Modelli Lineari

Corso di **Probabilità ed Inferenza**
a.a. 2009/2010 Secondo Periodo
Prof. Filippo DOMMA

Corso di Laurea Specialistica in
Economia Applicata
Facoltà di Economia – UniCal

Richiami di Algebra delle Matrici

Matrice. È una tabella rettangolare di elementi numeri disposti in righe e colonne. In particolare, se una matrice **A** costituita da n-righe e p-colonne diremo che la matrice **A** ha dimensione (nxp) – oppure è di ordine (nxp) – e scriveremo

$$\mathbf{A}_{(n\times p)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1p} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{np} \end{pmatrix} = \{a_{ij}\}$$

dove a_{ij} è l'elemento posto nella i-esima riga e nella j-esima colonna.

Casi particolari: sia **A** una matrice di dimensione (nxp).

- Se n=p allora **A** è detta matrice quadrata di dimensione (nxn).
- se p=1 oppure n=1 allora **A** è detta vettore colonna di dimensione (nx1) oppure vettore riga di dimensione (1xp), così definiti

$$\mathbf{a}_{(n\times 1)} = \begin{pmatrix} a_1 \\ a_2 \\ \ldots. \\ a_n \end{pmatrix} \qquad \mathbf{a}_{(1\times p)} = \begin{pmatrix} a_1 & a_2 & \ldots. & a_p \end{pmatrix}$$

- Se poniamo n=1 e a_{ij}=1 per ogni j=1,…,p nella matrice **A**, otteniamo il vettore riga unitario di dimensione (1xp)

$$\mathbf{1}_{(1\times p)} = \begin{pmatrix} 1 & 1 & \ldots. & 1 \end{pmatrix}$$

Casi particolari di matrici quadrate: sia **A** una matrice di dimensione (nxn).

- se a_{ij}=0 per ogni i≠j ed almeno un elemento a_{ii} diverso da zero allora **A** è detta <u>diagonale</u>

$$\mathbf{A}_{(n\times n)} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

- se a_{ij}=0 per ogni i≠j e a_{ii}=1 per ogni i=1,…,n allora **A** è detta <u>identità</u>

$$\mathbf{I}_{(n\times n)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

- se $a_{ij}=a_{ji}$ per ogni i,j=1,…,n allora **A** è detta <u>simmetrica</u>

$$\mathbf{A}_{(n\times n)} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix}$$

- se $a_{ij}=1$ per ogni i,j=1,…,n allora **A** è detta <u>matrice unità</u>

$$\mathbf{A}_{(n\times n)} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

Matrice Trasposta

Data una matrice $\mathbf{A}_{(nxp)}$, si definisce trasposta di **A**, e si indica con **A'**, la matrice di dimensione (pxn) ottenuta scambiando le righe con le colonne, cioè

$$\mathbf{A}'_{(nxp)} = \mathbf{A}_{(p\times n)} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1p} & a_{2p} & \dots & a_{np} \end{pmatrix}$$

Osservazione 1: $(\mathbf{A}')' = \mathbf{A}$

Osservazione 2: se **A** è simmetrica allora $\mathbf{A}' = \mathbf{A}$.

Somma e Differenza

Siano A e B due matrici di dimensione (nxp).

- $$\mathbf{A}_{(nxp)} + \mathbf{B}_{(nxp)} = \mathbf{C}_{(nxp)} = \left\{ c_{ij} \right\} = \left\{ a_{ij} + b_{ij} \right\}$$

- $$\mathbf{A}_{(nxp)} - \mathbf{B}_{(nxp)} = \mathbf{C}_{(nxp)} = \left\{ c_{ij} \right\} = \left\{ a_{ij} - b_{ij} \right\}$$

Prodotto

$$\lambda \mathbf{A}_{(nxp)} = \left\{ \lambda a_{ij} \right\}$$ se λ è uno scalare

Prodotto tra matrici

$$\mathbf{A}_{(nxp)}\mathbf{B}_{(pxm)} = \mathbf{C}_{(nxm)}$$

$$\mathbf{A}_{(nxp)}\mathbf{B}_{(pxm)} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{12} & a_{22} & \dots & a_{2p} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{p1} & b_{p2} & \dots & b_{pm} \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{h=1}^{p} a_{1h}b_{h1} & \sum_{h=1}^{p} a_{1h}b_{h2} & \dots & \sum_{h=1}^{p} a_{1h}b_{hm} \\ \sum_{h=1}^{p} a_{2h}b_{h1} & \sum_{h=1}^{p} a_{2h}b_{h2} & \dots & \sum_{h=1}^{p} a_{2h}b_{hm} \\ \dots & \dots & \dots & \dots \\ \sum_{h=1}^{p} a_{nh}b_{h1} & \sum_{h=1}^{p} a_{nh}b_{h2} & \dots & \sum_{h=1}^{p} a_{nh}b_{hm} \end{pmatrix} = \left\{ \sum_{h=1}^{p} a_{ih}b_{hj} \right\}$$

Prodotto tra un vettore ed una matrice

$$\mathbf{x}_{(1xn)}\mathbf{A}_{(nxp)} = \mathbf{b}_{(1xp)} \qquad \text{oppure} \qquad \mathbf{A}_{(nxp)}\mathbf{x}_{(px1)} = \mathbf{b}_{(nx1)}$$

$$\mathbf{x}_{(1xn)}\mathbf{A}_{(nxp)} = \left(x_{11}, x_{12}, \ldots, x_{1n}\right)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{12} & a_{22} & \ldots & a_{2p} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix} =$$

$$= \left(\sum_{h=1}^{n} x_{1h}a_{h1}, \sum_{h=1}^{n} x_{1h}a_{h2}, \ldots, \sum_{h=1}^{n} x_{1h}a_{hp} \right)$$

Prodotto tra vettori

$$\mathbf{x}_{(nx1)}\mathbf{y}_{(1xm)} = \mathbf{A}_{(nxm)} \quad \text{oppure} \quad \mathbf{x}_{(1xp)}\, y_{px1} = \lambda_{(1x1)}$$

$$\mathbf{x}_{(nx1)}\mathbf{y}_{(1xm)} = \begin{pmatrix} x_{11} \\ x_{21} \\ \dots \\ x_{n1} \end{pmatrix} \left(y_{11}, y_{12}, \dots, y_{1m} \right) = \begin{pmatrix} x_{11}\, y_{11} & x_{11}\, y_{12} & \dots & x_{11}\, y_{1m} \\ x_{21}\, y_{11} & x_{21}\, y_{12} & \dots & x_{21}\, y_{1m} \\ \dots & \dots & \dots & \dots \\ x_{n1}\, y_{11} & x_{n1}\, y_{12} & \dots & x_{n1}\, y_{1m} \end{pmatrix}$$

$$\mathbf{x}_{(1xp)}\, y_{px1} = \left(x_{11}, x_{12}, \dots, x_{1p} \right) \begin{pmatrix} y_{11} \\ y_{21} \\ \dots \\ y_{p1} \end{pmatrix} = \sum_{h=1}^{p} x_{1h}\, y_{h1}$$

Prodotto tra vettori: casi particolari

$$\mathbf{x}_{(nx1)}\left(\mathbf{x}'_{(nx1)}\right)=\begin{pmatrix} x_{11} \\ x_{21} \\ \cdots \\ x_{n1} \end{pmatrix}\left(x_{11}, x_{21}, \ldots, x_{n1}\right)=\begin{pmatrix} x_{11}^2 & x_{11}x_{21} & \ldots & x_{11}x_{n1} \\ x_{21}x_{11} & x_{21}^2 & \ldots & x_{21}x_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{n1}x_{11} & x_{n1}x_{21} & \ldots & x_{n1}^2 \end{pmatrix}$$

$$\left(\mathbf{x}'_{(nx1)}\right)\mathbf{x}_{(nx1)}=\mathbf{x}_{(1xn)}x_{nx1}=\left(x_{11}, x_{12}, \ldots, x_{1n}\right)\begin{pmatrix} x_{11} \\ x_{21} \\ \cdots \\ x_{n1} \end{pmatrix}=\sum_{h=1}^{p} x_{1h}^2$$

Osservazione 3: $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

Osservazione 4: $(\mathbf{A}+\mathbf{B})' = (\mathbf{A}'+\mathbf{B}')$

Osservazione 5: $(\lambda)' = \lambda$ se λ è uno scalare

Traccia di una matrice quadrata

E' la somma degli elementi posti sulla diagonale principale, cioè: $tr\left(\mathbf{A}_{(nxn)}\right) = \sum_{i=1}^{n} a_{ii}$

- $tr(\lambda) = \lambda$ se λ è uno scalare
- $tr(\lambda\mathbf{A}) = \lambda\, tr(\mathbf{A})$
- $tr(\mathbf{A} \pm \mathbf{B}) = tr(\mathbf{A}) \pm tr(\mathbf{B})$
- $tr(\mathbf{A}\,\mathbf{B}) = tr(\mathbf{BA})$
- $tr(\mathbf{A}\,\mathbf{B}\,\mathbf{C}) = tr(\mathbf{BCA}) = tr(\mathbf{BCA})$

Determinante

Sia **A** una matrice quadrata di dimensione (nxn). Si definisce determinante di **A** la quantità:

$$\left|\mathbf{A}\right| = \sum (-1)^k a_{1r(1)} \ldots\ldots a_{nr(1)}$$

dove la somma è estesa a tutte le permutazini r di (1,2,…,n).

Matrice Inversa

Sia **A** una matrice quadrata di dimensione (nxn), con determinante diverso dallo zero. Si definisce inversa di **A**, la matrice di dimensione (nxn), indicata con $\mathbf{A}^{-1}$, tale che

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

- è unica !
- se $\mathbf{D} = \mathbf{ABC} \Rightarrow \mathbf{D}^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$
- $\left|\mathbf{A}^{-1}\right| = \dfrac{1}{|\mathbf{A}|}$
- $(\mathbf{A}')^{-1} = \left(\mathbf{A}^{-1}\right)'$
- se A è simmetrica allora A^{-1} è simmetrica

Rango di una Matrice

Si definisce rango di una qualunque matrice **A** il numero di vettori colonna linearmente indipendenti, viene indicato con r(**A**).

- r(**A**) ≤min{n,m} dove *n* ed *m* sono le righe e le colonne di **A**.

- r(**AB**) ≤min{r(**A**), r(**B**)}

- se **A** è una matrice di dimensione (nxn), allora r(**A**)=n se e solo se il determinante di **A** è diverso dallo zero.

- se **A** è una matrice di dimensione (nxn) e se esiste $\mathbf{A}^{-1}$, allora r($\mathbf{A}^{-1}$)=r(**A**).

Matrici definite positive e semi-definite positive

Data una matrice simmetrica **A**.

- A è definita positiva se $\mathbf{x}'\mathbf{A}\mathbf{x}>0$ per ogni vettore **x** non nullo.
- A è semi-definita positiva se $\mathbf{x}'\mathbf{A}\mathbf{x}\geq 0$ per ogni vettore **x** non nullo.

Matrici Idempotente

La matrice quadrata **A** è idempotente se $\mathbf{AAA}\ldots\mathbf{A} = \mathbf{A}$.

- se **A** è simmetrica ed idempotente allora $r(\mathbf{A})=tr(\mathbf{A})$.
- Ogni matrice idempotente è semi-\definita positiva.

Fasi di costruzione di un Modello Statistico

Per costruire un modello statistico si seguono essenzialmente le seguenti tre fasi:

1. Specificazione
2. Stima
3. Diagnostica

Generalità

In molti campi della ricerca accade spesso di dover studiare le relazioni, eventualmente esistenti, tra le variabili analizzate. Gli economisti, ad esempio, potrebbero essere interessati a stabilire qual è la relazione esistente tra la produzione di un bene e i fattori produttivi, oppure la relazione esistente tra il PIL ed altre variabili macroeconomiche.
In generale, si tratta di studiare la relazione tra una variabile y, denominata variabile dipendente, ed un insieme di k variabili, X_1, X_2,…,X_k, denominate indipendenti (oppure esplicative o regressori). Formalmente abbiamo si ha:

$$y = f\left(X_1, X_2, \ldots, X_k\right) \longleftarrow \boxed{\text{Modello deterministico}}$$

Osserviamo che tra le variabili che rappresentano fenomeni del mondo reale non possiamo ipotizzare che possa esistere una perfetta relazione matematica del tipo $y=f(x_1, x_2,\ldots,x_k)$, perché in **primo luogo** la nostra conoscenza non verte su tutto $x_1, x_2,\ldots,x_k$ e su tutto y, ma si limita ad un campione di osservazioni su di essi estratti da una popolazione ipotetica. Tali valori riflettono solo con una certa probabilità la "vera" relazione che si può ipotizzare sussista per l'intera popolazione. In **secondo luogo**, le variabili del mondo reale sono spesso il risultato di fenomeni imprevedibili, errori di misurazione, scarti accidentali, decisioni individuali che contribuiscono ad alterate un legame deterministico. **Infine**, quando si afferma che il fenomeno Y "dipende" ovvero è "spiegato" dal fenomeno X, si semplifica certamente la dinamica reale ove intercorrono, con differenti pesi, una miriade di interrelazioni tra le variabili (Y,X) ed il resto del mondo dei fenomeni non-esplicitati nel modello deterministico.
Quindi, è necessario inserire nel modello deterministico una v.c. ε che riassume l'insieme degli effetti e delle cause che impediscono alla relazione tra la variabile dipendente e le variabili indipendenti di essere un legame teorico di tipo matematico. In tal modo si ottiene il cosiddetto modello statistico

$$y = f\left(X_1, X_2,\ldots, X_k\right) + \varepsilon \longleftarrow \boxed{\text{Modello Statistico}}$$

La v.c. ε qualifica il modello nel senso che se f(.) è una funzione matematica, allora sarà Y ad essere una variabile casuale risultante dalla somma della componente deterministica f(.) e di una componente stocastica ε.

La Specificazione

Specificazione del modello → forma funzionale di $f(.)$

Specificazione del modello → Ipotesi sulla v.c. ε

Le forme funzionali che consideriamo in questo corso sono:

-Lineari nei parametri, ad esempio

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + \varepsilon$$

$$y = a_0 + a_1 x_1 + a_2 x_2^2 + ... + a_k x_k^k + \varepsilon$$

$$y = c_0 + c_1 \ln(x_1) + c_2 e^{x_2} + ... + c_k x_k^{-1} + \varepsilon$$

....

-Linearizzabili a seguito di trasformazioni $y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} x_3^{\beta_3} ..... x_k^{\beta_k} \varepsilon$

$$\Rightarrow \quad \ln(y) = \ln(\beta_0) + \beta_1 \ln(x_1) + ... + \beta_k \ln(x_k) + \ln(\varepsilon)$$

Definizione ed Ipotesi Fondamentali

Definizione. Chiameremo modello lineare l'equazione

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon$$

dove:

y è un vettore casuale osservabile di dimensione (*nx1*);

X è una matrice di elementi noti di dimensione (*nxp*), con $p<n$, le cui colonne contengono le *n* osservazioni sulle variabili indipendenti (regressori);

β è un vettore di parametri incogniti di dimensione (*px1*);

ε è un vettore casuale di variabili non-osservabili di dimensione (*nx1*).

In particolare, si ha:

$$\mathbf{y}_{(n\times 1)} = \begin{pmatrix} y_1 \\ y_2 \\ \cdots\cdot \\ y_n \end{pmatrix} \qquad \mathbf{X}_{(n\times p)} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}$$

$$\boldsymbol{\beta}_{(p\times 1)} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \cdots\cdot \\ \beta_k \end{pmatrix} \qquad \boldsymbol{\varepsilon}_{(n\times 1)} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots\cdot \\ \varepsilon_n \end{pmatrix}$$

Per completare la <u>specificazione</u> del modello statistico è necessario effettuare delle ipotesi, dette **<u>ipotesi fondamentali</u>**, sulla v.c. **ε**.

1. $E(\varepsilon) = 0$

2. $$V(\varepsilon) = \boldsymbol{\Sigma} = \sigma^2 \mathbf{I} = \sigma^2 \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

La 2. racchiude in sé l'ipotesi di **omoschedasticità**, cioè

$$v(\varepsilon_i) = \sigma^2 \qquad i=1,\dots,n$$

e di **incorrelazione** degli errori

$$Cov(\varepsilon_i, \varepsilon_j) = 0 \qquad \text{per ogni } i \neq j$$

Stima dei parametri con il Metodo dei Minimi Quadrati

Il metodo più frequentemente utilizzato per la stima dei parametri del modello statistico ipotizzato

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

è quello dei minimi quadrati (m.q.).

Se il modello fosse una perfetta rappresentazione per la variabile dipendente ($\mathbf{y}$), dovremmo osservare la quantità ($\mathbf{X}\boldsymbol{\beta}$); invece, osserviamo $\mathbf{y}$. Di conseguenza, il modello ipotizzato avrà tanta più rilevanza nell'interpretare il fenomeno osservato, quanto meno determinante è il "peso" della v.c. accidentale $\boldsymbol{\varepsilon}$. Quindi, intuitivamente, si vorrebbe che gli scarti $\boldsymbol{\varepsilon}=\mathbf{y}-\mathbf{X}\boldsymbol{\beta}$ siano più piccoli possibili, e cioè lo scarto esistente tra i valori osservati $\mathbf{y}$ e quelli teorici $\mathbf{X}\boldsymbol{\beta}$ siano "piccoli" in qualche senso.
In tale contesto, trova giustificazione l'applicazione del metodo dei minimi quadrati per la stima del vettore $\boldsymbol{\beta}$, che consiste nel determinare il valore del vettore $\boldsymbol{\beta}$ tale che la somma degli scarti al quadrato minima.

A tal fine, costruiamo la seguente funzione (somma degli scarti al quadrato):

$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \boldsymbol{\varepsilon}^t \boldsymbol{\varepsilon} = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$$

La quale può essere riscritta nel seguente modo:

$$S(\boldsymbol{\beta}) = (\mathbf{y}^t - \boldsymbol{\beta}^t \mathbf{X}^t)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) =$$

$$= \mathbf{y}^t \mathbf{y} - \boldsymbol{\beta}^t \mathbf{X}^t y - \mathbf{y}^t \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^t \mathbf{X}^t \mathbf{X}\boldsymbol{\beta} =$$

$$= \mathbf{y}^t \mathbf{y} - 2\boldsymbol{\beta}^t \mathbf{X}^t y + \boldsymbol{\beta}^t \mathbf{X}^t \mathbf{X}\boldsymbol{\beta}$$

L'ultima relazione discenda dal fatto che il trasposto di uno scale è uguale allo scalare stesso, cioè

$$(y^t \mathbf{X}\boldsymbol{\beta})^t = \boldsymbol{\beta}^t \mathbf{X}^t y$$

Stimare $\boldsymbol{\beta}$ con il metodo dei m.q. significa individuare quel vettore tale che sostituito nella funzione S($\boldsymbol{\beta}$) la minimizza.
Imponiamo le condizioni del primo ordine

$$\frac{\partial S(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = -2\mathbf{X}^t\mathbf{y} + 2\mathbf{X}^t\mathbf{X}\boldsymbol{\beta} = 0$$

Da quest'ultima si ottengono le cosiddette <u>p-equazioni normali</u>

$$\mathbf{X}^t\mathbf{X}\boldsymbol{\beta} = \mathbf{X}^t\mathbf{y}$$

Se r($\mathbf{X}^t\mathbf{X}$)=p allora

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{y}$$

controllando le derivate seconde si verifica che si tratta effettivamente del vettore che minimizza la funzione S($\boldsymbol{\beta}$).

Caso particolare: modello lineare con un solo regressore

$$\underset{(nx\ 1)}{\mathbf{y}} = \underset{(nx\ 2)(2\ x\ 1)}{\mathbf{X}\boldsymbol{\beta}} + \underset{(nx\ 1)}{\boldsymbol{\varepsilon}}$$

$$\underset{(nx1)}{\mathbf{y}} = \begin{pmatrix} y_1 \\ y_2 \\ .. \\ y_n \end{pmatrix} \qquad \underset{(nx2)}{\mathbf{X}} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ .. & .. \\ 1 & x_n \end{pmatrix} \qquad \underset{(2x1)}{\boldsymbol{\beta}} = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \qquad \underset{(nx\ 1)}{\boldsymbol{\varepsilon}} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ .. \\ \varepsilon_n \end{pmatrix}$$

Determiniamo le stime ai minimi quadrati del vettore dei parametri $\boldsymbol{\beta}$. In questo contesto, si ha:

$$(\mathbf{X}^t\mathbf{X})^{-1} = \left[\begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix}\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ .. & .. \\ 1 & x_n \end{pmatrix}\right]^{-1} = \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}^{-1} = \frac{1}{n\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2}\begin{pmatrix} \sum_{i=1}^n x_i^2 & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & n \end{pmatrix}$$

$$\mathbf{X}^t\mathbf{y} = \begin{pmatrix} 1 & 1 & .. & 1 \\ x_1 & x_2 & .. & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ .. \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

Utilizzando le ultime due espressioni, otteniamo

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \frac{1}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & -\sum_{i=1}^{n} x_i \\ -\sum_{i=1}^{n} x_i & n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix} =$$

$$= \frac{1}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \begin{pmatrix} \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i x_i\right) \\ -\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) + n\left(\sum_{i=1}^{n} y_i x_i\right) \end{pmatrix}$$

Le stime dei due parametri risultano essere:

$$\hat{\beta}_0 = \frac{\left(\sum_{i=1}^{n} y_i\right)\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} x_i y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

$$\hat{\beta}_1 = \frac{n\sum_{i=1}^{n} y_i x_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

Da semplici calcoli algebrici si evince che:

$$\hat{\beta}_1 = \frac{\left[n\sum_{i=1}^{n} y_i x_i - \left(\sum_{i=1}^{n} x_i \right)\left(\sum_{i=1}^{n} y_i \right) \right] \frac{1}{n^2}}{\left[n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right] \frac{1}{n^2}} =$$

$$= \frac{\frac{1}{n}\sum_{i=1}^{n} y_i x_i - \left(\frac{1}{n}\sum_{i=1}^{n} x_i \right)\left(\frac{1}{n}\sum_{i=1}^{n} y_i \right)}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i \right)^2} = \frac{Cov(X,Y)}{V(X)}$$

In relazione all'intercetta, consideriamo il sistema delle equazioni normali

$$\begin{pmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} y_i x_i \end{pmatrix} \Rightarrow \begin{cases} n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases}$$

Dalla prima equazione, si deduce che:

$$n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \Rightarrow \beta_0 + \beta_1 \frac{1}{n} \sum_{i=1}^{n} x_i = \frac{1}{n} \sum_{i=1}^{n} y_i \Rightarrow \beta_0 + \beta_1 \bar{x} = \bar{y}$$

Proprietà degli stimatori m.q.

1. Se sono vere le ipotesi fondamentali sulla v.c, errore, lo stimatore ai minimi quadrati del vettore dei parametri β in un modello lineare è **<u>non distorto</u>**:

$$E\left[\hat{\beta}\right]=E\left\{\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{y}\right\}=E\left\{\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\left[\mathbf{X}\beta+\varepsilon\right]\right\}=$$

$$=\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{X}\beta+\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t E\left(\varepsilon\right)=\beta$$

2. Lo stimatore ai minimi quadrati del vettore dei parametri β è lineare in **y**:

$$\hat{\beta}=\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{Y}=\mathbf{M}\mathbf{y} \qquad \text{dove} \qquad \mathbf{M}=\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t$$

3. La matrice di varianze e covarianze dello stimatore ai m.q. di β è data da:

$$V\left(\hat{\beta}\right)=V\left[\mathbf{My}\right]=\mathbf{M}V\left(\mathbf{Y}\right)\mathbf{M}^t=\mathbf{M}\sigma^2\mathbf{I}\mathbf{M}^t=$$

$$=\sigma^2\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-1}=\sigma^2\left(\mathbf{X}^t\mathbf{X}\right)^{-1}$$

Si osservi che la matrice di varianze e covarianze dello stimatore ai m.q. del vettore dei parametri β, dipende dal parametro sconosciuto σ^2.

Stima di σ^2

Poiché $V(\varepsilon_\iota)=\sigma^2$ allora è ragionevole stimare σ^2 tramite i residui (cioè una stima dell'errore ε) i quali possono essere ottenuti nel seguente modo:

dal modello stimato: $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}$ determiniamo i valori stimati delle ordinate: $\hat{y}_i$

Per i=1,…,n. Una stima degli errori è: $\hat{\varepsilon}_i = y_i - \hat{y}_i$ i=1,…,n

In seguito dimostreremo che : $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i$

Di conseguenza, la media dei residui è zero, cioè: $\bar{\hat{\varepsilon}} = \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i = \sum_{i=1}^{n}\left(y_i - \hat{y}_i\right) = 0$

Lo stimatore naturale della varianza delle popolazione è: $S_\varepsilon^2 = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{\varepsilon}_i - \bar{\hat{\varepsilon}}\right]^2 = \frac{1}{n}\sum_{i=1}^{n}\hat{\varepsilon}_i^2$

La versione non-distorta è:

$$\hat{\sigma}^2 = \frac{1}{n-p}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 = \frac{1}{n-p}\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \frac{1}{n-p}\hat{\boldsymbol{\varepsilon}}^t\hat{\boldsymbol{\varepsilon}}$$

Teorema di Gauss-Markov

Sotto le ipotesi fondamentali del modello lineare, lo stimatore ai m.q. del vettore β sono i più efficienti nella classe degli stimatori lineari e non-distori per β.

Dimostrazione.

Sia **C** una matrice di costanti note. Consideriamo il seguente generico stimatore lineare in **y** del vettore β:

$$\tilde{\beta} = (\mathbf{C} + \mathbf{M})\mathbf{y}$$

dove $\mathbf{M} = \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t$

Lo stimatore considerato è non distorto per β se: $E\left[\tilde{\beta}\right] = \beta$

$$\begin{aligned}
\mathrm{E}\left[\tilde{\beta}\right] &= \mathrm{E}\left[(\mathbf{C}+\mathbf{M})\mathbf{y}\right] = \mathrm{E}\left\{(\mathbf{C}+\mathbf{M})(\mathbf{X}\beta+\boldsymbol{\varepsilon})\right\} = \\
&= \mathrm{E}\left\{(\mathbf{C}+\mathbf{M})\mathbf{X}\beta + (\mathbf{C}+\mathbf{M})\boldsymbol{\varepsilon}\right\} = (\mathbf{C}+\mathbf{M})\mathbf{X}\beta + (\mathbf{C}+\mathbf{M})\mathrm{E}\left(\boldsymbol{\varepsilon}\right) = \\
&= (\mathbf{C}+\mathbf{M})\mathbf{X}\beta = \left[\mathbf{C} + \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\right]\mathbf{X}\beta = \mathbf{C}\mathbf{X}\beta + \left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{X}\beta = \\
&= \mathbf{C}\mathbf{X}\beta + \beta = \beta \iff \mathbf{C}\mathbf{X} = \mathbf{0}
\end{aligned}$$

Calcoliamo, ora, la matrice di varianze e covarianze:

$$V\left(\tilde{\boldsymbol{\beta}}\right)=V\left[(\mathbf{C}+\mathbf{M})\mathbf{y}\right]=(\mathbf{C}+\mathbf{M})V(\mathbf{y})(\mathbf{C}+\mathbf{M})^t=(\mathbf{C}+\mathbf{M})\sigma^2\mathbf{I}(\mathbf{C}+\mathbf{M})^t=$$

$$=\sigma^2(\mathbf{C}+\mathbf{M})(\mathbf{C}+\mathbf{M})^t$$

Si evidenzia che:

$$(\mathbf{C}+\mathbf{M})(\mathbf{C}+\mathbf{M})^t=\left[C+\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\right]\left[C^t+\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\right]=$$

$$=CC^t+C\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-1}+\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^tC^t+\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-1}=$$

$$=CC^t+\left(\mathbf{X}^t\mathbf{X}\right)^{-1}$$

In definitiva, otteniamo:

$$V\left(\tilde{\beta}\right)=\sigma^2\left[\mathbf{C}\mathbf{C}^t+\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\right]=\sigma^2\mathbf{C}\mathbf{C}^t+\sigma^2\left(\mathbf{X}^t\mathbf{X}\right)^{-1}=$$

$$=\sigma^2\mathbf{C}\mathbf{C}^t+V\left(\hat{\beta}\right)$$

Quindi, $$V\left(\tilde{\beta}\right)-V\left(\hat{\beta}\right)=\sigma^2\mathbf{C}\mathbf{C}^t$$

è una matrice semidefinita positiva e, come tale, presenterà sulla diagonale principale termini certamente non negativi. Pertanto, la varianza dello stimatore ai m.q. di qualsiasi parametro del modello lineare sarà non superiore alla varianza di qualsiasi altro stimatore lineare e non distorto per β.

Somme di Quadrati – Indice di Determinazione

1. Dalle p-equazioni del sistema normale $\mathbf{X}^t\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}^t\mathbf{y}$

sostituendo ad $\mathbf{X}^t$ il vettore unitario $\mathbf{1}^t$, otteniamo la prima equazione del sistema:

$$(1,1,\ldots,1)\begin{pmatrix} 1 & x_{11} & x_{12} & \ldots & x_{1k} \\ 1 & x_{21} & x_{22} & \ldots & x_{2k} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{n1} & x_{n2} & \ldots & x_{nk} \end{pmatrix}\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \ldots \\ \hat{\beta}_k \end{pmatrix} = (1,1,\ldots,1)\begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ \\ y_n \end{pmatrix}$$

$$\left(n, \sum_{i=1}^{n} x_{1i}, \ldots\ldots, \sum_{i=1}^{n} x_{ki} \right) \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \ldots \\ \hat{\beta}_k \end{pmatrix} = \sum_{i=1}^{n} y_i$$

cioè

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{1i} + \ldots + \hat{\beta}_k \sum_{i=1}^{n} x_{ki} = \sum_{i=1}^{n} y_i$$

Dividendo tutto per n, si ottiene:

$$\hat{\beta}_0 + \hat{\beta}_1 \overline{x}_1 + \ldots + \hat{\beta}_k \overline{x}_k = \overline{y}$$

Da questa relazione si evince che il baricentro dei dati soddisfa l'equazione del modello stimato.

Inoltre, dalle equazioni del sistema normale, sostituendo ad $\mathbf{X}^t$ il Vettore unitario, abbiamo :

$$1^t \mathbf{X}\hat{\boldsymbol{\beta}} = 1^t \mathbf{y}$$

Dato che $\mathbf{X}\hat{\boldsymbol{\beta}} = \hat{\mathbf{y}}$

Si ha: $$1^t \hat{\mathbf{y}} = 1^t \mathbf{y} \Rightarrow \sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} y_i$$

cioè, la somma delle ordinate stimate è uguale alla somma delle ordinate osservate.

Dal modello stimato

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t\mathbf{y} = \mathbf{H}\mathbf{y}$$

dove $\mathbf{H} = \mathbf{X}\left(\mathbf{X}^t\mathbf{X}\right)^{-1}\mathbf{X}^t$ è simmetrica ed idempotente, si ottiene:

- Somma dei quadrati delle ordinate stimate

$$\sum_{i=1}^{n} \hat{y}_i^2 = \hat{\mathbf{y}}^t\hat{\mathbf{y}} = \left(\mathbf{H}\mathbf{y}\right)^t\left(\mathbf{H}\mathbf{y}\right) = \mathbf{y}^t\mathbf{H}^t\mathbf{H}\mathbf{y} = \mathbf{y}^t\mathbf{H}\mathbf{y}$$

- Somma dei prodotti tra ordinate stimate e ordinate osservate

$$\sum_{i=1}^{n} \hat{y}_i y_i = \hat{\mathbf{y}}^t\mathbf{y} = \left(\mathbf{H}\mathbf{y}\right)^t\mathbf{y} = \mathbf{y}^t\mathbf{H}^t\mathbf{y} = \mathbf{y}^t\mathbf{H}\mathbf{y}$$

quindi, la somma dei quadrati delle ordinate stimate è uguale alla somma dei prodotti tra ordinate stimate ed ordinate osservate, cioè:

$$\sum_{i=1}^{n} \hat{y}_i^2 = \mathbf{y}^t\mathbf{H}\mathbf{y} = \sum_{i=1}^{n} y_i$$

Definiamo, ora, le seguenti quantità:

- SOMMA dei QUADRATI TOTALI (**SQT**) o devianza totale:

$$SQT = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} \left(y_i^2 - 2 y_i \bar{y} + \bar{y}^2 \right) = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 = \mathbf{y}^t \mathbf{y} - n\bar{y}^2$$

- SOMMA dei QUADRATI del MODELLO (**SQM**) o devianza della regressione:

$$SQM = \sum_{i=1}^{n} \left(\hat{y}_i - \bar{\hat{y}} \right)^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \sum_{i=1}^{n} \hat{y}_i^2 - n\bar{y}^2 = \hat{\mathbf{y}}^t \hat{\mathbf{y}} - n\bar{y}^2 = \mathbf{y}^t \mathbf{H} \mathbf{y} - n\bar{y}^2$$

- SOMMA dei QUADRATI dei RESIDUI (**SQR**) o devianza residua:

$$SQR=\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2=\left(\mathbf{y}-\hat{\mathbf{y}}\right)^t\left(\mathbf{y}-\hat{\mathbf{y}}\right)=\mathbf{y}^t\mathbf{y}-\mathbf{y}^t\hat{\mathbf{y}}-\hat{\mathbf{y}}^t\mathbf{y}+\hat{\mathbf{y}}^t\hat{\mathbf{y}}=$$

$$=\mathbf{y}^t\mathbf{y}-\mathbf{y}^t\hat{\mathbf{y}}-\hat{\mathbf{y}}^t\mathbf{y}+\mathbf{y}^t\mathbf{H}\mathbf{y}=$$

essendo: $\hat{\mathbf{y}}^t\hat{\mathbf{y}}=\mathbf{y}^t\mathbf{H}\mathbf{y}=\hat{\mathbf{y}}^t\mathbf{y}$

$$=\mathbf{y}^t\mathbf{y}-\mathbf{y}^t\hat{\mathbf{y}}-\hat{\mathbf{y}}^t\mathbf{y}+\hat{\mathbf{y}}^t\mathbf{y}=\mathbf{y}^t\mathbf{y}-\mathbf{y}^t\hat{\mathbf{y}}$$

ma $\left(\mathbf{y}^t\hat{\mathbf{y}}\right)^t=\hat{\mathbf{y}}^t\mathbf{y}$

$$=\mathbf{y}^t\mathbf{y}-\mathbf{y}^t\mathbf{H}\mathbf{y}=\mathbf{y}^t\left(\mathbf{I}-\mathbf{H}\right)\mathbf{y}$$

Dalla scomposizione della devianza, sappiamo che: **SQT=SQM+SQR**, la quale può essere scritta in modo tale da far emergere che dipende solo dai dati osservati, cioè

$$\mathbf{y}^t\mathbf{y}=\mathbf{y}^t\mathbf{H}\mathbf{y}+\mathbf{y}^t\left(\mathbf{I}-\mathbf{H}\right)\mathbf{y}$$

Dividendo per n SQT, SQM e SQR otteniamo, rispettivamente, la varianza totale, la varianza spiegata dal modello stimato e la varianza residua, cioè

$$V(y) = \frac{SQT}{n} \qquad V(\hat{y}) = \frac{SQM}{n} \qquad V(\hat{\varepsilon}) = \frac{SQR}{n}$$

Di conseguenza, dalla scomposizione della devianza, si ha

$$V(y) = V(\hat{y}) + V(\hat{\varepsilon})$$

L'ultima relazione ci informa che parte della variabilità totale (osservata) è spiegata dal modello stimato mentre la varianza dei residui rappresenta la parte di variabilità delle ordinate osservate non spiegata dal modello stimato. Tanto più elevata è la varianza residua tanto meno il modello stimato si adatta ai dati. Un indice sintetico che misura la bontà di adattamento del modello stimato ai dati, denominato indice di determinazione, è il seguente:

$$R^2 = \frac{SQM}{SQT} = 1 - \frac{SQR}{SQT}$$

dato che è un rapporto di composizione, moltiplicato per 100, ci fornisce la percentuale di variabilità totale spiegata dal modello stimato.

Ipotesi di Normalità dei residui

Stimati i parametri del modello lineare con il metodo dei m.q., al fine di completare il processo di stima è necessario costruire intervalli di confidenza e/o test d'ipotesi sui parametri del modello.
L'obiettivo è quello di verificare se i regressori utilizzati hanno "senso" nello spiegare nello spiegare Y oppure le stime (valori numerici) ottenute sono solo il risultato di operazioni aritmetiche effettuate, senza alcune validità statistica. Evidentemente, è necessario conoscere la distribuzione dello stimatore ai m.q. del vettore $\boldsymbol{\beta}$.
A tal fine, ipotizzeremo che la v.c. $\boldsymbol{\varepsilon}$ segua una distribuzione Normale

Ipotesi: $$\boldsymbol{\varepsilon} \sim \mathbf{N}\left(\mathbf{0}, \sigma^2 \mathbf{I}\right)$$

Distribuzioni di Y e degli stimatori ai m.q. di β

Test sui parametri del modello lineare

Analisi dei Residui

Esempio Modello Lineare: Corsi di Pr-Inf e T.I.S. a.a. 2008-09
Dati tratti da S. Weisberg (1985): "Applied Linear Regression", pag. 35

Obiettivo dell'analisi: studiare il consumo di carburante in funzione di altre variabili;

Dati: rilevati in 48 Stati dell'USA, nel 1974
Matrice dei dati: oltre alla riga di intestazione delle colonne,
la matrice dei dati è formata da 48 righe e da 9 colonne.
Significato delle colonne:
prima: **STATE** (indica i 48 Stati considerati)
seconda: **POP** (popolazione in migliaia, relativa la 1971)
terza: **TAX** (tasse sul carburante, in centesimi per gallone; 1 Gallon = 3.7 Litri); X_1
quarta: **NLIC** (Numero di patenti nel 1971)
quinta: **INC** (reddito pro-capite, in migliaia di dollari, nel 1972); X_3
sesta: **ROAD** (miglia di autostrade, in migliaia nel 1971; 1 Miglio = 1.6 Km); X_4
settima: **FUELC** (consumo di carburante, in milioni di gallons);
ottava: **DLIC**=100*NLIC/POP (percentuale di popolazione con patente); X_2
nona: **FUEL**=1000*FUELC/POP (consumo di carburante per persona); Y

Modello da Stimare

$$Y = \beta_0 + \beta_1 * TAX + \beta_2 * DLIC + \beta_3 * INC + \beta_4 * ROAD$$

```
# Lettura dei dati
datW<-read.table(file="C:/Documents and
        Settings/F.Domma/Desktop/CorsoDURATA/
         DATI-MOD_LIN-AA200809.txt", header=TRUE,
        dec=".")

y<-datW[1:48,9];  x1<-datW[1:48,3]; x2<-datW[1:48,8]; x3<-
datW[1:48,5]; x4<-datW[1:48,6]
```

summary(lm(y ~ x1 + x2 + x3 + x4))

Residuals:

Min	1Q	Median	3Q	Max
-122.03	-45.57	-10.66	31.53	234.95

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	377.291	185.541	2.033	0.048207 *
x1	-34.790	12.970	-2.682	0.010332 *
x2	13.364	1.923	6.950	1.52e-08 ***
x3	-66.589	17.222	-3.867	0.000368 ***
x4	-2.426	3.389	**-0.716**	**0.477999**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Residual standard error: 66.31 on 43 degrees of freedom
Multiple R-Squared: 0.6787, Adjusted R-squared: 0.6488
F-statistic: 22.71 on 4 and 43 DF, p-value: 3.907e-10

summary(lm(y ~ x1 + x2 + x3))

Residuals:

Min	1Q	Median	3Q	Max
-110.10	-51.22	-12.89	24.49	238.77

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	307.328	156.831	1.960	0.056394 .
x1	-29.484	10.584	-2.786	0.007848 **
x2	13.748	1.837	7.485	2.24e-09 ***
x3	-68.023	17.010	-3.999	0.000240 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05

Residual standard error: 65.94 on 44 degrees of freedom
Multiple R-Squared: 0.6749, Adjusted R-squared: 0.6527
F-statistic: 30.44 on 3 and 44 DF, p-value: 8.235e-11

Modello Stimato

$$\hat{Y} = 307.328 - 29.484 * TAX + 13.748 * DLIC - 68.023 * INC$$

Residui stimati:

res2<-residuals(mod2)

Res2: residui stimati relativi sulla base del secondo modello stimato