Sistemi lineari sovradeterminati

I sistemi lineari sovradeterminati hanno un numero di equazioni superiore al numero di incognite, ossia sono della forma

$$Ax = b$$

dove $A \in R^{m\times n}, \; x \in R^n \; e \; b \in R^m, \; m > n$

$$\begin{bmatrix} \\ \\ \\ & & A & & \\ & & m>n & & \\ \\ \\ \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \\ \\ b \\ \\ \\ \\ \end{bmatrix}$$

e quindi potrebbero non avere soluzioni.

In particolare se

se $rank(A) \neq rank(A|b)$ → sistema incompatibile nessuna soluzione

se $rank(A) = rank(A|b)$ → sistema compatibile

$$\begin{cases} \boldsymbol{rank(A) = n} & \qquad \boldsymbol{soluzione\ unica} \\ \\ \boldsymbol{rank(A) < n} & \qquad \infty^{\boldsymbol{n-rank(A)}}\ \boldsymbol{soluzioni} \end{cases}$$

La risoluzione di un sistema lineare sovradeterminato risulta, quindi, essere un problema mal posto in quanto potrebbe accadere che la soluzione non esista o, se esiste, non sia unica.
Per rendere il problema ben posto (ossia per fare in modo che ammetta sempre una ed una sola soluzione che dipende con continuità dai dati del problema) si riconsidera una sua riformulazione del seguente tipo.

Risoluzione di sistemi lineari sovradeterminati nel senso dei minimi quadrati (Least Squares) .

Problema ben posto:

Cerchiamo la soluzione del sistema lineare sovradeterminato

$$Ax = b$$

dove $A \in R^{m\times n}$, $x \in R^n$ e $b \in R^m$, $m > n$

nel senso dei minimi quadrati, cioè, definito il vettore residuo come

$$r(x) := Ax - b$$

cerchiamo il vettore $x^* \in R^n$ che rende minima la norma 2 al quadrato del residuo, cioè

$$x^* = \underset{x\in R^n}{\arg\min}\ ||r(x)||_2^2 = \underset{x\in R^n}{\arg\min}\ ||Ax - b||_2^2 \qquad (1)$$

$$F(x) = ||Ax - b||_2^2 = (Ax - b)^T(Ax - b) = x^T A^T Ax - x^T A^T b - b^T A\, x + b^T b = x^T A^T Ax - 2x^T A^T b + b^T b$$

Poniamo $G = A^T A$. Si ha che G è simmetrica, infatti $G^T = (A^T A)^T = A^T A = G$.

$$F(x) = x^T G\, x - 2x^T A^T b + b^T b$$

Per calcolare il valore $x^* \in R^n$ che rende minimo $F(x)$, calcoliamo il gradiente della funzione $F(x)$ ed imponiamo che si annulli.

Ricordiamo che il gradiente rispetto ad x della forma quadratica $x^T G\, x$, con G simmetrica, è uguale a $2\, G\, x$.

Allora si ha:

$$\nabla F(x) = 2\, G\, x - 2A^T b = 0$$

Il vettore $x^* \in R^n$ che annulla il gradiente della funzione $F(x)$ è la soluzione del sistema lineare

$$\boldsymbol{G\, x = A^T b}$$

Queste ultime equazioni sono note come **equazioni normali** e la loro risoluzione porta alla risoluzione del problema dei minimi quadrati.

Abbiamo trasferito il problema della risoluzione di un sistema sovra determinato in quello della risoluzione di un sistema quadrato con matrice $G = A^T A$ di dimensione $n \times n$. .

Ovviamente affinché il sistema quadrato di dimensione $n \times n$ sia risolubile è necessario che il determinante della matrice $A^T A$ sia diverso da zero, e quindi è necessario che la matrice A abbia rango n .

Inoltre, affinché la soluzione cercata sia proprio un minimo è necessario che la matrice Hessiana valutata nella soluzione sia una matrice definita positiva.

Poichè

$$\nabla^2 F(x) = 2\, G = 2\, A^T A$$

occorre che la matrice $G = A^T A$, che abbiamo visto essere simmetrica, sia definita positiva.

Verifichiamo che è definita positiva, cioè:

$$\forall\ x \neq 0, \quad x^T A^T A x > 0$$

A tale scopo, si osserva che $x^T A^T A x$ può essere vista come

$$x^T A^T A x = (Ax)^T A x = ||Ax||_2^2.$$

Perciò, per le proprietà sulle norme, $\| Ax\|_2^2 > 0 \quad \forall\, Ax \neq 0$.

Se la matrice A è a rango massimo, cioè costituita da vettori colonna linearmente indipendenti, allora $Ax = 0$ se e solo se $x = 0$, che è escluso dalle ipotesi.

Pertanto la condizione che garantisce la risolubilità del sistema delle equazioni normali garantisce anche che la soluzione trovata sia un minimo.

TEOREMA

Dato il sistema lineare sovradeterminato

$$Ax = b$$

dove $A \in R^{m\times n},\ x \in R^n \quad e\ b \in R^m,\ m > n$

$x^* = \arg\min_{x\in R^n} ||Ax - b||_2^2 \iff$ è la soluzione di $A^T Ax = A^T b$

La soluzione è unica se e solo se la matrice A ha rango massimo $rank(A) = n$ (le colonne di A sono linearmente indipendenti).

Se $rank(A) < n$ invece le colonne di A sono linearmente dipendenti allora il problema ha infinite soluzioni; tuttavia quella di lunghezza (euclidea) minima è unica.

La soluzione del problema dei minimi quadrati mediante equazioni normali richiede solo che la matrice A del sistema sovradeterminato Ax=b abbia rango massimo.

Se questa condizione è verificata il vettore x* tale che

$$x^* = \arg\min_{x\in R^n} \|Ax - b\|_2^2$$

può essere sempre calcolato, teoricamente, come soluzione delle equazioni normali

$$A^T Ax = A^T b.$$

La matrice G = A^TA è simmetrica e definita positiva e quindi il sistema può essere risolto utilizzando il metodo di Cholesky.

ATTENZIONE AL MALCONDIZIONAMENTO

Poiché

$$K_2(A^TA) = (\,K_2(A))^2$$

questo implica che il sistema delle equazioni normali può risultare mal condizionato anche quando il problema nella sua forma originale non lo è.

Se quindi **A è ben condizionata** il metodo delle equazioni normali è un ottimo metodo per calcolare la soluzione del problema dei minimi quadrati. Se A è 'mediamente mal condizionata' il metodo delle equazioni normali è ***numericamente pericoloso*** perché può fornire risultati non affidabili.

Per ovviare a questo problema, nei prossimi paragrafi considereremo, per risolvere il problema dei minimi quadrati, due altri metodo, che si basano su un principio diverso: poiché sappiamo che le trasformazioni ortogonali lasciano inalterata la norma 2 di un vettore l'idea è di individuare una trasformazione ortogonale che, applicata al residuo r = Ax-b, lo trasformi in modo tale da rendere più facile la soluzione del problema di minimizzarne la norma.

Proprietà: Una matrice ortogonale $Q \in R^{m\times m}$ applicata ad un vettore $y \in R^m$, ne mantiene inalterata la norma 2 al quadrato, cioè:

$$||y||_2^2 = ||Qy||_2^2$$

Dim.

Ricordiamo che una matrice ortogonale è tale che $Q^TQ = QQ^T = I$

$$||y||_2^2 = y^Ty = y^TQ^T\,Qy = (Q\quad y)^T(Q\quad y) = ||Qy||_2^2$$

Vale anche

$$||y||_2^2 = y^Ty = y^T\,QQ^Ty = (Q^Ty)^T(Q^Ty) = ||Q^Ty||_2^2$$

Osservazione: Dato un vettore $y \in R^m$, sia $m>n$, consideriamo la sua norma 2 al quadrato:

$$||y||_2^2 = \sum_{i=1}^{m} y_i^2 = \sum_{i=1}^{n} y_i^2 + \sum_{i=n+1}^{m} y_i^2$$

Suddividiamo il vettore y in due vettori $\tilde{y}_1$, costituito dalle prime n componenti del vettore y ed $\tilde{y}_2$ formato dalle rimanenti m-n.

$$y = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{array}{l} \text{n componenti} \\ \text{m-n componenti} \end{array}$$

$$||y||_2^2 = ||\tilde{y}_1||_2^2 + ||\tilde{y}_2||_2^2$$

Metodo QR per la risoluzione del problema dei Minimi Quadrati (QRLS)

Abbiamo visto che assegnata una matrice rettangolare $A \in R^{m \times n}$, con **rango n**, essa può essere sempre fattorizzata, mediante successive trasformazioni ortogonali di Householder nella forma

$$A = QR = [Q] \cdot \begin{array}{c} \overbrace{}^{n} \\ \left[\begin{array}{ccc} \ddots & & R_1 \\ & \ddots & \\ 0 & & \ddots \\ \cdots & \cdots & \cdots \\ & 0 & \end{array}\right] \end{array} \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} n \\ \\ m-n \end{array}$$

dove Q è ortogonale mxm e R_1 è una matrice triangolare superiore nxn con elementi r_{ii} tutti diversi da 0.

Una volta calcolata la fattorizzazione QR di A, poiché Q è ortogonale e quindi anche Q^T è ortogonale, consideriamo

$$\|Q^T(Ax-b)\|_2^2 = \|Q^TAx - Q^Tb\|_2^2 = \|R_1x - Q^Tb\|_2^2.$$

(infatti premoltiplicando ambo i membri di $A = QR$ per Q^T, si ottiene

$$Q^TA = Q^TQR = R \qquad (\text{ essendo Q ortogonale e quindi } Q^TQ = I)$$

Posto

$$h = Q^T b = \begin{bmatrix} h_1 \\ \cdots \\ h_2 \end{bmatrix} \begin{matrix} \}n \\ \\ \}m-n \end{matrix}$$

si ha

$$\left\| \overbrace{\begin{bmatrix} \ddots & & R_1 \\ & \ddots & \\ 0 & & \ddots \\ \cdots & \cdots & \cdots \\ & 0 & \end{bmatrix}}^{n} x - \begin{bmatrix} h_1 \\ \\ \\ \cdots \\ h_2 \end{bmatrix} \begin{matrix} \}n \\ \\ m-n \end{matrix} \right\|_2^2 = \left\| \begin{matrix} (R_1 x - h_1) \\ \\ -h_2 \end{matrix} \right\|_2^2$$

$$\arg\min_{x \in R^n} \|Ax - b\|_2^2 = \arg\min_{x \in R^n} \{\|R_1 x - h_1\|_2^2 + \|h_2\|_2^2\}$$
$$= \arg\min_{x \in R^n} \|R_1 x - h_1\|_2^2 + \|h_2\|_2^2.$$

Quindi il minimo sarà ottenuto per x^* che risolve il sistema lineare

$$R_1 x = h_1$$

Il metodo QR per la soluzione del problema dei minimi quadrati consiste quindi nel calcolare i due fattori Q ed R di A, nel valutare h_1 e risolvere il sistema lineare $R_1 x = h_1$, con R_1 matrice triangolare superiore.

Il valore del residuo assunto per questo valore di x sarà dato da:

$$\min_{x \in R^n} \|Ax - b\|_2^2 = \|h_2\|_2^2$$

Due sono i vantaggi da un punto di vista dell'affidabilità della soluzione nell'utilizzare il metodo QR. Il primo è quello di lavorare sempre solo sulla matrice A, senza dover passare alla matrice A^TA che, come abbiamo visto, è molto peggio condizionata di A. Il secondo è che la fattorizzazione QR è, se pur non stabile in senso forte, *abbastanza stabile*. Più precisamente, ricordando che la definizione di stabilità forte di una fattorizzazione di A (che si suppone normalizzata $a_{ij} \leq 1$) chiede che esistano due costanti *a* e *b* indipendenti dalla dimensione e dagli elementi di A tali che gli elementi delle due matrici della fattorizzazione siamo maggiorabili da queste due costanti, abbiamo che per la fattorizzazione QR vale

$$\| q_{ij} \| \leq a \quad e \quad \| r_{ij} \| \leq b$$

con $a = 1 \quad e \quad b = \sqrt{n}$.

Questo ci dice che gli elementi di R possono crescere al massimo come $\sqrt{n}$ e quindi la stabilità della fattorizzazione è abbastanza buona.

Se quindi la matrice A mxn ha rango n ed è 'mediamente mal condizionata' la soluzione del problema dei minimi quadrati con il metodo QR produce in genere risultati sufficientemente accurati.

Abbiamo visto che sia il metodo delle equazioni normali sia il metodo QR richiedono che il rango di $A \in R^{m \times n}$, $m > n$ sia n, cioè la matrice abbia rango massimo. Se la matrice A non ha rango massimo n, allora è possibile utilizzare la **Decomposizione ai Valori Singolari di A** (denominata **SVD**) e aggiungere un' altra condizione sulla soluzione cercata affinchè essa risulti unica.

Decomposizione in Valori Singolari di una matrice: SVD

Data la matrice A $m \times n$, la sua decomposizione in valori singolari è caratterizzata dal seguente teorema:

<u>Teorema:</u>

Sia $A \in R^{mxn}$ a rango $k \leq min(m, n)$. Allora esistono due matrici ortogonali $U \in R^{mxm}$ *e* $V \in R^{nxn}$ *tali che*

$$U^T AV = \Sigma = diag(\sigma_1, \sigma_2, \cdots, \sigma_k, 0, \ldots, 0) \qquad (1)$$

I valori sulla diagonale di Σ sono detti valori singolari di A e soddisfano

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0.$$

Se la matrice A ha rango $k = n$, avremo $\sigma_n > 0$,

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n > 0$$

se ha rango $k < n$ avremo che $\sigma_k > 0$ e $\sigma_{k+1} = \sigma_{k+2} = \cdots \sigma_n = 0$.

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_k > 0$$

Se A ha rango k=n

$$\Sigma = \begin{array}{c} \overbrace{\hphantom{\sigma_1 \quad \sigma_2 \quad \ddots \quad \ddots \quad \ddots \quad \sigma_n}}^{n} \\ \left[\begin{array}{cccccc} \sigma_1 & & & & & \\ & \sigma_2 & & & 0 & \\ & & \ddots & & & \\ & & & \ddots & & \\ & 0 & & & \ddots & \\ & & & & & \sigma_n \\ - & - & - & - & - & - \\ & & & & & \\ & & & 0 & & \\ & & & & & \end{array}\right] \end{array} \begin{array}{l} \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\\1\end{array}}\right\} n \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} m-n \end{array}$$

Se A ha rango k<n

$$\Sigma = \overbrace{\left[\begin{array}{cccccc} \sigma_1 & & & & & \\ & \sigma_2 & & & 0 & \\ & & \ddots & & & \\ & & & \sigma_k & & \\ & 0 & & & \ddots & \\ & & & & & 0 \\ - & - & - & - & - & - \\ & & & & & \\ & & 0 & & & \\ & & & & & \end{array}\right]}^{n}$$

$U^T A V = \Sigma$ è detta decomposizione in valori singolari (o Singular Value Decomposition (SVD)) di A e può essere anche scritta come :

$$A = U \Sigma V^T$$

Le colonne di U e V sono dette, rispettivamente, **vettori singolari sinistri e destri di *A***.

I valori singolari di una matrice hanno le seguenti proprietà:

- σ_i sono sempre reali e ≥ 0;
- σ_1 è il massimo valore singolare, anche indicato con σ_{max} , e il $\sigma_i > 0$ più piccolo è detto σ_{min}.
- Il rapporto $\sigma_{max}/\sigma_{min}$ ci fornisce l'indice di condizionamento della matrice rettangolare A;
- il numero di valori singolari non nulli (cioè l'indice di σ_{min}) rappresenta il rango della matrice A;
- inoltre

$$\sigma_i(A) = \sqrt{\lambda_i(A^T A)} \quad i = 1, \cdots, n.$$

Utilizzando i valori singolari e i corrispondenti vettori singolari si può ottenere la decomposizione spettrale di A nella forma

$$A = \sum_{j=1}^{n} \sigma_j u_j v_j^T = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots \sigma_n u_n v_n^T.$$

Questa decomposizione è particolarmente importante per quei problemi in cui è necessario ottenere una approssimazione di A mediante una matrice A_r che abbia rango k < rango di A.

Facendo uso della fattorizzazione SVD di una matrice $A \in R^{mxn}$ a rango $k \leq min(m, n)$ è possibile calcolare la matrice **pseudo-inversa** di A, che si denota con A^+, che è una generalizzazione della matrice inversa al caso di matrici non quadrate.

$$A^+ = V\,\Sigma^+ U^T$$

dove

$$\Sigma^+ = \begin{bmatrix} \frac{1}{\sigma_1} & & & & & \\ & \frac{1}{\sigma_2} & & & 0 & \\ & & \ddots & & & \\ & & & \frac{1}{\sigma_k} & & \\ & 0 & & & \ddots & \\ & & & & & 0 \\ - & - & - & - & - & - \\ & & & & & \\ & & 0 & & & \\ & & & & & \end{bmatrix}$$

La matrice pseudoinversa A^+ gode della seguenti proprietà:

1) $A\,A^+ A = A$
2) $A^+ A\,A^+ = A^+$
3) $(A^+ A)^T = A^+ A$

4) $(AA^{+})^{T} = AA^{+}$

Se $A \in R^{mxn}$ con m>n è a rango massimo, cioè k=n, allora la pseudoinversa di A è la matrice

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

ed è un'inversa sinistra di A, cioè

$$A^{+}A = I$$

dove I è l'identità.

Se $A \in R^{mxn}$ con m<n è a rango massimo, cioè k=m, allora la pseudoinversa di A è la matrice

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

ed è un'inversa destra di A, cioè

$$AA^{+} = I$$

dove I è l'identità.

Decomposizione in Valori Singolari per la soluzione del problema dei minimi Quadrati

Sia il metodo delle equazioni normali sia il metodo QR richiedono che il rango di A sia n. Se questo non è vero dobbiamo usare la fattorizzazione SVD di A e aggiungere un' altra condizione sulla soluzione cercata perché essa risulti unica.

Possiamo quindi considerare, ricordando che trasformazioni ortogonali lasciano invariata la norma 2,

$$||Ax - b||_2^2 = ||U^{T}(Ax - b)||_2^2 = ||U^{T}AVV^{T}x - U^{T}b||_2^2 = ||\Sigma V^{T}x - U^{T}b||_2^2$$

Se $k \leq n$ è il rango di A,

$$c = V^{T}x \qquad \text{e} \qquad d = U^{T}b = \begin{bmatrix} {u_1}^{T}b \\ {u_2}^{T}b \\ {u_n}^{T}b \\ \\ {u_m}^{T}b \end{bmatrix} = \begin{array}{l} \begin{bmatrix} d_1 \\ \\ d_2 \end{bmatrix} \end{array} \begin{array}{l} \}\, k \\ \}\, m - k \end{array}$$

otteniamo

$$\arg\min_{x\in R^n} ||Ax-b||_2^2 = \arg\min_{c\in R^n} ||\Sigma c-d||_2^2 = \arg\min_{c\in R^n} ||\Sigma c-d_1||_2^2 + ||d_2||_2^2$$

dove

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & 0 & \\ & & \ddots & & \\ & 0 & & \sigma_k & \\ & & & & 0 \\ - & - & - & - & - \\ & & 0 & & \end{bmatrix}$$

Per rendere minimo il residuo dobbiamo calcolare il valore di c che annulla $(\Sigma c - d_1)$ cioè risolvere

$$\Sigma c = d_1 .$$

Si ha

$$c_i = \frac{d_i}{\sigma_i} \quad \text{i=1,…,k}$$

e poniamo c_i=0 per i=k+1, …,n che rappresenta una condizione aggiuntiva per ottenere, fra le infinite soluzioni, quella di minima norma.

Ricaviamo poi

$$x = Vc = \sum_{i=1}^{k} c_i v_i$$

dove $v_i, i = 1,\ldots,n$ indicano le colonne della matrice V,

e poichè

$$c_i = \frac{d_i}{\sigma_i} = \frac{{u_i}^T b}{\sigma_i}$$

dove $u_i, i = 1,\ldots,n$ indicano le colonne della matrice U,

otteniamo

$$x = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i,$$

che rappresenta la **soluzione di minima norma del problema dei minimi quadrati**.

(Nota che equivale a $x = A^+b = V\,\Sigma^+ U^T\, b$)

Il valore del residuo al minimo sarà

$$\min_{x \in R^n} ||Ax - b||_2^2 = ||d_2||_2^2$$

cioè $\sum_{i=k+1}^{m} {d_i}^2 = \sum_{i=k+1}^{m} ({u_i}^T b)^2$.

Approssimazione nel senso dei minimi quadrati di un insieme di dati sperimentali

Siano assegnate m coppie di valori sperimentali (x_i, y_i) i=0,…,*m-1* , (in cui tutte le ascisse siano distinte tra loro) , che descrivono in modo discreto l'andamento di un fenomeno reale (es. l'andamento della quotazione di un titolo azionario nel mese di marzo, o le misure di piovosità in una certa stazione di misura nei primi mesi dell'anno).

Si vuole determinare il polinomio di grado *n*, ***m>n***,

$$P_n(x) = \sum_{j=0}^{n} \alpha_j x^j$$

i cui coefficienti $\alpha_j, j = 0, .. n$ si ottengono risolvendo nel senso dei minimi quadrati il sistema lineare sovradeterminato ottenuto dalle equazioni:

$$P_n(x_i) = y_i \quad i=0,..,m-1 \qquad (2)$$

Indichiamo con α il vettore di dimensione *n+1* dei coefficienti della combinazione lineare

$$\alpha = \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_n \end{bmatrix},$$

con *y* il vettore, di dimensione *m* delle misure sperimentali

$$y = \begin{bmatrix} y_0 \\ \vdots \\ y_{m-1} \end{bmatrix}$$

e con B la matrice $(m) \times (n+1)$, con rango n+1, data da

$$B = \begin{bmatrix} x_0^0 & x_0^1 & \dots & x_0^n \\ x_1^0 & x_1^1 & \dots & x_1^n \\ x_2^0 & x_2^1 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \vdots \\ x_{m-1}^0 & x_{m-1}^1 & & x_{m-1}^n \end{bmatrix}$$

il cui elemento di posto (i,j) è dato da $b_{ij} = x_i^j$.

dove $B \in R^{(m)\times(n+1)}$, rank(B)=n+1, m>n

Individuare i coefficienti del polinomio $P_n(x) = \sum_{j=0}^{n} \alpha_j x^j$ in maniera tale che $P_n(x_i) = y_i$ *i=0,..,m-1* *equivale a* determinare la soluzione nel senso dei minimi quadrati del sistema lineare sovradeterminato

$$B\alpha = y\,.$$

che equivale a trovare il vettore dei coefficienti α

$$\alpha^* = \underset{\alpha \in R^n}{\arg\min} \; \|B\alpha - y\|_2^2$$

problema che può essere risolto con uno dei tre metodi precedentemente illustrati.

Poiché $B\alpha = \begin{bmatrix} \sum_{j=0}^{n} b_{0j}\ \alpha_j \\ \sum_{j=0}^{n} b_{1j}\ \alpha_j \\ \dots \\ \sum_{j=0}^{n} b_{ij}\ \alpha_j \\ \\ \sum_{j=0}^{n} b_{mj}\ \alpha_j \end{bmatrix}$, applicando la definizione di norma 2 al quadrato di un vettore, si può scrivere che

$$\boldsymbol{\alpha^* = \underset{\alpha \in R^n}{\arg\min} \; \|B\alpha - y\|_2^2} = \underset{\alpha \in R^n}{\arg\min} \; \sum_{i=0}^{m-1} \left(\sum_{j=0}^{n} b_{ij}\ \alpha_j - y_i \right)^2$$

ricordando che $b_{ij} = x_i^j$

$$= \arg\min_{\alpha \in R^n} \sum_{i=0}^{m-1} \left(\sum_{j=0}^{n} x_i^j \ \alpha_j - y_i \right)^2 = \mathbf{arg\,min}_{\boldsymbol{\alpha \in R^n}} \sum_{i=0}^{m-1} (\ P_n(x_i) - y_i)^2$$

ES: Retta dei minimi Quadrati

Siano dati i punti sperimentali (x_i, y_i) i=0,..,m-1, si vuole determinare la retta che approssima i dati nel senso dei minimi quadrati, cioè si vuole determinare il polinomio $P_1(x) = \alpha_0 + \alpha_1 x$ tale che $P_1(x_i) = \alpha_0 + \alpha_1 x_i = y_i \ \forall\, i = 0, .., m-1$

che significare imporre che

$$P_1(x_0) = \alpha_0 + \alpha_1 x_1 = y_0$$

.

.

$$P_1(x_m) = \alpha_0 + \alpha_1 x_m = y_m$$

che equivale al sistema $A\alpha = y$

$$A = \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Consideriamo per la soluzione nel senso dei minimi quadrati di questo sistema lineare il metodo delle Equazioni Normali.

Calcoliamo $A^T A$

$$A^T A = \begin{bmatrix} 1 & 1.... & 1 \\ x_0 & x_1..... & x_m \end{bmatrix} \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} = \begin{bmatrix} m+1 & \sum_{i=0}^{m} x_i \\ \sum_{i=0}^{m} x_i & \sum_{i=0}^{m} x_i^2 \end{bmatrix}$$

e

$$A^T y = \begin{bmatrix} \sum_{i=0}^{m} y_i \\ \sum_{i=0}^{m} x_i y_i \end{bmatrix}$$

e risolviamo il corrispondente sistema lineare

$$A^T A\, \alpha = A^T y$$

---

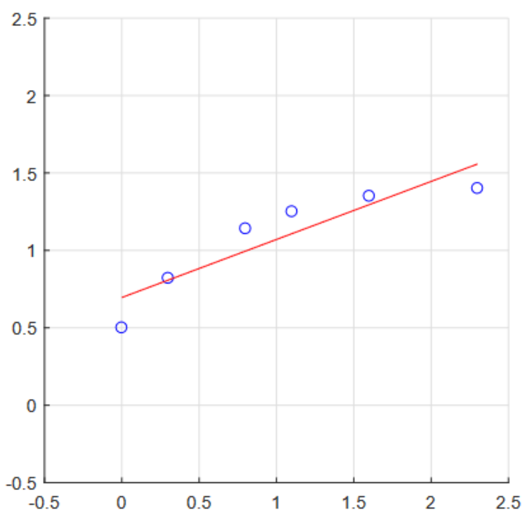
il cui vettore soluzione ci permette di individuare la retta richiesta

$$P_1(x) = \alpha_0 + \alpha_1\, x$$

Caso n = 1: polinomio lineare di approssimazione ai minimi quadrati (retta di regressione lineare)

Determinare la retta $P_1(x) = \alpha_0 + \alpha_1\, x$ che approssima nel senso dei minimi quadrati l'insieme di dati sperimentali.

x	y
0	0.5
0.3	0.82
0.8	1.14
1.1	1.25
1.6	1.35
2.3	1.40



$$p_1(x) = 0.6955 + 0.3749x$$

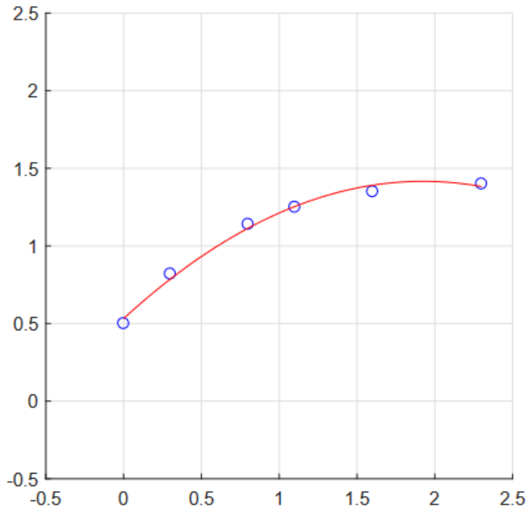
Calcolo del residuo:

$$\|\, \|r(\alpha)\|_2^2 = \|y - B\,\alpha\|_2^2 = \Sigma_{i=0}^{5}(\,(P_1(x_i) - y_i)^{\,2}) = 0.1073$$

Caso n = 2: polinomio quadratico di approssimazione ai minimi quadrati

Determinare la parabola $P_2(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2$ che approssima nel senso dei minimi quadrati l'insieme di dati sperimentali

x	y
0	0.5
0.3	0.82
0.8	1.14
1.1	1.25
1.6	1.35
2.3	1.40



$$p_2(x) = 0.5318 + 0.9191x - 0.2387x^2$$

$$\|r(\alpha)\|_2^2 = \|y - B\,\alpha\|_2^2 = \sum_{i=0}^{5}(\,(P_2(x_i) - y_i)^{\,2}) = 0.0048$$