrix} A^+_{11} & A^+_{21} & A^+_{n1} \\ A^+_{12} & A^+_{22} & A^+_{32} \\ A^+_{13} & A^+_{23} & A^+_{33} \end{bmatrix}$$

$$A^{-1} A = I$$

$$A = \begin{bmatrix} 1 & 3 & -2 \\ 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad A^{-1}\ \frac{-1}{14} \begin{bmatrix} 0 \cdot 2 - 1 \cdot 1 & -[3 \cdot 2 - (-2) \cdot 1] & 3 \cdot 1 - (-2) \cdot 0 \\ -[2 \cdot 2 - 1 \cdot 1] & 1 \cdot 2 - (-2) \cdot 1 & -[1 \cdot 1 - (-2) \cdot 2] \\ 2 \cdot 1 - 0 \cdot 1 & -[1 \cdot 1 - 3 \cdot 1] & 1 \cdot 0 - 3 \cdot 2 \end{bmatrix} = (-1/14) \begin{bmatrix} -1 & -8 & 3 \\ -3 & 4 & -5 \\ 2 & 2 & -6 \end{bmatrix}$$

$$\det A = -14 \qquad AA^{-1} = -1/14 \begin{bmatrix} 1 \cdot (-1) + 3(-3) - 2 \cdot 2 & 1(-8) + 3 \cdot 4 - 2(2) & 1 \cdot 3 + 3(-5) - 2(-6) \\ 2(-1) + 0(-3) + 1(2) & 2(-8) + 0 \cdot 4 + 1 \cdot 2 & 2 \cdot 3 + 0(-5) + 1(-6) \\ 1(-1) + 1(-3) + 2 \cdot 2 & 1(-8) + 1 \cdot 4 + 2 \cdot 2] & 1 \cdot 3 + 1(-5) + 2(-6) \end{bmatrix} = I$$

Se esiste, la matrice inversa è unica.

Rango di una matrice

Data una matrice A di ordine n (n x n),

una matrice A n x n ha rango m < n se e solo se esiste un suo minore di ordine m non nullo mentre sono nulli tutti i minori di ordine m + 1.

Una matrice A n x n ha rango n (rango pieno) se e solo se il suo determinante è diverso da 0

Rango di una matrice M x N è la dimensione massima di tutte le matrici quadrate estraibili da A e con determinante non nullo. Il rango è massimo quando non è inferiore alla dimensione minima della matrice.

Matrice inversa

$$A^{-1}A = I$$

La matrice inversa è definita per una matrice quadrata

Esiste ed è unica se $\det(A) \neq 0$

Numero di condizionamento di una matrice (quadrata): rapporto tra il valore singolare maggiore e minore (cf. Funzione cond in Matlab).
E' una misura di sensibilità della soluzione di un sistema lineare a variazioni nei dati.

Rango di una matrice

Data una matrice A di ordine n (n x n),

una matrice A n x n ha rango m < n se e solo se esiste un suo minore di ordine m non nullo mentre sono nulli tutti i minori di ordine m + 1.

Una matrice A n x n ha rango n (rango pieno) se e solo se il suo determinante è diverso da 0

Rango di una matrice M x N è la dimensione massima di tutte le matrici quadrate estraibili da A e con determinante non nullo. Il rango è massimo quando non è inferiore alla dimensione minima della matrice.

Altre proprietà delle matrici

det(AB) = det(A) det(B)

det(diag(W)) = $\prod_k w_{k,k}$

$(A^T)^{-1} = (A^{-1})^T$

$(A\ B\ C)^T = C^T\ B^T\ A^T$

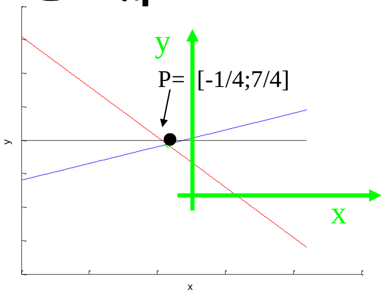
Una matrice U, si dice ortogonale se $U^T U = \mathrm{diag}(W)$.
Una matrice U, si dice ortonormale se $U^T U = I \Rightarrow U^{-1} = U^T$

Condizione di ortonormalità:
Il determinante è = 1.
La somma dei prodotti di due righe o di due colonne è = 0.
La somma dei quadrati degli elementi su righe e colonne = 1
Esempio notevole: **matrice di rotazione (cambio di sistema di riferimento)**.

Esempio

$y = x + 2$

$y = -3x + 1$



$1\ x_1 - 1\ x_2 = -2$

$-3\ x_1 - 1\ x_2 = -1$

$y = x_2$

$x = x_1$

Risolvo per sostituzione: $x_1 = -2 + x_2$.

$-3(-2 + x_2) - x_2 = -1 \Rightarrow x_2 = 7/4$

$x_1 - 1/4 = 2 \Rightarrow x_1 = -1/4$

Sistema lineare

$a_{11}x_1 + a_{12}\ x_2 + ... \quad a_{1N}\ x_N = b_1$

$a_{21}x_1 + a_{22}\ x_2 + ... \quad a_{2N}\ x_N = b_2$

....

$a_{M1}x_1 + a_{M2}\ x_2 + ... \quad a_{MN}\ x_N = b_M$

Esempio:

$3x_1 + 2\ x_2 + ... \quad 4x_N = 5$

$4x_1 - 2\ x_2 + ... \quad 0.5\ x_N = 3$

....

$2x_1 + 3\ x_2 + ... \quad -3\ x_N = -1$

$\mathbf{A\ x = b}$

M x 1

Vettore dei termini noti

M x N

(Matrice di disegno)

N x 1

Vettore delle incognite

Sistema quadrato (N x N)

$a_{11}x_1 + a_{12}\, x_2 + \ldots\ldots\ a_{1N}\, x_N = b_1$
$a_{21}x_1 + a_{22}\, x_2 + \ldots\ldots\ a_{2N}\, x_N = b_2$

....

$a_{N1}x_1 + a_{N2}\, x_2 + \ldots\ldots\ a_{NN}\, x_N = b_N$

Esempio:
$3x_1 + 2\, x_2 + \ldots\ldots\ 4x_N = 5$
$4x_1 - 2\, x_2 + \ldots\ldots\ 0.5\, x_N = 3$

....

$2x_1 + 3\, x_2 + \ldots\ldots\ -3\, x_N = -1$

Ammette 1, nessuna o ∞ soluzioni

A è N x N quadrata

$\mathbf{A\, x = b}$
$\mathbf{A^{-1}Ax = A^{-1}b}$
$\mathbf{x} = \mathbf{A^{-1}\, b}$ se $\mathbf{A^{-1}}$ esiste, 1 soluzione.

altrimenti, nessuna (rette parallele)
o
∞ soluzioni (rette coincidenti).

Overview

Modelli

Sistemi lineari

Soluzione ai minimi quadrati

Soluzione dei sistemi lineari

Scrivo il sistema lineare: $Ax = b$

$y = x + 2$
$y = -3x + 1$

$A = \begin{bmatrix} 1 & -1 \\ -3 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$

$1\,x_1 - 1\,x_2 = -2$
$-3\,x_1 - 1\,x_2 = -1$

X è una soluzione se soddisfa tutte le equazioni del sistema stesso.

Soluzioni:
- ! $\exists$ Soluzione (sistema impossibile)
- $\exists$ Soluzione (sistema possibile)
  - 1 soluzione (sistema determinato)
  - > 1 soluzione (∞^k soluzioni – sistema indeterminato).

Soluzione di sistemi lineari quadrati

$x = A^{-1}\, b$

Condizione di esistenza dell'inversa è $\det(A) \neq 0$

Il sistema ammette 1 ed 1 sola soluzione se $\det(A) \neq 0$

Altrimenti: nessuna o infinite soluzioni

Risoluzione di un sistema 2x2

$a_{11}x_1+a_{12}x_2 = b_1$
$a_{21}x_1+a_{22}x_2 = b_2$

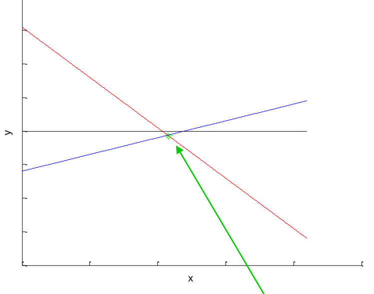
$\mathbf{y = Ax}$

$\mathbf{x = A^{-1} y}$

$$A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

$\det(A) = a_{11}*a_{22}-a_{12}*a_{21}$

Esempio



$y = x + 2$
$y = -3x + 1$

$1\ x_1\ - 1\ x_2 = -2$
$-3\ x_1 - 1\ x_2 = -1$

$x_1 = x$
$x_2 = y$

$$A = \begin{bmatrix} 1 & -1 \\ -3 & -1 \end{bmatrix} \quad b = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

Rango di A è pieno

$\det(A) = 1(-1)-(-1)(-3)=-1-3=-4$

$A^{-1} =$

P = [-1/4 7/4]

$x_1 = -1/4$
$x_2 = 7/4$

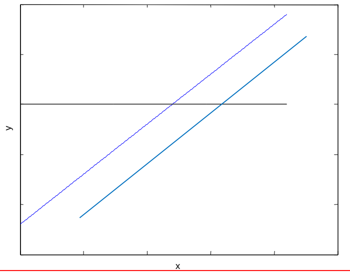
$$P = A^{-1}\, b = -\frac{1}{4}\begin{bmatrix} -1 & +1 \\ +3 & +1 \end{bmatrix} \begin{bmatrix} -2 \\ -1 \end{bmatrix}$$

Esempio di soluzione non univoca

$y = x + 2$
$2y = 2x + 3$

$A = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$ $b = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$

Non esistono soluzioni



$1\,x_1 - 1\,x_2 = -2$
$2\,x_1 - 2\,x_2 = -3$

$x_1 = x$
$x_2 = y$

det(A) = 1(-2)-(-1)(2)=-2+2=0   La soluzione non esiste o ∞ soluzioni.

$y = x + 2$
$2y = 2x + 4$

La soluzione, se esiste non è unica: tutti i punti della retta soddisfano contemporaneamente le 2 equazioni. In questo caso ∞ soluzioni: rette sovrapposte.

Sistema M x N, M > N

$a_{11}x_1 + a_{12}\,x_2 + \ldots\ldots\ a_{1N}\,x_N = b_1$
$a_{21}x_1 + a_{22}\,x_2 + \ldots\ldots\ a_{2N}\,x_N = b_2$

....

$a_{M1}x_1 + a_{M2}\,x_2 + \ldots\ldots\ a_{MN}\,x_N = b_M$

Esempio:
$3x_1 + 2\,x_2 + \ldots\ldots\ 4x_N = 5$
$4x_1 - 2\,x_2 + \ldots\ldots\ 0.5\,x_N = 3$

....

$2x_1 + 3\,x_2 + \ldots\ldots\ -3\,x_N = -1$

Ammette 1, nessuna o ∞ soluzioni

A x = b

A è M x N, M > N, non è una matrice quadrata.

1, nessuna, ∞ soluzioni.

Ho delle equazioni di troppo, devono essere correlate (combinate linearmente), perché il sistema ammetta soluzione.

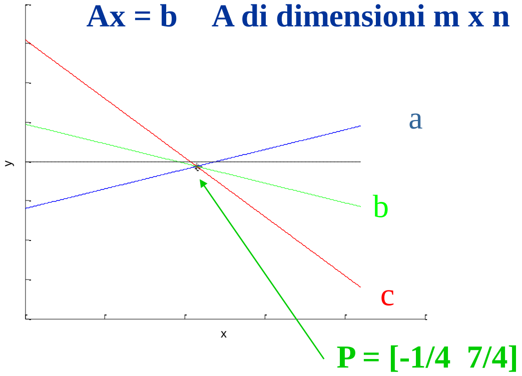
Posso sempre calcolare la soluzione in forma matriciale.

Sistemi lineari con m > n

J(W,L) è rettangolare: numero di righe maggiore del numero di colonne

Ax = b   A di dimensioni m x n

y = x + 2
y = -3x + 1
y = -x + 3/2



Una delle 3 righe di A è combinazione lineare delle altre.

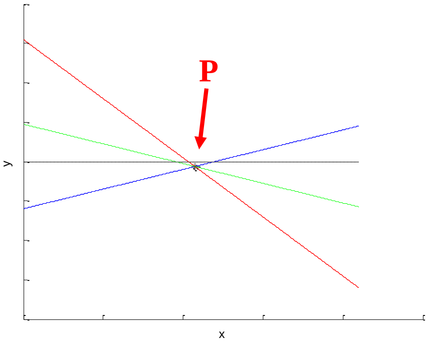
$$A=\begin{bmatrix}1 & -1\\ -3 & -1\\ -1 & -1\end{bmatrix} \qquad b=\begin{bmatrix}-2\\ -1\\ -1.5\end{bmatrix}$$

Esiste un'equazione "di troppo"

Nessuna, 1 o ∞ soluzioni

Rango di A è pieno

Relazione tra le equazioni (combinazione lineare)



α_1 (y - x - 2) +
α_2 (y +3x – 1) =
(y + x - 3/2)

In questo caso:
α_1 = -1/2
α_2 = -1/2

Tutte le rette per la soluzione P possono essere descritte come un fascio (di rette).
Un fascio di rette è univocamente identificato da due rette (che si incontrino in un punto).
La terza equazione è combinazione lineare delle prime due.

Sistema lineare: soluzione algebrica

Caso generale:

$$\mathbf{A\,x = b} \Longrightarrow \mathbf{A'\,A\,x = A'\,b} \Longrightarrow \mathbf{(A'\,A\,)^{-1}A'\,A\,x = (A'\,A\,)^{-1}A'\,b}$$

$$\Downarrow$$

(A'A) gioca il ruolo di A quadrata.

$$\mathbf{x = (A'A)^{-1}A'b}$$

$\boxed{\text{Quale criterio viene soddisfatto da } \mathbf{x}\text{?}}$

$C = (A^T * A)^{-1}$ è la matrice di covarianza (matrice quadrata n x n)

Sistemi lineari con m > n

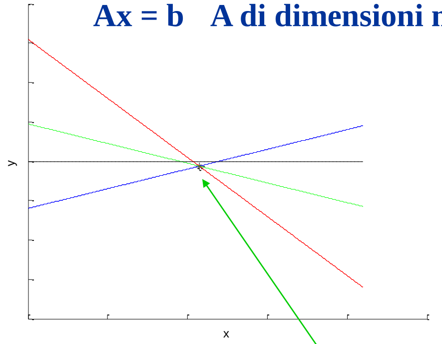
$y = x + 2$ $\qquad x_1 – x_2 = -2$

$y = -3x + 1$ $\qquad -3x_1 – x_2 = - 1$

$y = -x + 3/2$ $\qquad -x_1 – x_2 = -3/2$

Ax = b A di dimensioni m x n

$$A = \begin{bmatrix} 1 & -1 \\ -3 & -1 \\ -1 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} -2 \\ -1 \\ -1.5 \end{bmatrix}$$



P = [-1/4 7/4]

$$A^T * A = \begin{bmatrix} 11 & 3 \\ 3 & 3 \end{bmatrix} \qquad \det = 24$$

$$C = (A^TA)^{-1} = \begin{bmatrix} 0.1250 & -0.1250 \\ -0.1250 & 0.4583 \end{bmatrix}$$

$P = C * A^T * b$ **P = [-0.25 +1.75]**

intersezione

Riformulazione del problema con errore

$a_{11}x_1 + a_{12}\, x_2 + \ldots\ldots\quad a_{1N}\, x_N = b_1 + \nu_1$

$a_{21}x_1 + a_{22}\, x_2 + \ldots\ldots\quad a_{2N}\, x_N = b_2 + \nu_2$

....

$a_{M1}x_1 + a_{M2}\, x_2 + \ldots\ldots\quad a_{MN}\, x_N = b_M + \nu_M$

Modello **Misure**

Errore di modello (sistematico, randomico). M x 1 => **Residuo.**

A x = b + N

M x 1

Vettore dei termini noti

M x N
(Matrice di disegno)

N x 1
Vettore delle incognite

Quale criterio è soddisfatto da $\mathbf{x}=(x_1,x_2)$? Quali sono i parametri **w**, data la misura b? b + ν rappresenta una traslazione verticale della retta.

Soluzione come problema di ottimizzazione

Funzione costo: $(Ax - b)^2 = \sum_k \nu_k^{\,2} = \|Ax - b\|^2$

Assegno un costo al fatto che la soluzione x, non soddisfi tutte le equazioni, la somma dei residui associati ad ogni equazioni viene minimizzata. Geometricamente: viene trovato il punto a distanza (verticale) minima da tutte le rette.

$$\min_{\mathbf{x}} \sum_k \nu_k^{\,2} = \min_{\mathbf{x}} (\mathbf{Ax}-\mathbf{b})^2$$

$$\frac{d}{d\mathbf{x}}(\mathbf{Ax}-\mathbf{b})^2 = 2\mathbf{A}^T(\mathbf{Ax}-\mathbf{b}) = 0$$

$$\mathbf{A^TAx = A^Tb}$$

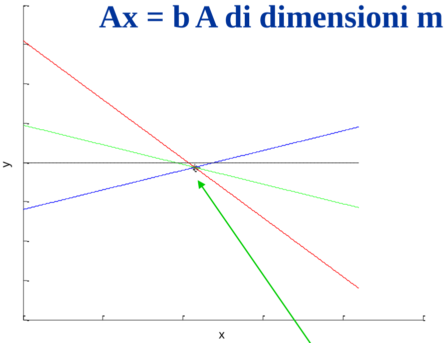
$$\mathbf{x = (A^TA)^{-1}A^Tb}$$

NB le funzioni costo sono spesso quadratiche (problemi di minimizzazione convessi) perchè il costo cresce sia che il modello sovrastimi che sottostimi le misure.
Inoltre, le derivate calcolate per imporre le condizioni di stazionarietà (minimo), sono relativamente semplici.

Sistemi lineari con m > n

$x_1 - x_2 = -2$
$-3x_1 - x_2 = -1$
$-x_1 - x_2 = -3/2$

Ax = b A di dimensioni m x n



P = [-1/4 7/4]

$$A = \begin{bmatrix} 1 & -1 \\ -3 & -1 \\ -1 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} -2 \\ -1 \\ -1.5 \end{bmatrix}$$

$$A^T * A = \begin{bmatrix} 11 & 3 \\ 3 & 3 \end{bmatrix} \qquad \det = 24$$

$$C = (A^T A)^{-1} = \begin{bmatrix} 0.1250 & -0.1250 \\ -0.1250 & 0.4583 \end{bmatrix}$$

$P = C * A^T * b$ **P = [-0.25 +1.75]**

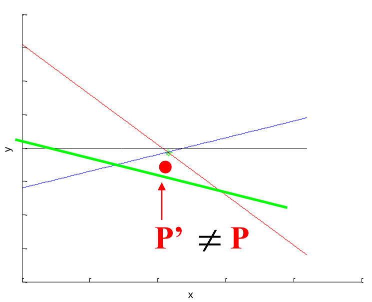
intersezione

$||Ax - b|| = 0$

Sistemi lineari con m > n - non esiste soluzione (matematica)

$x_1 - x_2 = -2$ **+ 0**
$3x_1 + x_2 = +1$ **- 0.5 = +0.5**
$-x_1 - x_2 = -3/2 +$ **0 = -3/2**

AX = b **A di dimensioni m x n**



$$A = \begin{bmatrix} 1 & -1 \\ 3 & 1 \\ -1 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} -2 \\ 0.5 \\ -1.5 \end{bmatrix}$$

$$A^T * A = \begin{bmatrix} 11 & 3 \\ 3 & 3 \end{bmatrix} \qquad \det = 24$$

$$\sum_k v_k^2 = ||Ax - b||^2 = 0.2041241$$

$$C = (A^T A)^{-1} = \begin{bmatrix} 0.1250 & -0.1250 \\ -0.1250 & 0.4583 \end{bmatrix}$$

$P = C * A^T * b$ **P' = [-0.375 +1.70833]**

No intersezione

Commenti

$$\sum_k \nu_k^{\;2} = ||Ax - b||^2 = \sum_k || A_{k,*}x - b_k ||^2 =$$

$$[(A_{11}x_1 + A_{12}x_2) - b_1]^2 + [(A_{21}x_1 + A_{22}x_2) - b_2]^2 +$$

$$[(A_{31}x_1 + A_{32}x_2) - b_3]^2$$

P' = [-0.375 +1.70833]

No intersezione

$x_1 - x_2 = -2$ $\quad$ $-0.375-1.70833+2 = \nu_1 = -0.083333$

$3x_1 + x_2 = \mathbf{+1/2}$ $\quad$ $-1.125+1.70833-0.5 = \nu_2 = 0.083333$

$-x_1 - x_2 = \mathbf{-}3/2$ $\quad$ $+0.375 -1.70833+1.5 = \nu_3 = 0.16666$

Lo scarto misura la somma quadratic delle distanze (verticali) tra il punto che rappresenta la soluzione e le 3 rette.

Overview

Modelli

Sistemi lineari

Soluzione ai minimi quadrati