

Le dimostrazioni da portare a esame sono le seguenti:

DIMOSTRAZIONI

A.1 MODELLO LINEARE

Slide 73 correttezza

Slide 74 efficienza Gauss Markov

B.1 VIOLAZIONI ERRORI

Slide 10/52/171 correttezza (identiche)

Slide 11/53/172(Aitken) efficienza (identiche)

Slide 69 valore Durbin Watson

Slide 77 Stime WLS

Slides 120-121 Stime GLS incorrelazione omoschedasticità

Slides 168-170 Stime GLS incorrelazione eteroschedasticità

B.2 COLL-LIN

Slides 66-69: solo slide 69

B.3 VIOLAZIONE NORMALITA

Slide 5 correttezza (identica a precedenti correttezza)

C.1 MODELLO LINEARE MULTIVARIATO CLASSICO

Slide 8 residui, var.esplicative, valori previsti

Slides 19-21: solo slide 21

Slide 26 efficienza Gauss Markov

Slide 47-48: omettere

C.1 MODELLO LINEARE MULTIVARIATO

Slide 12-13 Da modello lineare multivariato generalizzato a classico

Per il resto la trattazione essendo un esame di statistica deve riportare formule e grafici contenuti nelle slides

1) Correttezza stimatori OLS slide 66

Gli stimatori OLS B^{\wedge} di b sono corretti, infatti...

$$E[B^{\wedge}] = E[(X'X)^{-1}X'y] = (X'X)^{-1}X'(Xb + e) = b$$

2) Teorema Gauss Markov slide 68

$$\begin{aligned} E[(B - b)(B - b)'] &= E\{[(X'X)^{-1}X'y - b][(X'X)^{-1}X'y - b]'\} \\ &= E\{[(X'X)^{-1}X'Xb - b + (X'X)^{-1}X'E][(X'X)^{-1}(X'X)b - b + (X'X)^{-1}X'E]'\} \\ &= [(X'X)^{-1}X'] E\{(EE') [X(X'X)^{-1}]\} = \sigma^2 [(X'X)^{-1}X'X(X'X)^{-1}] = \sigma^2 (X'X)^{-1} \end{aligned}$$

3) Correttezza stimatori eteroschedastici slide 8

Valgono correttezza e la linearità perché non dipendono da omoschedasticità. Dato lo stimatore eteroschedastico B^* ...

$$E[B^*] = E[(X'X)^{-1}X'y] = (X'X)^{-1}X'E(y) = (X'X)^{-1}X'E[Xb + e^*] = (X'X)^{-1}X'XE(b) = b$$

4) Efficienza stimatori eteroschedastici slide 9

Lo stimatore B^* con errori eteroschedastici non è più BLUE e quindi efficiente.

$$\begin{aligned} Var(B^*) &= E\{[(X'X)^{-1}X'y - b][(X'X)^{-1}X'y - b]'\} \\ &= E\{[(X'X)^{-1}X'](Xb + e^*) - b\} \{[(X'X)^{-1}X'](Xb + e^*) - b\}' \\ &= E\{[(X'X)^{-1}X']e^*\} \{[(X'X)^{-1}X']e^*\}' = (X'X)^{-1}X'E\{(e^*e^{*\prime})[(X'X)^{-1}X']'\} \\ &= (X'X)^{-1}X'\Sigma_{e^*}[(X'X)^{-1}X']' \end{aligned}$$

5) Correttezza stimatori correlati slide 47

Proprietà stimatori OLS: valgono correttezza e linearità che non dipendono da autocorrelazione. Dato lo stimatore con osservazioni autocorrelate $B^{\#}$...

$$E[B] = E[(X'X)^{-1}X'y] = (X'X)^{-1}XE(y) = (X'X)^{-1}X'E(Xb + e^{\#}) = (X'X)^{-1}X'XE(b) = b$$

6) Efficienza stimatori correlati slide 48

Stimatore $B^{\#}$ con errori correlati non è più BLUE e quindi efficiente.

$$\begin{aligned} Var(B^{\#}) &= E\{[(X'X)^{-1}X'y - b][(X'X)^{-1}X'y - b]'\} \\ &= E\{[(X'X)^{-1}X'](Xb + e^{\#}) - b\} \{[(X'X)^{-1}X'](Xb + e^{\#}) - b\}' \\ &= E\{[(X'X)^{-1}X']e^{\#}\} \{[(X'X)^{-1}X']e^{\#}\}' = (X'X)^{-1}X'E\{(e^{\#}e^{\#\prime})[(X'X)^{-1}X']'\} \\ &= (X'X)^{-1}X'\Sigma_{e^{\#}}[(X'X)^{-1}X']' \end{aligned}$$

7) Incorrelazione stimatori residui incorrelati slide 108

In un'equazione OLS di tipo classico con errori incorrelati, dato che i w_i hanno anche essi valore atteso pari a 0

$$E(w_i) = E(\varepsilon_i^\# - \rho\varepsilon_{i-1}^\#) = E(\varepsilon^\#) - \rho E(\varepsilon_{i-1}^\#) = 0$$

si ha...

$$\begin{aligned} cov(w_i, w_{i-1}) &= cov(\varepsilon_i^\# - \rho\varepsilon_{i-1}^\#, \varepsilon_{i-1}^\# - \rho\varepsilon_{i-2}^\#) = cov(\varepsilon_i^\#\varepsilon_{i-1}^\# - \varepsilon_i^\#\rho\varepsilon_{i-2}^\# - \rho\varepsilon_{i-1}^\#\varepsilon_{i-1}^\# + \rho\varepsilon_{i-1}^\#\rho\varepsilon_{i-2}^\#) \\ &= cov(\varepsilon_i^\#\varepsilon_{i-1}^\#) - \rho * cov(\varepsilon_i^\#\varepsilon_{i-2}^\#) - \rho * cov(\varepsilon_{i-1}^\#\varepsilon_{i-1}^\#) + \rho^2 * cov(\varepsilon_{i-1}^\#\varepsilon_{i-2}^\#) \\ &= \rho - \rho^3 - \rho + \rho^3 = 0 \end{aligned}$$

8) Costruzione stimatori sferici GLS slides 155-157

Il modello intende eliminare correlazione e eteroschedasticità. In questo caso si utilizzano gli stimatori dei minimi quadrati generalizzati (GLS) interpretabili in modo analogo al modello classico quali stimatori OLS basati su variabili trasformate.

Si ha:

$$y = XB^{\wedge 0} + E^{\wedge 0} \quad (1)$$

La matrice di varianze covarianze campionaria degli errori è:

$$\frac{1}{n}(E^{\wedge 0}E^{\wedge 0'}) = S_{E^{\wedge 0}} \quad (2)$$

Si ipotizza a questo punto l'esistenza di una matrice $V(n,n)$ non singolare tale che...

$$S_{E^{\wedge 0}} = \sigma^2 VV'; \quad (S_{E^{\wedge 0}})^{-1} = (\sigma^2 VV')^{-1} = \frac{1}{\sigma^2} (V')^{-1}V^{-1} \quad (3)$$

Si definiscono gli errori trasformati ε^\wedge tali che:

$$E^\wedge = V^{-1}E^{\wedge 0} \quad (4)$$

Per la (2) si ha...

$$\begin{aligned} V^{-1}S_{E^{\wedge 0}}(V')^{-1} &= V^{-1}\sigma^2 VV'(V')^{-1} = \sigma^2 I_n \\ V'(S_{E^{\wedge 0}})^{-1}V &= V' \left(\frac{1}{\sigma^2} \right) (V')^{-1}V^{-1}V = \left(\frac{1}{\sigma^2} I_n \right) \end{aligned} \quad (5)$$

Moltiplicando entrambi i membri della (1) per V^{-1} si ha il modello delle variabili trasformate.

$$V^{-1}y = y^\circ = V^{-1}(XB^{\wedge 0} + E^{\wedge 0}) = X^\circ B + E^\wedge \quad (6)$$

$$\text{dove: } B^{\wedge 0} = (X^{\circ'}X^\circ)^{-1}X^{\circ'}y^\circ = [X'(S_{E^{\wedge 0}})^{-1}X]^{-1}X'(S_{E^{\wedge 0}})^{-1}y \quad (7)$$

Stima dei minimi quadrati generalizzati GLS con matrice di varianze covarianze degli errori correlati ed eteroschedastici (3). Da tutte le formule precedenti si ha:

$$\Sigma_y = B^{\circ'}\Sigma_{x^\circ}B^\circ + S_{E^\wedge} = B^{\circ'}\Sigma_{x^\circ}B^\circ + \sigma^2 I_n \quad (8)$$

9) Correttezza stimatori GLS slide 158

Lo stimatore GLS B° gode della proprietà di correttezza, infatti:

$$B^\circ = (X^\circ)'X^\circ)^{-1}X^\circ'y^\circ = (X^\circ)'X^\circ)^{-1}X^\circ'(X^\circ b^\circ + E^\circ) = b^\circ + (X^\circ)'X^\circ)^{-1}X^\circ'E^\circ \quad (1)$$

$$E(B^\circ) = E[b^\circ + (X^\circ)'X^\circ)^{-1}X^\circ'E^\circ] = b^\circ + (X^\circ)'X^\circ)^{-1}X^\circ'E(E^\circ) = b^\circ \quad (2)$$

Poiché il valore atteso è pari a 0 si può dimostrare anche che è consistente.

$$E(E^\circ) = E(V^{-1})E = V^{-1}E(E) = 0$$

10) Efficienza stimatori GLS slide 159

Lo stimatore GLS è efficiente per le variabili trasformate y°, X° .

Si dimostra il teorema di Aitken analogo a quello di Gauss Markov per gli stimatori OLS. La matrice varianze-covarianze dello stimatore OLS è data da:

$$\begin{aligned} E[(B^\circ - b^\circ)'(B^\circ - b^\circ)'] &= E\{[(X^\circ)'X^\circ)^{-1}X^\circ'y^\circ - b^\circ][(X^\circ)'X^\circ)^{-1}X^\circ'y^\circ - b^\circ]\}' \\ &= E\{[(X^\circ)'X^\circ)^{-1}X^\circ'X^\circ b^\circ - b^\circ + (X^\circ)'X^\circ)^{-1}X^\circ'E][(X^\circ)'X^\circ)^{-1}X^\circ'X^\circ b^\circ - b^\circ \\ &\quad + (X^\circ)'X^\circ)^{-1}X^\circ'E]\}' = [(X^\circ)'X^\circ)^{-1}X^\circ'E(EE')X^\circ(X^\circ)'X^\circ)^{-1}]' = \\ &\sigma^2[(X^\circ)'X^\circ)^{-1}X^\circ'\Sigma_{E^\circ}^{-1}\Sigma_{E^\circ}\Sigma_{E^\circ}^{-1}X^\circ(X^\circ)'X^\circ)^{-1}]' = \sigma^2[(X^\circ)\Sigma_{E^\circ}^{-1}]^{-1}X^\circ'\Sigma_{E^\circ}^{-1}X^\circ(X^\circ)\Sigma_{E^\circ}^{-1}]^{-1}]' = \sigma^2(X^\circ)'X^\circ)^{-1} \\ &= \sigma^2(X^\circ)\Sigma_{E^\circ}^{-1}X^\circ)^{-1} \end{aligned}$$

Tale formula è simile a quella per gli stimatori OLS con variabili y, X e che garantisce l'efficienza. Inoltre...

$$\sigma^2(X^\circ)'X^\circ)^{-1} - \sigma^2(X'X)^{-1}$$

...è matrice semidefinita positiva e quindi non è più VUE né BLUE. Tuttavia, il teorema di Aitken stabilisce che nella classe degli stimatori lineari per il modello di regressione generalizzato lo stimatore GLS è efficiente in quanto caratterizzato dalla minima varianza.

11) Correttezza stimatori OLS non normali slide 6

Per gli stimatori OLS, non essendo più equivalenti agli stimatori di massima verosimiglianza, non vale più il teorema di Cramer-Rao, quindi gli stimatori OLS non sono più gli stimatori corretti a minima varianza fra tutti gli stimatori corretti (VUE).

Tuttavia, gli stimatori OLS rimangono corretti infatti...

$$E[B^\wedge] = E[(X'X)^{-1}X'y] = (X'X)^{-1}XE(y) = (X'X)^{-1}X'E(X\beta + \varepsilon) = (X'X)^{-1}X'XE(\beta) = b$$

Pertanto, poiché vale anche il teorema di Gauss Markov rimangono efficienti e BLUE.

12) Incorrelazione errori variabili esplicative osservate e previste slide 8

Nel modello multivariato di tipo descrittivo per ogni equazione le stime OLS di β_j sono trovate in modo analogo al caso univariato.

Infatti, data l'incorrelazione degli errori e_j rispetto a tali variabili esplicative Z la dipendenza di una variabile dipendente y_j da Z non influisce sulla dipendenza empirica di una altra variabile dipendente y_m .

In ogni equazione si ha:

$$\beta_j^\wedge = y_j Z'(ZZ')^{-1} \quad (1)$$

Nel complesso invece si ha:

$$B^\wedge = YZ'(ZZ')^{-1} \quad (2)$$

Quest'ultima equazione ci permette di ricavare la matrice dei valori predetti:

$$Y^{\wedge} = B^{\wedge}Z = YZ'(ZZ')^{-1}Z \quad (3)$$

La risultante della matrice di errori è:

$$E^{\wedge} = Y - B^{\wedge}Z = Y[I - Z'(ZZ')^{-1}Z] \quad (4)$$

Permangono le condizioni di ortogonalità del caso univariato poiché i residui sono incorrelati con le variabili esplicative (5) e con i valori predetti (6).

$$E^{\wedge}Z' = Y[I - Z'(ZZ')^{-1}ZZ'] = YZ' - YZ' = 0 \quad (5)$$

$$Y^{\wedge}E^{\wedge'} = YZ(Z'Z)^{-1}ZZ'[I - Z'(ZZ')^{-1}Z]Y' = Y'Z[I - Z'(ZZ')^{-1}Z]Y' = Y'ZY - Y'ZY = 0 \quad (6)$$

13) Efficienza stimatori OLS slide 28

Gli stimatori OLS multivariati, come nel caso multiplo, sono corretti consistenti ed efficienti.

$$E[B^{\wedge}] = E[YZ'(ZZ')^{-1}] = E[(BZ + E)Z'(ZZ')^{-1}] = B$$

$$P(|B^{\wedge} - B| < \lambda) = 1 \text{ per } n \rightarrow \infty$$

Il teorema di Gauss-Markov vale quindi anche nel contesto multivariato, lo stimatore B^{\wedge} dei minimi quadrati è il miglior stimatore lineare corretto BLUE. Si dimostra anche che lo stimatore di ML è UNVUE (e quindi anche l'OLS con normalità).

$$\begin{aligned} E[(B^{\wedge} - B)'(B^{\wedge} - B)] &= E\{[YZ'(ZZ')^{-1} - B]'[YZ'(ZZ')^{-1} - B]\} \\ &= E\{[BZZ'(ZZ')^{-1} - B + EZ'(ZZ')^{-1}]'[BZZ'(ZZ')^{-1} - B + EZ'(ZZ')^{-1}]\} \\ &= (ZZ')^{-1}Z E(E^{\wedge}E^{\wedge}')Z'(ZZ')^{-1}Z = \sigma^2[(ZZ')^{-1}ZZ'(ZZ')^{-1}] = \sigma^2[(ZZ')^{-1}] \end{aligned}$$

14) Costruzione soluzioni con errori sferici slides 10-13

Come nel caso del modello lineare classico multiplo non è possibile ottenere soluzioni in termini di stime OLS dato che la matrice degli errori non è diagonale. La soluzione viene ricavata anche in questo caso in termini di minimi quadrati generalizzati GLS

Ricordando quanto detto a proposito del modello lineare multiplo, con matrice di varianze covarianze della popolazione conosciuta Σ_e si ipotizzi l'esistenza di una matrice $W(nm, nm)$ non singolare tale che :

$$\Sigma_{E^{\circ}} = E[E^{\circ*'} E^{\circ*}] = \sigma^2 W'W; \quad (\Sigma_{E^{\circ}})^{-1} = \frac{1}{\sigma^2} W^{-1}(W')^{-1} \quad (20)$$

Si definiscano gli errori trasformati E tali che:

$$E^{\circ} = E^{\circ*}(W)^{-1} \quad (21)$$

In modo tale che:

$$\Sigma_{E^{\circ}} = E(E^{\circ'} E^{\circ}) = E[(W')^{-1} E^{\circ*'} E^{\circ*} W^{-1}] \quad (22)$$

Poiché W è matrice non stocastica, per la (21) si ha che...

$$\Sigma_{E^{\circ}} = (W')^{-1} E(E^{\circ*'} E^{\circ*}) W^{-1} = (W')^{-1} \sigma^2 W' W W^{-1} = \sigma^2 I_{nm} \quad (23)$$

Gli errori E° trasformati sono quindi sferici, omoschedastici e incorrelati rispetto a ogni variabile dipende e fra le variabili dipendenti.

Il modello lineare multivariato può essere riscritto come $Y^\circ = B^\circ Z^\circ + E^\circ$.

Moltiplicando entrambe le parti di questa equazione per W^{-1} si ha il modello nelle variabili trasformate.

$$W^{-1}Y^\circ = Y^{*\circ} = (B^{*\circ}Z^\circ + E^\circ)W^{-1} = B^{*\circ}Z^{*\circ} + E \quad (24)$$

con $Y^{*\circ} = Y^\circ W^{-1}$ e $Z^{*\circ} = Z^\circ W^{-1}$

Dalla (24) si ha pertanto...

$$B^{*\circ} = Y^{*\circ} Z^{*\circ\prime} (Z^{*\circ} Z^{*\circ\prime})^{-1} = Y^\circ \Sigma_{E^*}^{-1} Z^{\circ\prime} [Z^\circ \Sigma_{E^*}^{-1} Z^{\circ\prime}]^{-1} \quad (25)$$

E per ogni equazione

$$\beta_j^{*\circ} = y_j^\circ \Sigma_{E^*}^{-1} Z^{\circ\prime} [Z^\circ \Sigma_{E^*}^{-1} Z^{\circ\prime}]^{-1} \quad (26)$$

Si vede la sostanziale differenza dal modello multivariato classico. Anche per calcolare la soluzione per la singola equazione j in (25) occorre tenere presente tutte le equazioni attraverso la matrice di varianze-covarianze degli errori Σ_{E^*} .

Come si ricava una matrice W ? Per le proprietà degli autovettori e degli autovalori attraverso decomposizione spettrale della matrice Σ_{E^*} , W può essere definita come...

$$W = \sigma KL^{\frac{1}{2}}K'; \quad W^{-1} = \frac{1}{\sigma}(K')^{-1}L^{-\frac{1}{2}}K^{-1} \quad (27)$$

...dove K è la matrice degli autovettori e L è la matrice diagonale degli autovalori di Σ_{E^*} . Infatti, data l'ortogonalità degli autovettori per cui $K'K = I$, si ha...

$$\sigma KL^{\frac{1}{2}}K' = WW' = \Sigma_{E^*}$$

15) Soluzioni analisi covarianza slides 9-14

Se le caratteristiche x delle osservazioni appartenenti ai diversi gruppi sono differenti tra gruppo e gruppo l'analisi della varianza viene distorta: si attribuiscono alla varianza fra gruppi ente effetti che dipendono da tali caratteristiche

Esempio: non si può attribuire all'efficacia di un farmaco la diminuzione del colesterolo se il gruppo a cui è somministrato il farmaco è caratterizzato da pazienti con diversa età e peso da quelli a cui non è somministrato.

È necessario quindi un modello che tenga conto delle caratteristiche della popolazione X in modo da eliminare prima l'effetto di tali caratteristiche sulla variabile dipendente e quindi effettuare l'analisi della varianza.

Il modello di analisi della covarianza è dato da:

$$y_{ij} = (\beta_0 + \Sigma_k \beta_{kj} x_{ijk}) + \lambda_{00} + v_j + e_{ij} \quad (11)$$

Con $y_{ij} = \beta_0 + \Sigma_k \beta_{kj} x_{ijk} + y_{eij}$ che è il modello lineare di y sulle X (x_1, x_2, \dots, x_p) supposte centrate per semplicità.

$$y_{ij} - (\beta_0 + \Sigma_k \beta_{kj} x_{ijk}) = \lambda_{00} + v_j + e_{ij} \quad (13)$$

L'equazione sopra è il modello di analisi della varianza al netto del modello lineare.

In termini matriciali il modello di analisi della covarianza è:

$$y - \gamma_{00} = Xb + Av + e \quad (14)$$

che da origine a un modello lineare multiplo...

$$y - \gamma_{00} = Xb + y_e \quad (15)$$

Dalla formula soprastante abbiamo il residuo y_x e analisi della varianza pari a:

$$y - \gamma_{00} - Xb = y_e = Av + e \quad (16)$$

A questo punto si possono ricavare le soluzioni OLS del modello lineare multiplo nel modo consueto con:

$$b^{\wedge} = (X'X)^{-1}X'(y - \gamma_{00}) \quad (17)$$

$$b^{\wedge}'X'Xb^{\wedge} = SSX \quad (\text{devianza spiegata di regressione}) \quad (18)$$

Dalle formule (15) e (17) otteniamo...

$$y - \gamma_{00} - Xb^{\wedge} = y_e^{\wedge} = Av + e \quad (19)$$

Per risolvere la (19) ottenendo v e e occorre tener presente che A , a differenza di X , non è di rango pieno per come è costruita: il rango è inferiore all'ordine. Esistono molti metodi per risolvere il problema e un vasto dibattito in letteratura. Uno di questi metodi è porre dei vincoli sui parametri. Nel caso più semplice in cui il rango sia $p-1$ si pone il vincolo

$$\sum_j u_j = 0 \quad (20)$$

Nel caso in cui il rango sia ancora inferiore si pongono altri vincoli nel modo più opportuno.

A questo punto basta eliminare la riga l -esima di A ottenendo la matrice A° di dimensione $(n, p-1)$ il vettore u° in cui è stato eliminato di converso il rispettivo parametro u_k che sarà ottenuto dagli altri $p-1$ grazie al vincolo (20). Il modello (19) diviene:

$$y_e^{\wedge} = A^{\circ}v^{\circ} + e \quad (21)$$

che ha soluzione OLS pari a:

$$v^{\circ\wedge} = (A^{\circ}'A^{\circ})^{-1}A^{\circ}'y_e^{\wedge} \quad (22)$$

A questo punto il modello di analisi della covarianza (formula 14) diventa:

$$SST = (y - \gamma_{00})'(y - \gamma_{00})' = b^{\wedge}'X'Xb^{\wedge} + v^{\circ\wedge}'A^{\circ}'A^{\circ}v^{\circ\wedge} + e'e \quad (23)$$

con devianza spiegata dell'analisi della covarianza pari a $SSV = v^{\circ\wedge}'A^{\circ}'A^{\circ}v^{\circ\wedge}$
e devianza residui dell'analisi della covarianza pari a $SSE = e'e$

La devianza residua di regressione o devianza y_e^{\wedge} al netto della varianza è...

$$SST - SSX = SSY_E = SSV + SSE \quad (26)$$

La devianza di totale di SST viene suddivisa nella devianza spiegata di regressione SSX e residua di regressione SSY_E . A sua volta SSY_E è suddivisa in devianza fra i gruppi SSV e nei gruppi SSE .

$$SST = SSX + SSY_E = SSX + SSV + SSE \quad (27)$$

