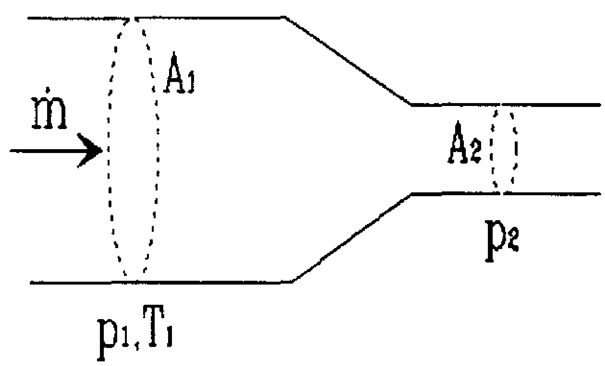
# MODELLI LINEARI E NON LINEARI NEI PARAMETRI

- Esempi di problemi di stima
- Definizione di modelli lineari e non lineari nei parametri
- Impostazione del problema di stima ai Minimi Quadrati

# FLUSSO DI UN GAS COMPRIMIBILE IN UN CONDOTTO




La portata d'aria $\dot{m}(t)$ è data da

$$\dot{m}(t) = c\ f(A_1, A_2, p_1(t), p_2(t), T_1(t))$$

dove $f(A_1, A_2, p_1(t), p_2(t), T_1(t))$ è una funzione nota e non lineare nei suoi argomenti, supposti noti o misurati, mentre

$$c = \text{coefficiente di efflusso}$$

è un parametro da stimare

E' evidente che il modello di $\dot{m}(t)$ è lineare nel parametro $c$, mentre è non lineare nelle grandezze note o misurate $p_1(t)$, $p_2(t)$, $T_1(t)$.

# FORMULAZIONE DEL PROBLEMA DI STIMA E SIMBOLOGIA

Supponiamo di aver effettuato $N$ prove di flussaggio in corrispondenza delle quali sono disponibili i valori delle grandezze $\dot{m}$, $p_1$, $p_2$, $T_1$.

Definiamo ora il vettore delle variabili dipendenti, o delle misure, $Y$, il vettore delle variabili indipendenti $U$ e il vettore $\Phi$ come segue:

$$Y = \begin{bmatrix} \dot{m}(1) \\ \dot{m}(2) \\ \dots \\ \dot{m}(N) \end{bmatrix} \quad , \quad U(i) = \begin{bmatrix} p_1(i) \\ p_2(i) \\ T_1(i) \end{bmatrix} \quad , \quad U = \begin{bmatrix} U(1) \\ U(2) \\ \dots \\ U(N) \end{bmatrix} \quad , \Phi(U) = \begin{bmatrix} f(U(1)) \\ f(U(2)) \\ \dots \\ f(U(N)) \end{bmatrix}$$

Pertanto, definendo con $\Theta$ il vettore dei parametri da stimare (in questo caso $\Theta=c$), risulta

$$Y = \Phi(U)\,\Theta$$

che evidenzia la dipendenza lineare di $Y$ da $\Theta$. Si noti che la dipendenza di $Y$ da $U$ è non lineare.

In realtà, anche se il modello fisico è corretto, la presenza di disturbi ed errori di misura fa sì che la relazione precedente non sia in generale verificata. Il problema di stima consiste quindi nel determinare un valore di $\Theta$ secondo un criterio opportuno.

# IL RENDIMENTO VOLUMETRICO η

In condizioni di regime la portata d'aria attraverso la valvola a farfalla $\dot{m}$ è uguale alla portata d'aria che passa attraverso la porta di alimentazione ai cilindri $\dot{m}_c$. Quest'ultima è data da

$$\dot{m}_c(t) = \frac{n(t)Vp_c(t)\eta}{120RT_c(t)}$$

dove

$n(t)$ = numero di giri dell'albero motore

$V$ = cilindrata del motore

$p_c(t)$ = pressione media nel collettore

$T_c(t)$ = temperatura media nel collettore

Supponendo di aver misurato $\dot{m}$, $n$, $p_c$, $T_c$, in un certo numero $N$ di condizioni stazionarie, e soprattutto ipotizzando che η sia non noto, ma costante (ipotesi poco realistica), l'equazione di bilancio di portata negli $N$ punti può essere riscritta come

$$Y = \Phi(U)\Theta$$

dove

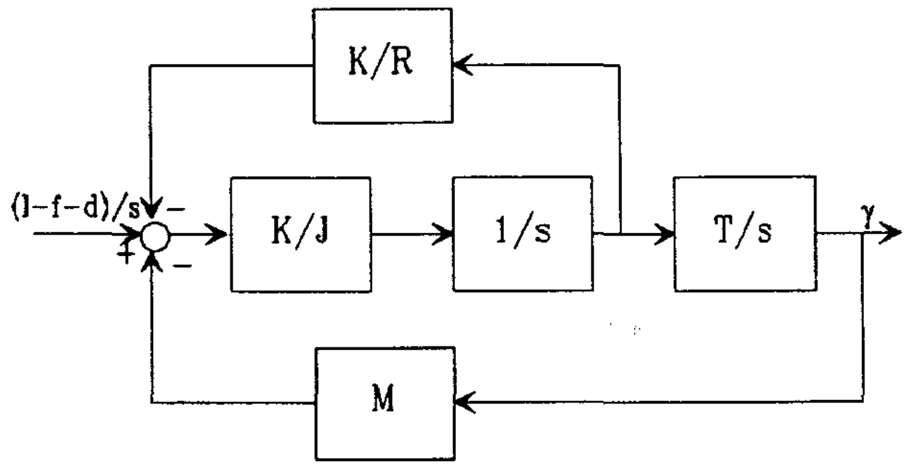
$$\Theta = \eta,\ Y = \begin{bmatrix} \dot{m}(1) \\ \dot{m}(2) \\ \cdots \\ \dot{m}(N) \end{bmatrix},\ U(i) = \begin{bmatrix} p_c(i) \\ T_c(i) \\ n(i) \end{bmatrix},\ U = \begin{bmatrix} U(1) \\ U(2) \\ \cdots \\ U(N) \end{bmatrix},\ \Phi(U) = \begin{bmatrix} \phi(U(1)) \\ \phi(U(2)) \\ \cdots \\ \phi(U(N)) \end{bmatrix},$$

$$\phi(U(i)) = \frac{n(i)Vp_c(i)}{120RT_c(i)}$$

**è anche questo un modello lineare nel parametro da stimare Θ**

# STIMA DEI PARAMETRI CARATTERISTICI DEL CORPO FARFALLATO VDO

Alcuni parametri caratteristici del corpo farfallato VDO sono stati stimati confrontando la risposta allo scalino del sistema reale con quella del modello descritto dallo schema a blocchi seguente (le coppie e forze sono trattate come disturbi in corrente e i momenti d'inerzia come masse)




*Parametri noti o misurati:*

$\gamma$ = posizione angolare

$T$ = rapporto di trasmissione

$R$ = resistenza di armatura

$I$ = corrente di armatura

$d$ = precarico delle molle

*Parametri da stimare:*

$K$ = costante di velocità

$J$ = momento di inerzia

$f$ = forza di attrito cinetico

$M$ = costante elastica delle molle

Posto

$$\omega = \sqrt{\frac{KTM}{J}} \quad , \quad \xi = \frac{1}{2R}\sqrt{\frac{K^3}{TMJ}}$$

l'andamento nel tempo della posizione angolare $\gamma$ nel tempo è dato da

$$\gamma(t) = \frac{I - f - d}{M}\left(1 - \frac{1}{\sqrt{1-\xi^2}} e^{-\xi\omega t} sin(\omega\sqrt{1-\xi^2}\,t + \arg(\xi + j\sqrt{1-\xi^2})\right) \quad (*)$$

Valutato allora il valore di $\gamma$ in $N$ istanti di tempo $t_1$, $t_2$, ..., $t_N$, e posto

$$Y = \begin{bmatrix} \gamma(t_1) \\ \gamma(t_2) \\ \dots \\ \gamma(t_N) \end{bmatrix} \quad , \quad \Theta = \begin{bmatrix} K \\ J \\ M \\ f \end{bmatrix} \quad , \quad U = \begin{bmatrix} t_1 \\ t_2 \\ \dots \\ t_N \end{bmatrix}$$

si può scrivere

$$Y = \Phi(U, \Theta)$$

dove ovviamente l'*i-esimo* elemento del vettore $\Phi \in R^N$ è dato dalla (*) valutata per $t = t_i$

**il modello è quindi non lineare nel vettore dei parametri $\theta$**

# ATTENZIONE

A volte il modello "sembra" non lineare nei parametri, ma basta una trasformazione per renderlo lineare

*Esempio*

Il modello del sistema è

$$y(t) = (\alpha + \beta u(t))^{\Theta}$$

con $\alpha$ e $\beta$ parametri noti e $\Theta$ da stimare. Si può anche scrivere

$$\ln y(t) = \ln(\alpha + \beta u(t))^{\Theta} = \Theta \ln(\alpha + \beta u(t))$$

e il modello diventa lineare in $\Theta$. A questo punto si possono definire i vettori

$$Y = \begin{bmatrix} \ln y(1) \\ \ln y(2) \\ \dots \\ \ln y(N) \end{bmatrix} \quad , \quad \Phi(\mathrm{U}) = \begin{bmatrix} \ln(\alpha + \beta \mathrm{u}(1)) \\ \ln(\alpha + \beta \mathrm{u}(2)) \\ \dots \\ \ln(\alpha + \beta \mathrm{u}(\mathrm{N})) \end{bmatrix}$$

e il modello lineare nel parametro

$$Y = \Phi(U)\,\Theta$$

# REGRESSIONE LINEARE

Supponiamo che un parametro $y$ dipenda da $q$ variabili indipendenti $u_1$, $u_2$, ..., $u_q$, (ad esempio il rendimento volumetrico $\eta$ dipende dal numero di giri $n$, dalla pressione media nel collettore $p_c$, dalla temperatura nel collettore $T_c$). Se la forma funzionale della dipendenza non è nota da considerazioni di tipo fisico, in prima approssimazione si può ipotizzare una dipendenza lineare del tipo

$$y = \theta_1 u_1 + \theta_2 u_2 + \ldots + \theta_q u_q$$

dove i parametri $\theta_i$ sono da stimarsi.

Se sono disponibili $N$ misure della variabile dipendente e delle variabili indipendenti, definendo i vettori

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ \ldots \\ y(N) \end{bmatrix} \quad , \quad U = \left[u_1(1) \ldots u_1(N)\ u_2(1) \ldots u_q(N)\right]' \quad , \quad \Theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \ldots \\ \theta_q \end{bmatrix}$$

e la matrice

$$\Phi(U) = \begin{bmatrix} u_1(1) & u_2(1) & \ldots & u_q(1) \\ u_1(2) & u_2(2) & \ldots & u_q(2) \\ \ldots & \ldots & \ldots & \ldots \\ u_1(N) & u_2(N) & \ldots & u_q(N) \end{bmatrix}$$

vale la relazione

$$Y = \Phi(U)\,\Theta$$

## CONCLUSIONI

- E' (ovviamente) fondamentale definire il problema di stima: le variabili dipendenti ($Y$), le variabili indipendenti ($U$), i parametri da stimare ($\Theta$)

- La non linearità del modello va valutata rispetto a $\Theta$, non a $U$

- Se possibile è meglio (come si vedrà in seguito) ricondursi al caso di modelli lineari in $\Theta$

# IL METODO DEI MINIMI QUADRATI

## (Least Squares - LS)

- definizione del problema di stima per modelli lineari
- soluzione del problema
- caratteristiche probabilistiche dello stimatore
- stima della varianza del disturbo
- Minimi Quadrati ponderati
- Stimatore di Markov (minimizza la varianza della stima)
- Minimi Quadrati con vincoli di uguaglianza

Supponiamo di avere un modello lineare nei $q$ parametri raccolti nel vettore

$$\Theta = \begin{vmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_q \end{vmatrix}$$

e di avere $N$ rilevazioni della variabile dipendente $y$, elementi del vettore

$$Y = \begin{vmatrix} y(1) \\ y(2) \\ \dots \\ y(N) \end{vmatrix}$$

e delle variabili indipendenti $u_1$, $u_2$, , $u_m$ raccolte nel vettore $U$ di dimensione opportuna. Come visto, è possibile scrivere la relazione

$$Y = \Phi(U)\Theta$$

dove $\Phi(U)$, o $\Phi$ per brevità, è una matrice

$$\Phi(U) = \begin{vmatrix} \phi_1(U) \\ \phi_2(U) \\ \dots \\ \phi_N(U) \end{vmatrix}$$

chiamata *matrice di sensitività*.

Supponiamo inoltre di avere almeno tante equazioni quanti sono i parametri, cioè $q \leq N$. Per $N > q$ è in generale impossibile trovare un vettore di parametri $\Theta$ che soddisfi la relazione $Y = \Phi\Theta$ per la presenza di errori di modello, disturbi, errori di misura.

## *Il problema di stima*

Definendo allora il vettore errore $\varepsilon$ (a $N$ componenti $\varepsilon(i)$)

$$\varepsilon = Y - \Phi\Theta$$

lo stimatore ai minimi quadrati è quello che minimizza la cifra di merito

$$J^{LS}(\Theta) = \varepsilon'\varepsilon = (Y - \Phi\Theta))'(Y - \Phi\Theta)) = \sum_{i=1}^{N} \varepsilon(i)^2 = \sum_{i=1}^{N} (y(i) - \phi_i \Theta)^2$$

cioè la somma dei quadrati dell'errore nelle $N$ rilevazioni disponibili.

## *Risultato*

Lo stimatore ai minimi quadrati $\Theta^{LS}$ si ottiene risolvendo rispetto a $\Theta$ il sistema di equazioni normali

$$(\Phi'\Phi)\Theta = \Phi' Y$$

e, se la matrice $(\Phi'\Phi)$ è non singolare, la stima ai minimi quadrati è data da

$$\Theta^{LS} = (\Phi'\Phi)^{-1} \Phi' Y$$

### *Dimostrazione*

Dalla definizione risulta

$$J_{LS}(\Theta)=Y'Y-2Y'\Phi\Theta+\Theta'\Phi'\Phi\Theta$$

questa è una forma quadratica in $\Theta$ semidefinita positiva. Il minimo si ottiene uguagliando a zero la derivata prima di $J_{LS}(\Theta)$ rispetto a $\Theta$.

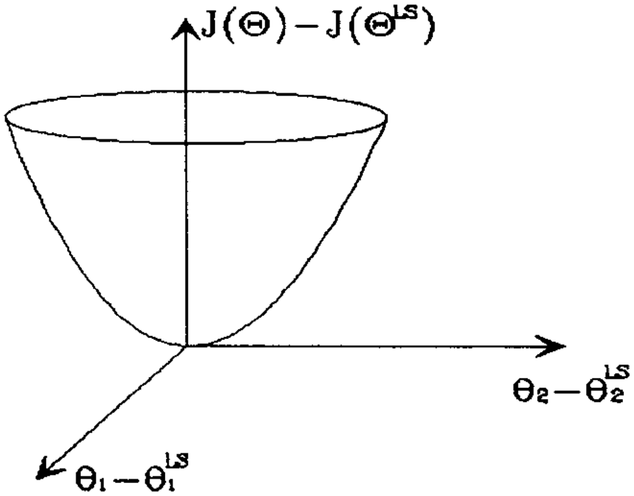
$$\partial J_{LS}(\Theta)/\partial\Theta=-2Y'\Phi+2\Theta'\Phi'\Phi=0$$

da cui segue il risultato.

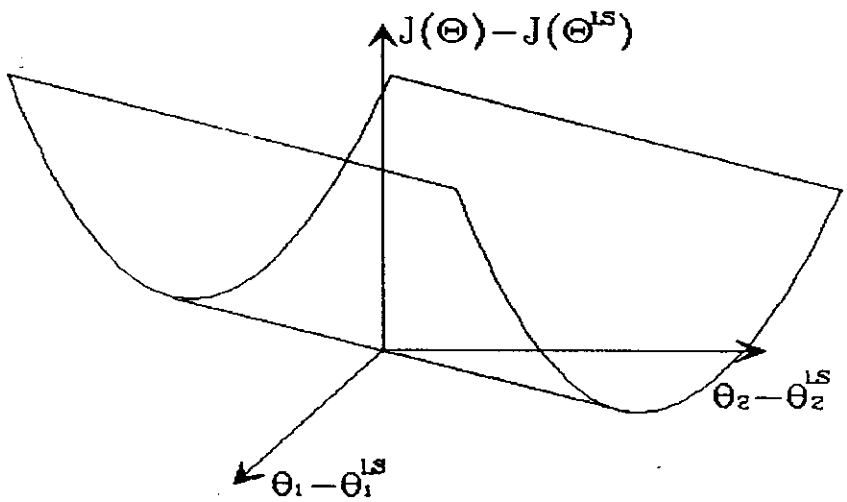
# INTERPRETAZIONE GRAFICA DELLA CONDIZIONE SUL DETERMINANTE DI $(\Phi'\Phi)$

*caso q=2*

$\det(\Phi'\Phi) \neq 0$




$\det(\Phi'\Phi) = 0$



## STIMA DELLA PORTATA D'ARIA ATTRAVERSO LA VALVOLA A FARFALLA

- E' un caso particolare di flusso attraverso un condotto e risulta

$$\dot{m}(t) = c\,f(A(a), p, T)$$

$c$ = coefficiente di efflusso (incognito)

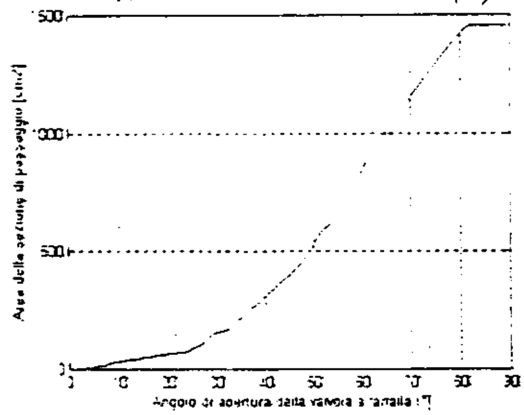
$A$ = area di passaggio

$\alpha$ = angolo di apertura della valvola a farfalla

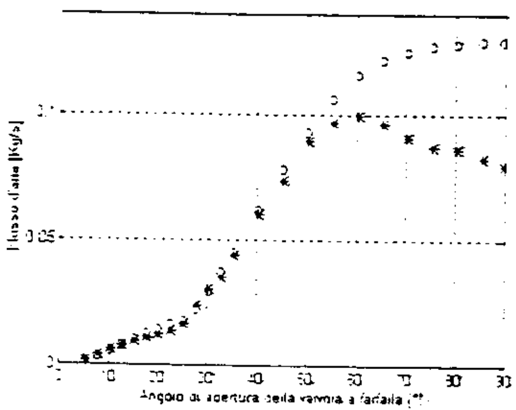
$p$ = pressione nella sezione di passaggio

$T$ = temperatura

- con considerazioni geometriche si ricava $A(\alpha)$



- da misure della portata (considerando solo condizioni di flusso sonico) si stima $c$ con il metodo dei minimi quadrati



- l'errore per alti valori di $p$ è dovuto al fatto che si usa un valore impreciso di $p$

# IL PROBLEMA DELL'IDENTIFICABILITA'

Poiché si è ipotizzato $N \geq q$ e $\Phi \in R^{N \cdot q}$, la condizione $\det(\Phi'\Phi) = 0$ è equivalente alla condizione

$$rango(\Phi) < q$$

Vediamo con un semplice esempio come interpretare questa condizione. Supponiamo $q=3$ e studiamo il problema della stima dei coefficienti della regressione lineare

$$y = \theta_1 u_1 + \theta_2 u_2 + \theta_3 u_3$$

Risulta

$$\Phi = \begin{vmatrix} u_1(1) & u_2(1) & u_3(1) \\ u_1(2) & u_2(2) & u_3(2) \\ \dots & \dots & \dots \\ \dots & \dots & \dots \\ u_1(N) & u_2(N) & u_3(N) \end{vmatrix}$$

Se $rango(\Phi)=2$, una colonna è linearmente dipendente dalle altre. Ad esempio

$$u_3(i) = au_1(i) + bu_2(i) \quad , \quad i=1,\dots,N$$

quindi

$$y = (\theta_1 + a\theta_3)u_1 + (\theta_2 + b\theta_3)u_2$$

**$u_3$ è inutile!**

Si può infatti definire un nuovo vettore dei parametri

$$\overline{\Theta} = \begin{vmatrix} \overline{\theta}_1 \\ \overline{\theta}_2 \end{vmatrix} = \begin{vmatrix} \theta_1 + a\theta_3 \\ \theta_2 + b\theta_3 \end{vmatrix}$$

e identificare un modello più semplice per il quale la condizione di identificabilità $\det(\Phi'\Phi) \neq 0$ è soddisfatta. Ciò significa che vi sarà un'unica stima di $\overline{\theta}_1$ e $\overline{\theta}_2$.

*Quando* $\det(\Phi'\Phi) = 0$?

- Se gli ingressi sono fissati, o il modello è sovraparametrizzato rispetto al fenomeno che si vuole descrivere (ad esempio si sono inserite nella regressione lineare due variabili indipendenti che hanno lo stesso effetto sulla variabile dipendente), o è sovraparametrizzato rispetto all'informazione contenuta nei dati disponibili.

- Se gli ingressi sono manipolabili e non si è nella prima delle condizioni precedenti, è necessario scegliere il valore delle variabili indipendenti in modo i dati utilizzati nella stima siano "il più possibile" ricchi di informazione sui parametri da stimare.

Queste considerazioni aprono due tematiche di notevole importanza:

***Determinazione della struttura del modello***

***Progetto dell'esperimento di stima***

entrambe verranno analizzate nel seguito.

# CARATTERISTICHE PROBABILISTICHE DELLA STIMA LS

Supponiamo che la sorgente casuale dei dati (il "modello vero") sia descritto dalle equazioni

$$Y = \Phi\Theta + V$$

dove $Y$, $\Phi$, $\Theta$ hanno il significato precedente e

$$V = \begin{vmatrix} v(1) \\ v(2) \\ \dots \\ v(N) \end{vmatrix}$$

L'ipotesi di lavoro è che $V \sim G(0, \Sigma)$ con $\Sigma$ matrice diagonale, cioè si assume che gli errori di misura o di modello siano incorrelati tra loro. Si osservi che l'aver posto il valore atteso nullo rappresenta il fatto che il sensore, benché affetto da disturbi di misura, è non polarizzato.

Si dimostra che

$$\Theta^{LS} = \Theta + (\Phi'\Phi)^{-1}\Phi' V$$

e

**Lo stimatore $\Theta^{LS}$ è non polarizzato ($E[\Theta^{LS}] = \Theta$)**

*Dimostrazione*

$$\Theta^{LS}=(\Phi'\Phi)^{-1}\Phi'Y=(\Phi'\Phi)^{-1}\Phi'(\Phi\Theta+V)=\Theta+(\Phi'\Phi)^{-1}\Phi'V$$

L'errore di stima è quindi

$$\Theta^{LS}-\Theta=(\Phi'\Phi)^{-1}\Phi'V$$

Da cui segue che

$$E[\Theta^{LS}-\Theta]=E[(\Phi'\Phi)^{-1}\Phi'V]=(\Phi'\Phi)^{-1}\Phi'E[V]=0$$

## CALCOLO DELLA VARIANZA DELLE STIME

Per quanto riguarda la varianza dello stimatore, facciamo l'ulteriore ipotesi

$$\Sigma=\sigma^2 I$$

con $I$ matrice identità di dimensioni $(N,N)$.

Dai risultati precedenti risulta

$$Var(\Theta^{LS}) = \sigma^2(\Phi'\Phi)^{-1}$$

La varianza della stima dipende dunque dalla varianza del disturbo e dalla matrice $\Phi'\Phi$, cioé dal valore assunto dalle variabili indipendenti nell'esperimento di identificazione. Ancora una volta questo va progettato con cura per rendere $(\Phi'\Phi)^{-1}$ "piccola".

Si osservi tuttavia che per valutare $Var(\Theta^{LS})$ è necessario conoscere $\sigma^2$, a priori non nota. E' quindi importante poterla stimare dai dati.

*Dimostrazione*

$$Var(\Theta^{LS}) = E[(\Theta^{LS}-E[\Theta^{LS}])\,(\Theta^{LS}-E[\Theta^{LS}])']= E[(\Theta^{LS}-\Theta)\,(\Theta^{LS}-\Theta)'] =$$

$$= E[((\Phi'\Phi)^{-1}\Phi'V)((\Phi'\Phi)^{-1}\Phi'V)'] = (\Phi'\Phi)^{-1}\Phi'E[VV']\Phi(\Phi'\Phi)^{-1} =$$

$$= \sigma^2(\Phi'\Phi)^{-1}\Phi'\Phi(\Phi'\Phi)^{-1} = \sigma^2(\Phi'\Phi)^{-1}$$

# STIMA DELLA VARIANZA DEL DISTURBO

Per stimare $\sigma^2$, determiniamo dapprima il valore della cifra di merito valutata per $\Theta = \Theta^{LS}$, cioè di $J(\Theta^{LS})$. Si noti che tale valore può essere determinato numericamente senza difficoltà una volta effettuata la stima di $\Theta^{LS}$.

Si ha

$$J(\Theta^{LS}) = Y'(I - \Phi(\Phi'\Phi)^{-1}\Phi')Y$$

e risulta

$$E[J(\Theta^{LS})] = (N-q)\sigma^2$$

Perciò

$$\hat{\sigma}^2 = \frac{J(\Theta^{LS})}{N-q}$$

che è la stima cercata.

*Dimostrazione*

Il valore di $J(\Theta^{LS})$ è

$$J(\Theta^{LS})=(Y-\Phi\Theta^{LS})'(Y-\Phi\Theta^{LS})'=$$
$$=(Y-\Phi(\Phi'\Phi)^{-1}\Phi'Y)'(Y-\Phi(\Phi'\Phi)^{-1}\Phi'Y)=Y'(I-\Phi(\Phi'\Phi)^{-1}\Phi')'(I-\Phi(\Phi'\Phi)^{-1}\Phi')Y=$$
$$=Y'(I-\Phi(\Phi'\Phi)^{-1}\Phi')Y$$

Ricordando che $Y=\Phi\Theta+V$, si può inoltre scrivere che

$$J(\Theta^{LS})=(\Phi\Theta+V)'(I-\Phi(\Phi'\Phi)^{-1}\Phi')(\Phi\Theta+V)=V'(I-\Phi(\Phi'\Phi)^{-1}\Phi')V$$

Definiamo ora la matrice $D=(I-\Phi(\Phi'\Phi)^{-1}\Phi')$ e notiamo che $J(\Theta^{LS})=tr(V'DV)=tr(VV'D)$, da cui

$$E[J(\Theta^{LS})]=E[tr(VV'D)]=tr(E[VV'D])=\sigma^2 tr(D)$$

da questa espressione si conclude che una stima di $\sigma^2$ si ottiene da

$$\hat{\sigma}^2=\frac{J(\Theta^{LS})}{tr(D)}$$

resta quindi da determinare il valore di $tr(D)$. Ricordando la definizione $D=(I-\Phi(\Phi'\Phi)^{-1}\Phi')$, in cui $I$ è una matrice $(N,N)$, si ha

$$tr(D)=tr(I-\Phi(\Phi'\Phi)^{-1}\Phi')=tr(I)-tr(\Phi(\Phi'\Phi)^{-1}\Phi')=N-tr((\Phi'\Phi)^{-1}\Phi'\Phi)=N-q$$

e in definitiva

$$\hat{\sigma}^2=\frac{J(\Theta^{LS})}{N-q}$$

# RIASSUNTO SUL METODO LS PER PROBLEMI DI REGRESSIONE LINEARE

Richiamiamo i principali risultati visti: dato il modello di regressione lineare

$$y = \theta_1 u_1 + \theta_2 u_2 + \dots + \theta_q u_q$$

e le $N$ misure della variabile dipendente e delle variabili indipendenti, raccolte nel vettore $\Theta$ e nella matrice $\Phi$

$$Y = \begin{vmatrix} y(1) \\ y(2) \\ \dots \\ y(N) \end{vmatrix} \quad , \quad \Phi(U) = \begin{vmatrix} u_1(1) & u_2(1) & \dots & u_q(1) \\ u_1(2) & u_2(2) & \dots & u_q(2) \\ \dots & \dots & \dots & \dots \\ u_1(N) & u_2(N) & \dots & u_q(N) \end{vmatrix}$$

la stima $LS$ del vettore $\Theta = \begin{vmatrix} \theta_1 \\ \theta_2 \\ \dots \\ \theta_q \end{vmatrix}$ è data da

$$\Theta^{LS} = (\Phi'\Phi)^{-1}\Phi'Y$$

la stima della varianza del disturbo può essere stimata con la formula

$$\hat{\sigma}^2 = \frac{J(\Theta^{LS})}{N-q}$$

e la varianza della stima è valutabile come

$$Var(\Theta^{LS}) = \Lambda = \hat{\sigma}^2(\Phi'\Phi)^{-1}$$

In molti casi il modello identificato è scritto nella forma

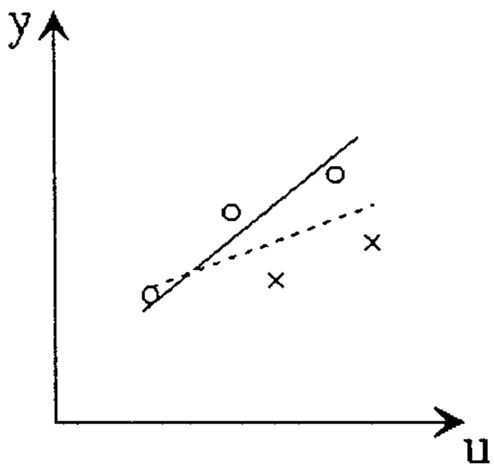
$$y = \theta_1^{LS}\,[\lambda_1]\,u_1 + \theta_2^{LS}\,[\lambda_2]\,u_2 + \ldots + \theta_q^{LS}\,[\lambda_q]\,u_q + v\,[\sigma]$$

dove i $\lambda_i$, $i=1,\ldots,N$, sono le radici quadrate degli elementi sulla diagonale della matrice $\Lambda$ (deviazioni standard). E' bene però ricordare le ipotesi sul disturbo che hanno consentito di pervenire alla formula per la stima della varianza del disturbo e a quella della varianza delle stime. L'attendibilità di queste ipotesi può essere verificata da un'analisi sui dati.

# I MINIMI QUADRATI PONDERATI (WLS)

Supponiamo di avere a disposizione alcuni dati affidabili ed altri meno. Ciò può succedere ad esempio perché alcune misure sono state fatte con maggiore accuratezza o con sensori più precisi. Nella determinazione della cifra di merito da minimizzare è naturale cercare di dare maggior peso agli errori $\varepsilon(i)$ relativi ai dati più affidabili.

*Esempio - regressione lineare*




i pallini corrispondono a misure precise, le crocette no. La linea continua corrisponde al modello ricavato considerando solo (o soprattutto) i dati precisi, la linea tratteggiata al modello stimato impiegando allo stesso modo tutti i dati.

Sembra sensato minimizzare il funzionale

$$J^{WLS}(\Theta) = \sum_{i=1}^{N} q(i)\varepsilon(i)^2 = \sum_{i=1}^{N} q(i)(y(i) - \phi_i\Theta)^2$$

dove i $q(i)$, $i=1,\ldots,N$ sono i pesi attribuiti ad ogni errore.

In forma matriciale, definendo la matrice diagonale

$$Q = \begin{vmatrix} q(1) & 0 & \dots & 0 \\ 0 & q(2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & q(N) \end{vmatrix}$$

la cifra di merito si può scrivere come

$$J^{WLS}(\Theta) = \varepsilon' Q \varepsilon$$

Ripetendo gli sviluppi precedenti si conclude facilmente che lo stimatore ottimo è dato da

$$\Theta^{WLS} = (\Phi' Q \Phi)^{-1} \Phi' Q Y$$

inoltre, sempre supponendo che la sorgente casuale dei dati (il "modello vero") sia descritto da $Y = \Phi\Theta + V$ e che le ipotesi precedenti su $V$ siano soddisfatte, si possono dedurre le seguenti proprietà

$$E[\Theta^{WLS}] = \Theta$$

$$Var(\Theta^{WLS}) = (\Phi' Q \Phi)^{-1} \Phi' Q \Sigma Q \Phi (\Phi' Q \Phi)^{-1}$$

# BEST LINEAR UNBIASED ESTIMATE (BLUE) o STIMA DI MARKOV

Ci si può chiedere quale sia una scelta opportuna negli WLS della matrice di ponderazione $Q$. Una possibilità è quella di scegliere $Q$ in modo da minimizzare $Var(\Theta^{WLS})$. Con questa scelta, si può dimostrare che risulta

$$Q = \Sigma^{-1}$$

corrispondentemente

$$\Theta^{WLS} = (\Phi' \Sigma^{-1} \Phi)^{-1} \Phi' \Sigma^{-1} Y$$

e

$$Var(\Theta^{WLS}) = (\Phi' \Sigma^{-1} \Phi)^{-1}$$

*Un esempio*

Supponiamo di dover stimare una costante $\Theta$ di cui sono note $N$ misure $y$ affette da incertezza

$$y(i)=\Theta+v(i) \quad , \quad i=1,\dots,N$$

Posto $\Phi = \begin{vmatrix} 1 \\ 1 \\ \dots \\ 1 \end{vmatrix}$ e applicando il metodo $LS$ la stima ottenuta è

$$\Theta^{LS}=(y(1)+y(2)+\dots+y(N))/N$$

che corrisponde alla media aritmetica

Supponendo invece che gli errori di misura siano indipendenti, a valore atteso nullo e con diverse varianze, cioè che

$$E[v(i)]=0, \quad Var(v(i))=\sigma^2(i), \quad i=1,\dots,N$$

ponendo $Q = \begin{vmatrix} 1/\sigma^2(1) & 0 & \dots & 0 \\ 0 & 1/\sigma^2(2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1/\sigma^2(N) \end{vmatrix}$

la BLUE di $\Theta$ è

$$\Theta^{WLS} = \frac{1}{\sum_{i=1}^{N}(1/\sigma^2(i))} \sum_{j=1}^{N} y(j)/\sigma^2(j)$$

che corrisponde alla media ponderata delle misure disponibili.

## MINIMI QUADRATI CON VINCOLI DI UGUAGLIANZA

In alcuni casi è necessario risolvere problemi di stima LS in presenza di vincoli di uguaglinza. Ad esempio nel problema della determinazione dei parametri nella regressione lineare

$$y = \theta_1 u_1 + \theta_2 u_2 + \ldots + \theta_q u_q$$

si può voler minimizzare la consueta cifra di merito

$$J^{LS}(\Theta) = \varepsilon'\varepsilon = (Y-\Phi\Theta))'(Y-\Phi\Theta)) = \sum_{i=1}^{N}\varepsilon(i)^2 = \sum_{i=1}^{N}(y(i)-\phi_i\Theta)^2$$

dove i simboli hanno il consueto significato e con in aggiunta i vincoli

$$\begin{gathered}
\gamma_{11}\theta_1 u_1(1) + \gamma_{12}\theta_2 u_2(1) + \ldots + \gamma_{1q}\theta_q u_q(1) = \delta(1) \\
\gamma_{21}\theta_1 u_1(2) + \gamma_{22}\theta_2 u_2(2) + \ldots + \gamma_{2q}\theta_q u_q(2) = \delta(2) \\
\ldots \\
\gamma_{m1}\theta_1 u_1(m) + \gamma_{m2}\theta_2 u_2(m) + \ldots + \gamma_{mq}\theta_q u_q(m) = \delta(m)
\end{gathered}$$

con $m<q$. Questi vincoli, definendo la matrice $\Gamma$ e il vettore $\Delta$ come

$$\Gamma = \begin{vmatrix} \gamma_{11}u_1(1) & \gamma_{12}u_2(1) & \ldots & \gamma_{1q}u_q(1) \\ \gamma_{21}u_1(2) & \gamma_{22}u_2(2) & \ldots & \gamma_{2q}u_q(2) \\ \ldots & \ldots & \ldots & \ldots \\ \gamma_{m1}u_1(m) & \gamma_{m2}u_2(m) & \ldots & \gamma_{mq}u_q(m) \end{vmatrix} \quad , \quad \Delta = \begin{vmatrix} \delta(1) \\ \delta(2) \\ \ldots \\ \delta(m) \end{vmatrix}$$

possono essere espressi nella forma matriciale

La soluzione del problema

$$\min J^{LS}(\Theta) = (Y-\Phi\Theta))'(Y-\Phi\Theta))'$$

con i vincoli

$$\Gamma\Theta=\Delta$$

si ottiene risolvendo rispetto al vettore delle incognite $\begin{vmatrix}\Theta\\ \mu\end{vmatrix}$ il sistema di equazioni lineari

$$\begin{vmatrix}2\Phi'\Phi & \Gamma' \\ \Gamma & 0\end{vmatrix} \begin{vmatrix}\Theta \\ \mu\end{vmatrix} = \begin{vmatrix}2\Phi' Y \\ \Delta\end{vmatrix}$$

Dove μ sono i moltiplicatori di Lagrange. Al risultato si perviene appunto mediante l'uso della tecnica dei moltiplicatori di Lagrange.

# CONCLUSIONI

- Nel caso di modelli lineari nei parametri la stima a minimi quadrati comporta la risoluzione di un sistema di equazioni lineari (le equazioni normali)

- Perché questo sistema sia risilubile restano aperti due problemi fondamentali: la determinazione della struttura del modello (qualora non sia noto da considerazioni di tipo fisico) e il progetto dell'esperimento di stima, cioè la generazione di un insieme di dati che consenta effettivamente la stima dei parametri)

- E' possibile calcolare la varianza delle stime, di fondamentale importanza per sapere che affidabilità hanno le stime stesse e per selezionare il modello più opportuno. Per far questo è necessaria la stima della varianza del disturbo

- Note le caratteristiche del disturbo è possibile (e facile) determinare lo stimatore che minimizza la varianza delle stime