Vittorio Casella

DIET – Università di Pavia

email: vittorio.casella@unipv.it

I Minimi Quadrati (MQ) lineari
Introduzione

Dispense

L'esempio del fenomeno lineare

Molti fenomeni seguono la legge

$$y = px + q \tag{1}$$

Si tratta di un fenomeno lineare, il cui grafico, nello spazio (O,x,y) è rappresentato da una retta

Esempi di fenomeni lineari: la molla

Una molla è un sistema fisico avente equazione

$l = kF + l_r$ dove

l lunghezza della molla

l_r lunghezza a riposo della molla (non sollecitata da alcuna forza)

k coefficiente elastico della molla

F forza con cui la molla viene sollecitata

In questo caso p e q si identificano rispettivamente con il coefficiente di elasticità e con la lunghezza a riposo; x e y rispettivamente con la forza applicata e la lunghezza della molla

Esempi di fenomeni lineari: la dilatazione termica

Nel caso della dilatazione termica, p e q rappresentano rispettivamente il coefficiente di dilatazione termica e la lunghezza iniziale; x e y rispettivamente la variazione di temperatura e la lunghezza dell'oggetto considerato

Il problema diretto e quello inverso

La legge (1) può essere usata in due modi: il problema diretto e quello inverso.

Il **problema diretto,** dalla teoria all'esperimento. Si conoscono i parametri p e q del fenomeno e si deve prevedere il suo comportamento: che lunghezza assumerà la molla se sollecitata con una certa forza; che dilatazione subirà un certo oggetto a causa del riscaldamento.

Il **problema inverso**, dall'esperimento alla teoria. In un contesto sperimentale, la misura di alcune coppie di valori $\left(x_i, y_i\right)$, fatta l'ipotesi che il fenomeno segua la legge (1), permette di ricavare i parametri incogniti p e q.

Formulazione matriciale del problema diretto

In corrispondenza dei valori $x_1, x_2, \ldots, x_m$ il fenomeno (1) assumerà i valori

$$y_i = px_i + q \qquad i = 1,2,\ldots,m \tag{2}$$

che equivale a dire

$$\begin{aligned} y_1 &= px_1 + q \\ y_2 &= px_2 + q \\ &\vdots \\ y_m &= px_m + q \end{aligned} \tag{3}$$

Formulazione matriciale del problema diretto - 2

Introduciamo

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \qquad \mathbf{A} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} p \\ q \end{pmatrix} \tag{4}$$

La (3) è equivalente a

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \tag{5}$$

Se si vuole conoscere il comportamento di un fenomeno noto, in corrispondenza di certi valori di x_i, si costruiscono la matrice **A** e il vettore **X** e si calcola il loro prodotto. Nel vettore **Y** così trovato si leggono i valori y_i cercati.

Il problema inverso

Di un certo fenomeno che segue la legge (1) non si conoscono i parametri; si *misurano* alcune coppie

$$\left(x_i, y_i\right) \qquad i = 1,2,\ldots,m \tag{6}$$

e si cerca di stimare le proprietà del fenomeno a partire dal suo comportamento. In termini matematici: trovare l'equazione delle retta.

Nel caso della molla per esempio si potrebbe pensare di applicare una serie di forze, materializzate da pesi tarati, e di misurare la lunghezza della molla.

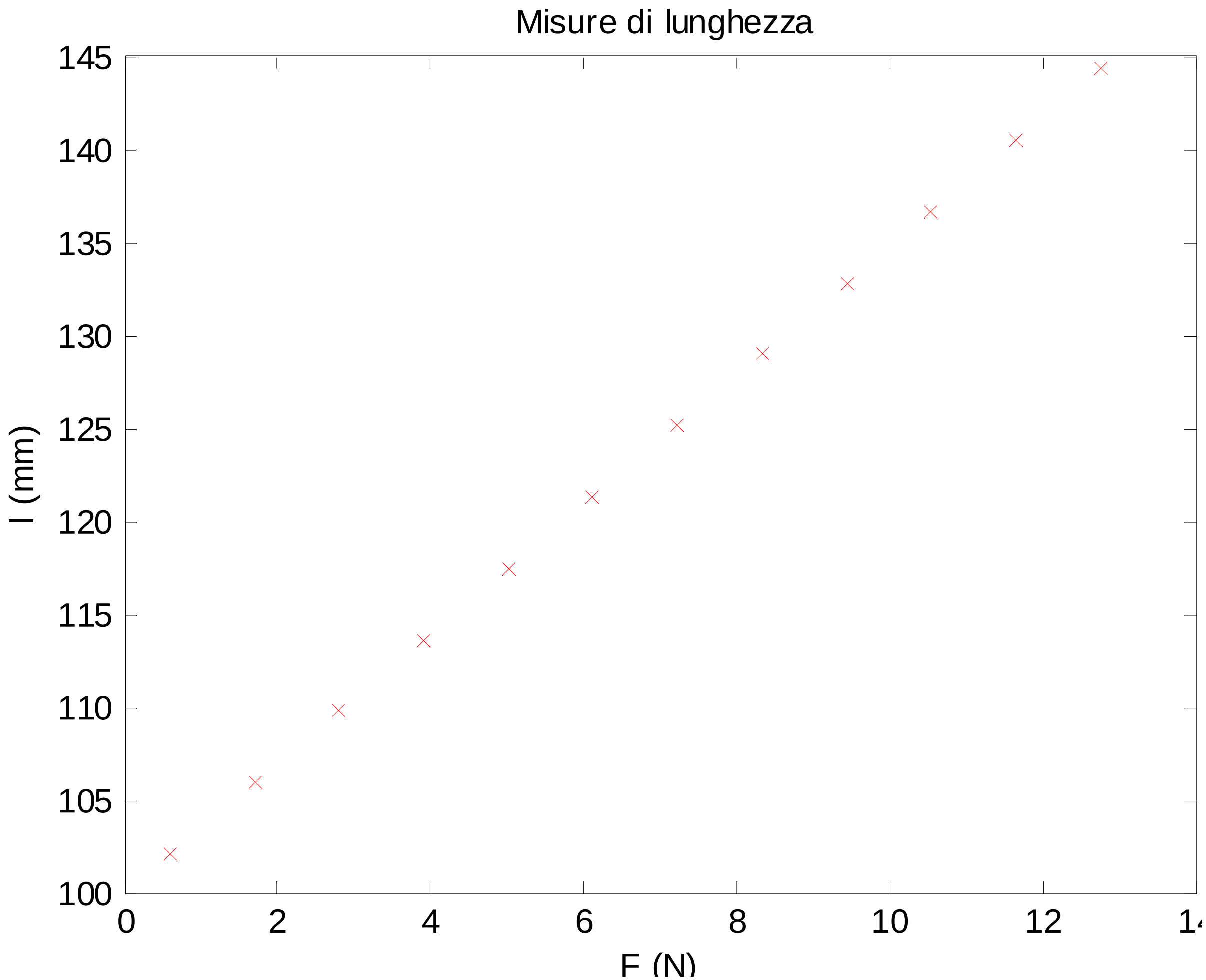
Il problema inverso - 2

Nel problema inverso, il vettore **Y** non rappresenta ciò che deve essere calcolato, ma ciò che si conosce in forza di misure. Si indica $\mathbf{Y}_0$ per sottolineare che esso contiene *m* ben precise misure

$$\mathbf{Y}_0 = \begin{pmatrix} y_{01} \\ y_{02} \\ \vdots \\ y_{0m} \end{pmatrix}$$

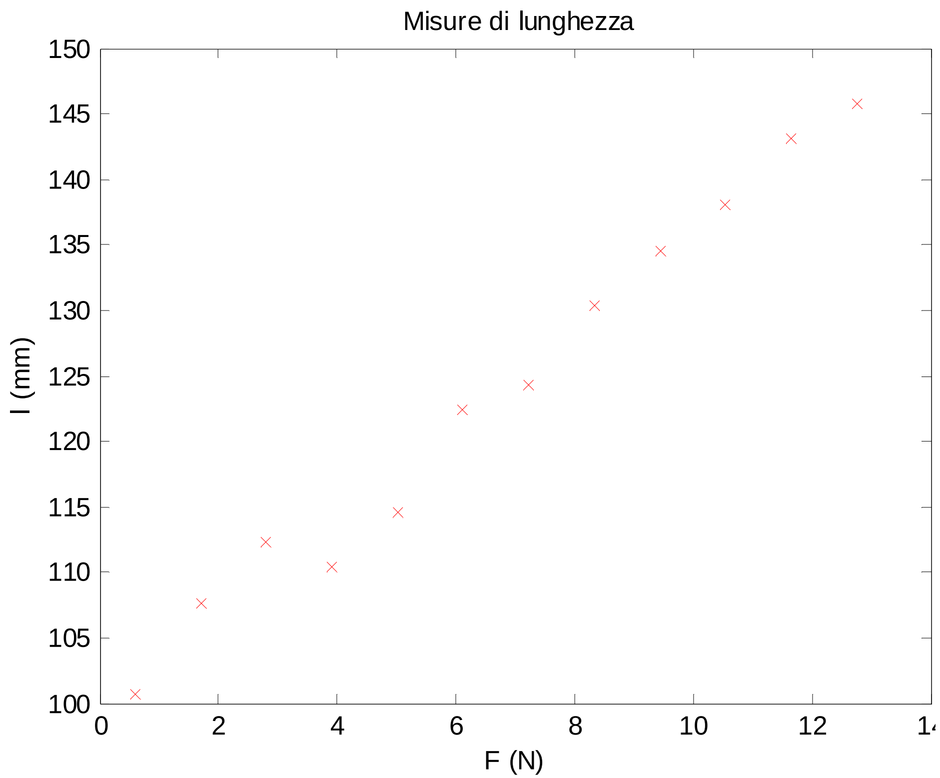
Il problema inverso - 3

In un mondo perfetto, in assenza di errori di misura, gli *m* punti sono esattamente su una linea. L'equazione delle retta può essere determinata con i metodi della geometria analitica, usando due degli *m* punti disponibili.



Il problema inverso - 4

Nel mondo reale le misure di lunghezza della molla contengono errori acciden-tali e la loro rappresentazione è del tipo



Non esiste una retta che passa per gli *m* punti.

Il problema inverso - 5

Possiamo scegliere due degli *m* punti e usare la geometria analitica? Si tratta di una spreco di informazione, che può condurre a risultati di bassa qualità.

Vogliamo usare tutta l'informazione contenuta degli *m* punti.

Interpretazione geometrica. Dati gli *m* punti del piano (6) bisogna trovare la retta che passa per essi (retta interpolante). Ma i punti (6), pur rappresentando un fenomeno lineare, a causa degli errori con cui la grandezza *y* viene misurata, non sono allineati in senso geometrico. Dunque non esiste, in senso geometrico classico, la retta passante per i punti.

Il problema inverso - 6

La cosa può essere analizzata in termini algebrici.

Si conosce anche la matrice $\mathbf{A}$ che è nota in quanto si conoscono i valori x_i in corrispondenza dei quali le osservazioni sono state effettuate. E' invece incognito il valore del vettore $\mathbf{X}$. Si cerca di dare una stima $\hat{\mathbf{X}}$ delle incognite

Interpretazione algebrica: si tratta di risolvere il sistema lineare

$$\mathbf{Y}_0 = \mathbf{A}\mathbf{X}$$

rispetto a $\mathbf{X}$. Ciò è impossibile con i metodi dell'algebra classica se il numero m delle osservazioni è maggiore del numero n delle incognite.

La soluzione generalizzata o ai MQ

Abbiamo dunque un vettore $\mathbf{Y}_0$ che contiene le grandezze osservate

$$\mathbf{Y}_0 = \begin{pmatrix} y_{01} \\ y_{02} \\ \vdots \\ y_{0m} \end{pmatrix}$$

(usiamo il termine generico *grandezze* perché la trattazione vale per la lunghezza della molla, l'allungamento di un oggetto e molti altri fenomeni) e vogliamo determinare il vettore delle incognite **X** come la soluzione del sistema

$$\mathbf{Y}_0 = \mathbf{A}\mathbf{X}$$

soluzione evidentemente non nel senso algebrico classico, ma generalizzato

La soluzione generalizzata o ai MQ - 2

Si considera il vettore

$$\mathbf{Y}(\mathbf{X}) = \mathbf{AX} \tag{7}$$

delle osservazioni stimate, cioè la funzione che, a partire da un ipotetico vettore $\mathbf{X}$ delle soluzioni, determina quali sarebbero le osservazioni corrispondenti.

Si considera il vettore differenza o *vettore degli scarti*

$$\begin{aligned} \mathbf{v}(\mathbf{X}) &= \mathbf{Y}_0 - \mathbf{Y}(\mathbf{X}) \\ &= \mathbf{Y}_0 - \mathbf{AX} \end{aligned} \tag{8}$$

Si determina la stima delle incognite $\hat{\mathbf{X}}$ in modo che il vettore differenza o *vetto-re degli scarti* $\mathbf{v}(\mathbf{X})$ abbia norma minima.

Nel caso in questione minimizzare la norma o il suo quadrato è equivalente (vale in generale?) e la soluzione ai MQ è definita come

La soluzione generalizzata o ai MQ – 3

La norma di **U**(**X**)

$$\left\|\mathbf{v}(\mathbf{X})\right\|^2 = \left\|\mathbf{Y}_0 - \mathbf{Y}(\mathbf{X})\right\|^2 =$$

$$= \sum_{i=1}^{m} \left(\mathbf{Y}_{0i} - \mathbf{Y}(\mathbf{X})_i\right)^2$$

Definiamo la funzione G (goal)

$$G : \mathbb{R}^n \to \mathbb{R}$$

$$G(\mathbf{X}) = \left\|\mathbf{v}(\mathbf{X})\right\|^2 = \left\|\mathbf{Y}_0 - \mathbf{Y}(\mathbf{X})\right\|^2 = \sum_{i=1}^{m} \left(\mathbf{Y}_{0i} - \mathbf{Y}(\mathbf{X})_i\right)^2$$

La soluzione generalizzata o ai MQ – 4

Definizione della soluzione MQ

$$
\begin{aligned}
\hat{\mathbf{X}} &= \min_{\mathbf{X}} G(\mathbf{X}) = \\
&= \min_{\mathbf{X}} \left\| \mathbf{Y}_0 - \mathbf{Y}(\mathbf{X}) \right\|^2 = \\
&= \min_{\mathbf{X}} \sum_{i=1}^{m} \left(\mathbf{Y}_{0i} - \mathbf{Y}(\mathbf{X})_i \right)^2
\end{aligned}
\qquad (9)
$$

La soluzione generalizzata o ai MQ - 5

Le seconda delle (9) evidenzia come si tratti di minimizzare la somma dei quadrati delle componenti di un vettore e questo spiega il nome attribuito a tale metodo, detto dei *minimi quadrati*.

La soluzione del problema MQ lineare

La funzione da minimizzare

$$G(\mathbf{X}) = \left\| \mathbf{Y}_0 - \mathbf{Y}(\mathbf{X}) \right\|^2$$

è una funzione di più variabili a valori reali e può essere minimizzata con le u-suali tecniche della analisi, imponendo che il gradiente si annulli nel punto di minimo (condizione che in questo caso è necessaria e sufficiente). Svolgendo i calcoli si conclude

$$\hat{\mathbf{X}} = \left(\mathbf{A}^t \mathbf{A} \right)^{-1} \mathbf{A}^t\, \mathbf{Y}_0$$

dove $\hat{\mathbf{X}}$ rappresenta la stima ai MQ delle incognite. Il vettore

$$\hat{\mathbf{Y}} = \mathbf{A}\hat{\mathbf{X}} \qquad (10)$$

rappresenta le altezze stimate (*vere*), da confrontare con quelle misurate.

La soluzione del problema MQ lineare - 2

Il vettore degli scarti

$$\hat{\mathbf{v}} = \mathbf{Y}_0 - \hat{\mathbf{Y}}$$

rappresenta un primo importante elemento per valutare l'adattamento del modello (1) ai dati (6).

Il più generale problema di MQ lineari

Rapporti fra le grandezze osservate $\mathbf{Y}$ e quelle incognite $\mathbf{X}$

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$$

Soluzione MQ

$$\hat{\mathbf{X}} = \left(\mathbf{A}^t\mathbf{A}\right)^{-1}\mathbf{A}^t\left(\mathbf{Y}_0 - \mathbf{b}\right)$$

$$\hat{\sigma}_0^2 = \frac{\mathbf{v}^t\mathbf{v}}{m-n}$$

$\hat{\sigma}_0^2$ è una stima della varianza delle osservazioni $\mathbf{Y}$, equidi sperse

Dispersione dei parametri stimati: matrice di varianza-covarianza di $\hat{\mathbf{X}}$

$$\mathbf{C}_{\hat{\mathbf{x}}\hat{\mathbf{x}}} = \sigma_0^2\left(\mathbf{A}^t\mathbf{A}\right)^{-1}$$