ECONOMETRIA

7/10/15

• Modello di REGRESSIONE LINEARE MULTIPLA (CAP 4, 6 L2)

$Y_i = \beta_0 X_{0i} + \beta_1 X_{1i} + \dots + \beta_K X_{Ki} + u_i \qquad i = 1, \dots, n$

errore — dimensione campione

con $X_{0i} = 1 \ \forall i$ (costante)

Idea è che ci sia una relazione lin (approssimata) nella pop e quindi nel campione.

HP MODELLO:

1) $E[u_i | X_{1i}, \dots, X_{Ki}] = 0$ q.c. (con prob 1)

OSS Sia Y, sia X, sia U sono V.A. (anche X non sono considerate fisse come in stat medica)

$\Rightarrow$ 1) va letta come F. REGRESSIONE dell'ERRORE u_i RISPETTO ALLE V.A. $X_{1i}, \dots, X_{Ki}$. La F REG è $g(X_{1i}, \dots, X_{Ki})$ è V.C. (che ci dice per ogni valore $x_{1i}, \dots, x_{Ki}$ il valore atteso dell'errore)

$\Rightarrow$ le determinazioni sono tutti i valori attesi condizionati $E[u_i | X_{1i} = x_{1i}, \dots, X_{Ki} = x_{Ki}]$ a cui (nel caso continuo) è associata la DENSITÀ DI PROBABILITÀ CONGIUNTA DELLE X:

$f(x_{1i}, \dots, x_{Ki}) \qquad \forall\, x_{1i}, \dots, x_{Ki}$

$\Rightarrow$ SI DIM che

• $E[Y_i | X_{1i}, \dots X_{Ki}] = \beta_0 + \sum_{j=1}^{K} \beta_j X_{ji}$

VEDO PROPRIETÀ F. REG. su FILE WORD $\boxed{\text{(NOW)}}$

2) $(Y_i, X_{1i}, \dots, X_{Ki}) \sim$ I.I.D. $\forall i$ (rispetto a indice i)

(quindi le loro determinazioni sono un CAMPIONE CASUALE SEMPLICE)

cioè $(Y_i, X_{1i}, \dots, X_{Ki}) \overset{\text{ind}}{\perp} (Y_j, X_{1j}, \dots, X_{Kj}) \quad i \neq j$

(HP + FORTE DI I.I.D su una sola V.A. vista a stat)

Questa HP è accettabile (a volte) IN DATI CROSS-SECTIONAL, ma ⚠ NON HA SENSO SU S.S.

ES S.S modello semplice:

$Y_t = \beta_0 + \beta_1 X_t + u_t$ se (A.A.) $(Y_t, X_t) \sim$ IID $\forall t$

$\Rightarrow t \neq s \Rightarrow (Y_t, X_t) \perp (Y_{t+1}, X_{t+1})$

es Y_t = consumo al trimestre 3 e X_t = Reddito disponibile al trimestre t

$\Rightarrow$ nella pratica spesso le Y_t sono correlate nel tempo e pure le X_t

$\Rightarrow$ IN S.S. (2) NON HA SENSO!

- IN DATI Cross-Sectional (2) I.I.D. a volte va bene.

$\Rightarrow$ anche le distr. congiunte sono uguali

$F_i(Y, X_1, \dots, X_k) = F_j(Y, X_1, \dots, X_k)$

3) (3) tecnica per avere buone proprietà OLS (es. asintoticamente N)

$0 < E[Y_i^4] < +\infty \Rightarrow$ VARIANZA $Var(Y_i)$ FINITA

CURTOSI $= k = \dfrac{E[(Y_i - E(Y_i))^4]}{[Var(Y_i)]^2}$ FINITA

(OSS code spesse $\Rightarrow k \nearrow$) $k = 3$ se $f(Y) \sim N(\mu, \sigma^2)$

$k > 3 \Rightarrow$ LEPTOCURTICA

$k < 3 \Rightarrow$ PLATICURTICA

e pure (3)

$0 < E[X_{k_i}^4] < +\infty$ $\forall k = 1, \dots, j$ (analogamente)

OSS Distr. X_i e Y possono avere $k \neq 3 \Rightarrow$ possono non essere normali $\Rightarrow$ le code possono essere pesanti, ma non così grasse da rendere il momento $4° = +\infty$. Sulle code troviamo i valori estremi (che non possono avere prob. troppo alte) outliers = valori anomali = lontani dalla massa.

$\Rightarrow$ (3) $\Rightarrow$ OUTLIER SONO POCO PROBABILI.

(Se salta 3 $\Rightarrow$ OLS non va molto bene perché sensibili a outliers.)

$\Rightarrow$ Varianza Finita

Scanned by CamScanner

4) NON MULTICOLLINEARITÀ PERFETTA:

$(X_{0i}, X_{1i}, \dots, X_{ki})$ non sono perfettamente collineari tra loro. $\Rightarrow$ Un regressore non può essere scritto come C.L. dei restanti.

Può esserci o meno un'ulteriore (HP):

(5) $VAR[u_i | X_{1i}, \dots, X_{ki}] = \sigma_u^2$ q.c. $\forall X_{1i}, \dots, X_{ki}$

OMOSCHEDASTICITÀ CONDIZIONALE (uguale $\forall$ individuo)

$\Rightarrow E[u_i^2 | X_{1i}, \dots, X_{ki}] \overset{(1)}{=} VAR[u_i | X_{1i}, \dots, X_{ki}] \overset{(5)}{=} \sigma^2$ q.c.

f. ecc. in generale è V.C. $h(X_{1i}, \dots X_{ki})$ se (HP) 5 questa non dipende dalle X_k

(perché OSS $E[u_i | X_{1i}, \dots X_{ki}] = 0 \Rightarrow E[u_i] = 0$ per la legge del valore atteso iterato $E[E[u_i | X_{1i}, \dots, X_{ki}]]$

$\Rightarrow VAR(u_i) \overset{(1)}{=} E[u_i^2] = E[E[u_i^2 | X_{1i}, \dots, X_{ki}]] \overset{(5)}{=} \sigma_u^2$)

OSS OMOSCHEDASTICITÀ CONDIZIONALE $\Rightarrow$ OMOSCHEDASTICITÀ NON CONDIZIONALE

OSS Nei dati economici la (5) salta spesso $\Rightarrow$ Non l'assumiamo di default, se (5) $\Rightarrow$ abbiamo + proprietà per OLS.

- DATE LE (HP) e MODELLO $\Rightarrow$ PROBLEMA TROVARE OLS (in forma matriciale lo ricavo con gradiente etc...)

Scanned by CamScanner

STIMATORE OLS (lo ricaviamo in generale come metodo di approssimazione di Y tramite X, se poi vale modello e (H) lo stimatore ha buone proprietà.

° METODO OLS:

$$S(\beta_0, \beta_1, \dots, \beta_k) \doteq \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_{1i} - \dots - b_k X_{ki})^2 = $$

$$= \sum_{i=1}^{n} a_i = \text{funzione paraboloide}$$

(ed è regolare) da minimizzare.

$\Rightarrow \min_{\{b_0, \dots, b_k\}} S(\beta_0, \dots, \beta_k)$ la si risolve con F.O.C.

$$\begin{cases} \frac{\partial S(\cdot)}{\partial b_0} = 0 \\ \vdots \\ \frac{\partial S(\cdot)}{\partial b_k} = 0 \end{cases} \Leftrightarrow \begin{cases} (-2)\sum_{i=1}^{n} (Y_i - b_0 - b_1 X_{1i} - \dots - b_k X_{ki}) = 0 \\ \vdots \\ (-2)\sum_{i=1}^{n} (Y_i - b_0 - \dots - b_k X_{ki}) \cdot X_{ki} = 0 \end{cases}$$

$\Rightarrow$ conti

$\boxed{\text{se } k=1}$ $b_0 = \bar{y} - b_1 \bar{x}$ (con $X_{1i} = x_i$ è l'unica)

$$b_1 = \frac{S_{xy}}{S_x^2} \doteq \frac{\frac{1}{n}\sum_{i=1}^{n} (Y_i - \bar{Y})(X_i - \bar{X})}{\frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2}$$ — cov campionaria / varianza campionaria

$\boxed{\text{se } k \geq 1}$ $\Rightarrow$ FORMULA MATRICIALE

RIPASSO MATRICI (MATRIX APPENDIX PDF)

- MATRICE IDENTITÀ $= I_{n \times n} = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \end{bmatrix}$ è matrice diagonale con diagonale $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$
- In generale $AB \neq BA$ l'ordine dei fattori conta!
- PRODOTTO TRA MATRICI (righe-colonne) $\underset{n \times n}{A} \cdot \underset{n \times p}{B} = \underset{n \times p}{D}$

  con $d_{ij} = \underset{1 \times n}{(a_{i1}, \dots, a_{in})} \underset{n \times 1}{\begin{bmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{bmatrix}} = \sum_{k=1}^{n} a_{ik} b_{kj}$

Scanned by CamScanner

• PROPRIETÀ TRASPOSTA

• $(A^T)^T = A$

• $(A+B)^T = A^T + B^T$

• $(AB)^T = B^T A^T$ inverte l'ordine ⚠

• L.I. : m vettori sono LINEARMENTE INDIPENDENTI ⟺ la loro C.L. è $= 0$ solo ponendo tutti i pesi $= 0$.

⇒ OSS Se uno degli m vettori è nullo ⇒ non sono L.I.

OSS Se 2 vettori sono uguali o proporzionali tra loro ⇒ i 2 vettori sono L.D.

OSS Se m vettori sono L.D. ⇒ Almeno uno di essi può essere scritto come C.L. degli altri.

• RANGO DI UNA MATRICE

Data una Matrice Rettangolare $A_{(m \times n)}$ si dicono

• RANGO RIGA di A il numero di vettori riga L.I.

• RANGO COLONNA di A il numero massimo di vettori colonna L.I.

⇒ Th RANGO RIGA e RANGO COLONNA DI UNA MATRICE COINCIDONO ⇒ sono detti RANGO di A

OSS $RANK(A) \leq MIN(m, n)$

• SE $RANK(A) = MIN(m, n)$ ⇒ A è DETTA A RANGO PIENO

• SE $RANK(A) \lneq MIN(m, n)$ ⇒ A è DETTA SINGOLARE

OSS Una MATRICE QUADRATA $A_{(n \times n)}$ è detta a RANGO PIENO ⟺ tutti i suoi vettori colonna (⟺ riga) sono L.I.

• PROPRIETÀ RANGHI

• $RANK(A) = RANK(A^T)$

• $RANK(AB) \leq MIN[RANK(A), RANK(B)]$

• $RANK(A \cdot A^T) = RANK(A^T A) = RANK(A)$

OSS Per convenzione il RANGO DI UNA MATRICE NULLA $\doteq 0$.

OSS UN VETTORE con almeno un elemento $\neq 0$, ha RANGO $= 1$

Scanned by CamScanner

• MATRICE INVERSA di A se ∃ è la matrice t.c.

$AA^{-1} = A^{-1}A = I$

OSS L'INVERSA DI UNA MATRICE ∃ ⟺ LA MATRICE È QUADRATA E NON SINGOLARE (A RANGO PIENO) ⇒ LA MATRICE SI DICE INVERTIBILE

• INVERSA DI UNA MATRICE $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (2x2) NON SINGOLARE È:

$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

• PROPRIETÀ DELLA MATRICE INVERSA

• $I^{-1} = I$

• $\text{DIAG}(a_1, \dots, a_m)^{-1} = \text{DIAG}(1/a_1, \dots, 1/a_m)$

• $(A^T)^{-1} = (A^{-1})^T$

• $(AB)^{-1} = B^{-1}A^{-1}$ (se $\exists A^{-1}, B^{-1}$)

Scanned by CamScanner

12/10/15

• MODELLO di REGRESSIONE LINEARE

• $y_i = \beta_0 + \beta_1 (X_1)_i + \dots + \beta_k (X_k)_i + u_i$ (relazione lineare) $i = 1, \dots, n$

• $y_i = \beta_1 \cdot (X_1)_i + \dots + \beta_k \cdot (X_k)_i + u_i$ $i = 1, \dots, n$ $\Rightarrow$

Riscrivo relazione in forma + compatta:

con $\underset{k\times 1}{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}$ $\underset{k\times 1}{X_i} = \begin{bmatrix} (X_1)_i \\ \vdots \\ (X_k)_i \end{bmatrix}$

$\Rightarrow$ • $y_i = X_i^T \beta + u_i$ $i = 1, \dots, n$ $\Rightarrow$

→ riscrivo questo sistema di n equazioni ($i = 1, \dots, n$) in forma matriciale.

$\Rightarrow \begin{cases} y_1 = X_1^T \cdot \beta + u_1 \\ \vdots \\ y_n = X_n^T \cdot \beta + u_n \end{cases} \Rightarrow \underset{n\times 1}{y} = \begin{bmatrix} X_1^T \cdot \beta + u_1 \\ \vdots \\ X_n^T \cdot \beta + u_n \end{bmatrix} = \begin{bmatrix} X_1^T \cdot \beta \\ \vdots \\ X_n^T \cdot \beta \end{bmatrix} + \underset{n\times 1}{u}$

con $\underset{n\times 1}{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$; $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ $\Rightarrow$

$\Rightarrow \underset{n\times 1}{y} = \underset{n\times k}{X} \cdot \underset{k\times 1}{\beta} + \underset{n\times 1}{u}$ con $\underset{n\times k}{X} = \begin{bmatrix} X_1^T \\ \vdots \\ X_n^T \end{bmatrix} = \begin{bmatrix} (X_1)_1 & (X_2)_1 & \dots & (X_k)_1 \\ (X_1)_2 & (X_2)_2 & \dots & (X_k)_2 \\ \vdots & \vdots & & \vdots \\ (X_1)_n & (X_2)_n & \dots & (X_k)_n \end{bmatrix}$

• OLS

$\underset{n\times 1}{\hat{y}} = \underset{n\times k}{X} \underset{k\times 1}{\hat{\beta}}$ stime = approssimazioni lineari delle y (con $\hat{\beta}$ = coef OLS stimato)

• $y = \hat{y} + \hat{u}$ (residui) $\Rightarrow \hat{u} = y - \hat{y}$ ($\Rightarrow \hat{u}_i = y_i - \hat{y}_i$ $i = 1, \dots, n$)

• PROBLEMA: TROVARE $\hat{\beta}$ OLS cioè t.c. $\min_{\{\beta\}} \left[\sum_{i=1}^{n} u_i^2 \right] =$

$= \min_{\{\beta\}} \underset{1\times n}{u^T} \underset{n\times 1}{u} \doteq \min_{(\beta)} S(\beta)$ $\Rightarrow$ RISOLVO con F.O.C.

$S(\beta) = u^T u = (y - X\beta)^T (y - X\beta) = \sum_{i=1}^{n} (y_i - \beta_1 - \beta_2 X_{2i} - \dots - \beta_k X_{ki})^2$

$\Rightarrow$ FOC

FOC $\begin{cases} \frac{\partial S(\cdot)}{\partial \beta_1} = 0 \\ \vdots \\ \frac{\partial S(\cdot)}{\partial \beta_k} = 0 \end{cases} \Leftrightarrow \begin{cases} -2\sum_{i=1}^{n} u_i \cdot 1 = 0 \\ \vdots \\ -2\sum_{i=1}^{n} u_i \cdot x_{ki} = 0 \end{cases} \Leftrightarrow \underset{k \times n}{X^T} \cdot \underset{n \times 1}{\underline{u}} = \underset{k \times 1}{\underline{0}} \Leftrightarrow$

13/10/15

$\Leftrightarrow X^T(\underline{y} - X\underline{\beta}) = \underline{0} \Leftrightarrow$

$\Leftrightarrow X^T\underline{y} = X^TX\underline{\beta} \Leftrightarrow$ ($\exists (X^TX)^{-1}$) (SSE rango$(X) = k$ = pieno)

$\Leftrightarrow \hat{\underline{\beta}}_{OLS} = (X^TX)^{-1}X^T\underline{y}$ OLS stimatore dei MINIMI QUADRATI)

OSS Si può fare solo se X^TX è invertibile, ma $\underset{k \times n}{(X^T} \underset{n \times k}{X)}$ è matrice quadrata $k \times k$ $\Rightarrow$

Se rango$(X^TX) = k \Rightarrow$ è invertibile $\Rightarrow \exists!$ soluzione OLS.

(ma RANK(A^TA) = RANK(A)) ma rango$(X) = ?$

$\Rightarrow$ Verifichiamo nel nostro problema (H1):

(H1) è RANGO$(X) = k$ $\Rightarrow \exists! \hat{\beta}_{OLS}$ (perché $\exists! (X^TX)^{-1}$)

OSS Per fare stime serve anche $n >> k$ o almeno $n \geq k$

OSS (H1) MODELLO RANGO$(X) = k \Leftrightarrow$ (H1) COVARIATE NON PERFETTAMENTE COLLINEARI (perché rango$(X) = k$ significa colonne X L.I.) (= NON SI PUÒ SCRIVERE C.L. delle X = 0 con pesi $\neq \underline{0}$ tra loro) è equivalente

$\Rightarrow \hat{\underline{y}} = X\hat{\underline{\beta}} = [X(X^TX)^{-1}X^T]\underline{y}$ = stima lin di $\underline{y}$ = $\stackrel{\triangle}{=} M\underline{y}$ (con $M \stackrel{\triangle}{=} X(X^TX)^{-1}X^T$)

$\Rightarrow \hat{\underline{u}} = \underline{y} - \hat{\underline{y}} = \underline{y} - M\underline{y} = (I - M)\underline{y}$

Scanned by CamScanner

⇒ ● PROPRIETÀ DEI RESIDUI OLS $\hat{\underline{u}}$

Poiché $X^T \hat{\underline{u}} = \underline{0}$ ⟺ $\begin{cases} \sum_{i=1}^{m} (X_1)_i \hat{u}_i = 0 \\ \vdots \\ \sum_{i=1}^{m} (X_k)_i \hat{u}_i = 0 \end{cases}$ ⇒ RESIDUI OLS HANNO MEDIA = 0

⟺ $\begin{cases} \hat{\underline{u}}^T \underline{X}_1 = 0 \\ \vdots \\ \hat{\underline{u}}^T \underline{X}_k = 0 \end{cases}$. cioè il vettore dei residui OLS è ORTOGONALE A TUTTE LE COLONNE DELLA MATRICE X: $\hat{\underline{u}} \perp \underline{X}_j$ $j=1,\dots,k$

▲ VETTORE ALEATORIO $\underset{(k\times 1)}{\underline{X}} = \begin{bmatrix} X_1 \\ \vdots \\ X_k \end{bmatrix}$ con $X_j \sim$ V.C, $j=1,\dots,k$

▲ VETTORE VALORE ATTESO $\triangleq \underset{(k\times 1)}{E[\underline{X}]} = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_k] \end{bmatrix}$ con $E(X_j)$ = VALORE ATTESO di X_j $j=1,\dots,k$

(ampliamento)

▲ MATRICE ALEATORIA $\underset{(m\times m)}{A} = [a_{ij}]$ con $a_{ij} \sim$ V.C $i=1,\dots,m$ $j=1,\dots,m$

▲ MATRICE VALORI ATTESI $\triangleq \underset{(m\times m)}{E[A]} = \underset{(m\times m)}{[E(a_{ij})]}$ con $E(a_{ij})$ = valore atteso a_{ij} $i=1,\dots,m$ $j=1,\dots,m$

▲ MATRICE (DI) VARIANZA DI $\underset{(k\times 1)}{\underline{X}}$ vettore ALEATORIO. $\triangleq \underset{(k\times k)}{\Sigma} =$ (matrice k×k) (vettore aleatorio k×1)

$= \underset{(k\times k)}{VAR(\underline{X})} \overset{\text{ALTERNAT.}}{=} \underset{(k\times k)}{COV(\underline{X})} \triangleq E[[\underline{X}-E(\underline{X})][\underline{X}-E(\underline{X})]^T]$

OSS con $\underline{a}^T \underline{a}$ = scalare (è prodotto interno a)
con $\underline{a}\,\underline{a}^T$ = MATRICE $k\times k$ (è prodotto esterno a)

OSS $VAR(X) = E[(X-E(X))^2]$ ⇒ $VAR(\underline{X})$ generalizzazione

es $K=2$
⇒ $VAR(\underline{X}) = E\left[\begin{bmatrix} X_1-E(X_1) \\ X_2-E(X_2) \end{bmatrix} \cdot [X_1-E(X_1) \;,\; X_2-E(X_2)]\right] =$

$= E\begin{bmatrix} (X_1-E(X_1))^2 & (X_1-E(X_1))(X_2-E(X_2)) \\ (X_2-E(X_2))(X_1-E(X_1)) & (X_2-E(X_2))^2 \end{bmatrix} = \begin{bmatrix} VAR(X_1) & COV(X_1,X_2) \\ COV(X_2,X_1) & VAR(X_2) \end{bmatrix}$

● PROPRIETÀ $\underset{k\times k}{\Sigma} = VAR(\underline{X}) = [\sigma_{ij}] = \begin{bmatrix} VAR(X_1) & COV(X_1,X_2) & \\ COV(X_2,X_1) & VAR(X_2) & \\ & & \ddots \\ & & & VAR(X_k) \end{bmatrix}$

- È SIMMETRICA (perché $\sigma_{ij} = \sigma_{ji}$)
- HA LE VARIANZE $VAR(X_1),\dots,VAR(X_k)$ sulla diagonale principale
  HA LE COVARIANZE altrove.

Scanned by CamScanner

▲ FUNZIONE VETTORIALE = VETTORE DI FUNZIONI

$$\underset{(m\times 1)}{\underline{y}} = \underset{(m\times 1)}{\underline{f}}(\underline{X}) \overset{def}{=} \begin{bmatrix} f_1(\underline{X}) \\ \vdots \\ f_m(\underline{X}) \end{bmatrix}$$ (scalare: $\underset{1\times 1}{y} = f(\underset{m\times 1}{\underline{x}})$)

$$\underline{f}: \mathbb{R}^m \to \mathbb{R}^m \qquad \underline{X} \in \mathbb{R}^m \xrightarrow{\underline{f}} \underline{y} \in \mathbb{R}^m$$

OSS caso + facile è:

▲ F. VETTORIALE LINEARE AFFINE

$$\underline{f}(\underline{x}) = \begin{bmatrix} f_1(\underline{x}) \\ \vdots \\ f_m(\underline{x}) \end{bmatrix} \Rightarrow \left(\begin{array}{l} \text{con } f_i(\underline{X}) \text{ f. lin. affini di } \underline{X} \quad \forall_{i=1,\dots,m} \\ \text{cioè } f_i(\underline{X}) = c_i + \sum_{j=1}^{m} a_{ij} X_j \quad i=1,\dots,m \end{array} \right.$$

vettore costante, matrice A

$$\Rightarrow \underset{m\times 1}{\underline{f}}(\underline{x}) = \underset{m\times 1}{\underline{c}} + \underset{m\times m}{A}\,\underset{m\times 1}{\underline{X}}$$ (se $\underline{c} = \underline{0} \Rightarrow$ F. VETTORIALE LINEARE)

Th se $\underline{f}$ F. VETTORIALE LIN. AFFINE DI $\underline{X}$ SI PUÒ SCRIVERE SEMPRE

$\Leftrightarrow \underline{f}(\underline{X}) = \underline{c} + A\underline{X}$ (TRASF. LIN. DI VETTORE $\underline{X}$)

$$\Rightarrow \boxed{\underset{m\times 1}{\underline{y}} = \underset{m\times 1}{\underline{c}} + \underset{m\times m}{A}\,\underset{m\times 1}{\underline{X}}} = \underset{m\times 1}{\underline{f}}(\underline{X})$$

▲ SE $\underline{c}$, A SONO DETERMINISTICHE (costanti) e $\underset{m\times 1}{\underline{X}}$ è VETTORE ALEATORIO $\Rightarrow \underset{m\times 1}{\underline{y}}$ è VETTORE ALEATORIO (TRASF. LIN. $\underline{X}$)

$$\Rightarrow E[\underline{y}] \overset{LIN}{=} \underline{c} + A \cdot E(\underline{X})$$ (N.B. $\neq \underline{c} + E(\underline{X}) \cdot A$ NO COMMUTATIVA PRODOTTO MATRICI)

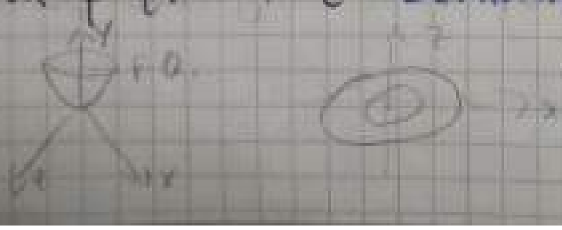
$$VAR[\underset{m\times 1}{\underline{y}}] = VAR[\underset{m\times 1}{\underline{c}} + \underset{m\times m}{A}\,\underset{m\times 1}{\underline{X}}] \overset{\text{invarianza per traslazione}}{=} VAR[\underset{m\times m}{A}\,\underset{m\times 1}{\underline{X}}] \overset{\text{Th}}{=} \underset{m\times m}{A}\,\underset{m\times m}{VAR(X)}\,\underset{m\times m}{A^T}$$

▲ FORME QUADRATICHE (F.Q.) CON 2 VARIABILI

$$\underset{1\times 1}{y} = f(X,Z) = aX^2 + bZ^2 + \frac{2c}{2}XZ$$ (F.Q. IN 2 VARIABILI X, Z)

$\Rightarrow (X=0, Z=0) \Rightarrow y = f(0,0) = 0$

• LA f (F.Q.) È DEFINITA POSITIVA SE $f(X,Z) > 0 \quad \forall (X,Z) \neq (0,0)$



curve di livello ellittiche.

- LA F.Q. F $\bar{e}$ DEFINITA NEGATIVA $\Leftrightarrow$ $f(x,z)<0 \quad \forall \binom{x}{z} \neq \binom{0}{0}$
- LA F.Q. F $\bar{e}$ SEMIDEFINITA POSITIVA (NEGATIVA) $\Leftrightarrow$
$f(x,y) \geq (\leq)\ 0 \quad \forall (x,z)$
- LA F.Q. F $\bar{e}$ INDEFINITA $\Leftrightarrow$
$\exists \binom{x}{y}$ t.c. $f(x,y) \gtrless 0$

▲ FORMA QUADRATICA NOTAZIONE MATRICIALE

$\underset{2\times1}{v} = \binom{x}{z} \qquad \underset{2\times2}{Q} = \begin{bmatrix} a & c \\ c & b \end{bmatrix} = \bar{e}$ SIMMETRICA

$\Rightarrow \underset{1\times1}{y} = f(v) = \underset{1\times2}{v^T}\ \underset{2\times2}{Q}\ \underset{2\times1}{v} = ax^2+bz^2+2cxz$ c.v.d.

OSS La forma quadratica $\bar{e}$ DEFINITA POSITIVA $\Leftrightarrow$ Q MATRICE SIMMETRICA $\bar{e}$ DEFINITA POSITIVA (Analogo SEMIDEFINITA POSITIVA).

▲ FORMA QUADRATICA (in generale con n variabili) Analogo.

$y = f(\underline{x}) = \underset{1\times n}{\underline{x}^T}\ \underset{n\times n}{Q}\ \underset{n\times1}{\underline{x}}$ con $\underset{n\times n}{Q}$ = MATRICE DEI PESI SIMMETRICA

$\cdot f(\underline{x}) = 0$ se $\underline{x} = 0$.

e analogamente DEFINIZIONI F.Q. DEFINITA (SEMI DEFINITA) POSITIVA/NEGATIVA

■ Th $cov(X) \doteq \Sigma_x$ ($\bar{e}$ SIMMETRICA) è DEFINITA POSITIVA $\Leftrightarrow$
$\Leftrightarrow x_1, \dots, x_n$ L.I. ($\bar{e}$ def. positiva salvo casi patologici)

■ Th $Q_{m\times m}$ $\bar{e}$ (SIMMETRICA) DEFINITA POSITIVA $\Rightarrow$ RANGO$(Q) = m \Rightarrow$
$\Rightarrow \exists!\ Q^{-1}$ e inoltre Q^{-1} $\bar{e}$ DEFINITA POSITIVA

■ Th se $\underset{m\times m}{Q}$ $\bar{e}$ SEMIDEFINITA-POSITIVA, MA NON DEF. POSITIVA,
$\Rightarrow$ RANGO$(Q) < m \Rightarrow Q$ $\bar{e}$ SINGOLARE

▲ FUNZIONE DI REGRESSIONE $E[Y|X] = g(X)$ V.C. UNIDIMENSIONALE

È F. DI REGRESSIONE DI Y RISPETTO A X.

CHE ASSUME COME POSSIBILI DETERMINAZIONI I VALORI ATTESI CONDIZIONATI PER OGNI POSSIBILE VALORE x DI X: $E[Y|X=x]$ = n° certo.

VEDO PROPRIETÀ F. REGRESSIONE SUI PDF

• Analogamente $\underset{1\times1}{E}[Y|\underline{X}] = g(\underline{X})$ V.C. unidim. è F. REG. DI Y RISPETTO A VETTORE ALEATORIO $\underset{n\times1}{\underline{X}}$ e $f_{\underline{X}}(x)$ è densità di $\underline{X}$.

(Oss: Si chiama ASSETTO INFORMATIVO l'info disponibile a cui mi condiziono)

che assume determinazioni $E[Y|\underline{X}=\underline{x}]$

• Analogamente $\underset{1\times1}{E}[Y|\underset{m\times m}{X}] = \underset{1\times m}{g(X)}$, con la differenza che l'informazione disponibile è una MATRICE X

(può cmq essere vista come caso precedente, mettendo tutti gli elementi della matrice in un vettore: ? = $\begin{bmatrix} x_{11} \\ \vdots \\ x_{1m} \\ x_{21} \\ \vdots \\ x_{2m} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{mm} \end{bmatrix}$)

▲ $\underset{m\times1}{E}[\underset{m\times1}{\underline{y}}|X] \overset{def}{=} \begin{bmatrix} E[Y_1|X] \\ \vdots \\ E[Y_m|X] \end{bmatrix}$ = F. REG. VETTORIALE (Vettore delle F. REG. UNIDIMENSIONALI)

con $\underset{m\times1}{\underline{y}} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_m \end{bmatrix}$ $\underset{q\times p}{X}$

Scanned by CamScanner

• $COV(\underline{X}) = VAR(\underline{X}) = E[(\underline{X}-E[\underline{X}])(\underline{X}-E(\underline{X}))^T]$ = MATRICE COVARIANZA.

▲ COVARIANZA TRA DUE VETTORI ALEATORI $\underline{X}, \underline{Y}$

$$\boxed{COV(\underset{m\times 1}{\underline{X}}, \underset{m\times 1}{\underline{Y}})} \triangleq E[(\underline{X}-E(\underline{X}))(\underline{Y}-E[\underline{Y}])^T] \triangleq C = [c_{ij}]$$

$m \times m$ il cui generico elemento $c_{ij} = COV(X_i, Y_j)$

OSS C contiene solo COVARIANZE (INCROCIATE) e non VARIANZE.

NOB $\underset{m\times m}{COV(\underline{X}, \underline{Y})} \neq \underset{m\times m}{COV(\underline{Y}, \underline{X})}$ ($\neq$ CASO UNIDIMENSIONALE)

ma $\exists$ RELAZIONE C , C^T (da una ricavo l'altra trasponendola)

OSS $VAR(\underline{X}) = COV(\underline{X}, \underline{X})$ (come unidimensionale)

PROP. X COST.

1) $COV(\underset{\text{costante}}{\underline{a}}, \underset{\text{vettore aleatorio}}{\underline{X}}) = 0$

2) $COV(\underset{q\times m}{A}\underset{m\times 1}{\underline{X}}, \underset{p\times m}{B}\underset{m\times 1}{\underline{Y}}) = \underset{q\times m}{A} \cdot \underset{m\times m}{COV(\underline{X}, \underline{Y})} \cdot \underset{m\times p}{B^T}$ (matrici costanti)

$\Rightarrow$ 3) $COV(\underset{q\times 1}{\underline{a}} + \underset{q\times m}{A}\underset{m\times 1}{\underline{X}} + C\underline{Z}; B\underline{Y}) = A\,COV(\underline{X}, \underline{Y})B^T + C\,COV(\underline{Z}, \underline{Y})B^T$.

1) $COV(\underline{a}, \underline{X}) = E[(\underline{0}) \cdot (\underline{X}-E(\underline{X}))^T] = E(\underline{0}) = 0$

2) $COV(A\underline{X}, B\underline{Y}) \overset{\text{DEF}}{=} E\{[A(\underline{X}-E(\underline{X}))] \cdot [B(\underline{Y}-E(\underline{Y}))]^T\} = A \cdot COV(\underline{X}, \underline{Y}) \cdot B^T$

3) $COV(a + A\underline{X} + C\underline{Z}; B\underline{Y}) = E\{[a + A\underline{X} + C\underline{Z}][B\underline{Y}]^T\} =$

$= 0 + A\,COV(\underline{X}, \underline{Y})B^T + C\,COV(\underline{Z}, \underline{Y})B^T$ c.v.d.

RIPASSO F. REGRESSIONE (NEW_LE F. REG... _PARTE II.DOCX)

▲ F. REG.: Considerati 2 V.A. Y e X, è detta "F. REGRESSIONE DI Y RISPETTO AD X" IL V.A. $E[Y|X]$ dipendente da X secondo una funzione $g(\cdot)$ (cioè $E[Y|X] = g(X)$) determinata dalla distribuzione subordinata di Y rispetto ad X.

• I valori del V.A. $g(X)$ ($\triangleq E[Y|X]$) sono le speranze matematiche condizionate $E[Y|X=x]$. Pertanto l'evento $\{g(X) = E[Y|X=x]\}$ coincide con l'evento "IL V.A. $E[Y|X]$ assume il valore $E[Y|X=x]$" ad esso ha probabilità pari a $Pr(X=x)$ quando il V.A. X è discreto; altrimenti ad esso è associata la densità di probabilità marginale del v.A. X nel punto x, cioè $f(x)$.

→ la conoscenza della f. reg $E[Y|X]$ permette di conoscere ∀ possibile valore di x assunto dal v.A. X quale sarà il valore medio (condizionato) di Y. Inoltre al variare di x, saremo in grado di capire come varia in media (condizionata) Y.

PROPRIETÀ DELLA F. REG.

1) $E[aY+bZ|X] = aE[Y|X]+bE[Z|X]$ LINEARITÀ

2) $E[E[Y|X]] = E[Y]$ LEGGE DEL VALORE ATTESO ITERATO

3) $E[Y \cdot h(X)|X] = h(X) \cdot E[Y|X]$ (⑥ $E[E[Y|X]|X] = E[Y|X]$)

4) $X \perp Y \Rightarrow E[Y|X] = E[Y]$

5) $E\{[Y-E[Y|X]]^2\} \leq E\{[Y-\varphi(X)]^2\}$ ∀ f. reale $\varphi(\cdot)$ t.c. $E\{[\varphi(X)]^2\} < +\infty$

6) $E[E[Y|X]|X,Z] = E[E[Y|X,Z]|X] = E[Y|X]$

OSS Una def. ALTERNATIVA della f. reg. $E[Y|X]$ è v.A. $f(X)$ t.c.
$E\{[Y-E[Y|X]] \cdot \ell(X)\} = 0$ ∀ $\ell(\cdot)$ t.c. ∃ $E(\cdot)$.

• MODELLO DI REG. LIN. (leggere file Word)

Scanned by CamScanner

• MODELLO di REGRESSIONE: $Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + u_i$ $i=1,\dots,n$

$\Leftrightarrow$ $Y_i = \underline{X}_i^T \underline{\beta} + u_i$ $i=1,\dots,n$ con $\underline{X}_i = \begin{bmatrix} 1 \\ (X_1)_i \\ \vdots \\ (X_k)_i \end{bmatrix}$ $\underline{\beta} = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix}$

ASSUNZIONI MODELLO:

⓪ Assumiamo che ∃ una relazione LINEARE (APPROSSIMATIVA) TRA LA Y e le $X_1, \dots, X_k$ e di avere CAMPIONE CASUALE SEMPLICE
$\Rightarrow$ che i coefficienti β della relazione siano = (stessa relazione) $\forall i$ = INDIVIDUO del campione.
$\Rightarrow$ Vogliamo trovare relazione approssimata che vale nel campione e nella Popolazione

① $E[u_i \mid X_{0i}, X_{1i}, \dots, X_{ki}] \overset{NOTAZ.}{=} E[u_i \mid \underline{X}_i] = 0$ $\forall i$

② CAMPIONE I.I.D. cioè $(Y_i, X_{1i}, \dots, X_{ki}) \sim$ I.I.D. $\forall i$
(Notazione impropria che useremo $(Y_i, \underline{X}_i^T) \sim$ I.I.D. $\forall i$ in cui in realtà consideriamo solo parte aleatoria di $\underline{X}_i$)

③ CURTOSI FINITA
$0 < E[Y_i^4] < +\infty$ e anche per X_{ji}

④ ASSENZA DI PERFETTA COLLINEARITÀ

Calcoliamo $E[Y_i \mid \underline{X}_i]$ sfruttando ⓪ e ① e proprietà F. REG.:

$E[Y_i \mid \underline{X}_i] = E[\underline{X}_i^T \underline{\beta} + u_i \mid \underline{X}_i] \overset{LIN}{=} E[\underline{X}_i^T \underline{\beta} \mid \underline{X}_i] + E[u_i \mid \underline{X}_i] =$
$\overset{HP ①}{=} E[\underbrace{\underline{X}_i^T \underline{\beta}}_{\text{UNIVOCO} \mid X_i} \mid \underline{X}_i] =$ (perché $E[m(\underline{X}_i) \mid \underline{X}_i] = m(\underline{X}_i)$) $= \underline{X}_i^T \underline{\beta}$ c.v.d.

$\Rightarrow E[Y_i \mid \underline{X}_i] = \overset{g_{LIN}}{g}(\underline{X}_i) = \underline{X}_i^T \underline{\beta} = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}$
MODELLO DI REGRESSIONE LINEARE (SUB HP F. REG. è LINEARE AFFINE assumiamo forma funzionale)

Scanned by CamScanner

$\Rightarrow$ INTERPRETAZIONE PARAMETRI (SUB HP)

$$\beta_j = \frac{\partial E[Y_i | \underline{X}_i]}{\partial X_j} \qquad j = 1, \dots, K$$

Se $\Delta X_j = +1$ CETERIS PARIBUS CI ASPETTIAMO VARIAZIONE DELLA MEDIA DI $Y = \beta_j$.

$\beta_j =$ DERIVATA F. REG. IN X_j. = EFFETTO CAUSALE PARZIALE

(ha senso se X_j continue.)

OSS Se invece voglio interpretare come varia f reg al variare di tutti i dati. $\underset{n \times k}{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{1k} \\ 1 & X_{n1} & \cdots & X_{nk} \end{bmatrix}$ è in generale $\neq$ CAMBIA INFO DISPONIBILE $X \neq \underline{X}_i$

$E[Y_i | \underset{n \times (k+1)}{X}] \neq E[Y_i | \underline{X}_i]$ in generale.

$\Big[$ Se HP v.c. Y, X, Z (HP $Y \perp\!\!\!\perp Z$)

$\Rightarrow E[Y | X, Z] = E[Y | X]$ $\Big]$

$\Big($ Ma HP ② $\Rightarrow (Y_i, X_{1i}, \dots, X_{ki}) \underset{i \neq j}{\perp\!\!\!\perp} (Y_j, X_{1j}, \dots, X_{kj})$

$\Rightarrow Y_i \underset{j \neq i}{\perp\!\!\!\perp} (X_{1j}, \dots, X_{kj})$ $\Big)$

$\Rightarrow E[Y_i | X] \overset{HP 2}{=} E[Y_i | \underline{X}_i] \overset{0,1}{=} \underline{X}_i^T \underline{\beta}$ c.v.d.

- $E[u_i | X] = E[u_i | \underline{X}_i] = 0$ (SUB HP 0, 1, 2)

DIM

OSS Se $E[Y_i | X] = \underline{X}_i^T \underline{\beta}$

$\Rightarrow Y_i = E[Y_i | X] + V_i$

con V_i t.c. $E[V_i | X] = 0$

$\Rightarrow Y_i = \underline{X}_i^T \underline{\beta} + V_i$

$\Rightarrow V_i = Y_i - \underline{X}_i^T \underline{\beta} = u_i$

$\Rightarrow E[u_i | X] = 0 = E[u_i | \underline{X}_i]$ c.v.d.

Scanned by CamScanner

[illegible] $E[y_i | \underline{X}_i] = \underline{X}_i^T \cdot \underline{\beta} = E[y_i | X]$ $\forall i$ 19/04/16

$$\Rightarrow E[\underline{y}|X] = \begin{bmatrix} E[y_1|X] \\ \vdots \\ E[y_n|X] \end{bmatrix} = X\underline{\beta}$$

Modello $y_i = \underline{X}_i^T \underline{\beta} + u_i$ $i = 1, \dots, n$ $\Rightarrow \underline{y} = X\underline{\beta} + \underline{u}$

$E[u_i | \underline{X}_i] = E[u_i | X] = 0$ $\forall i$ $\Rightarrow E[\underline{u}|X] = \underline{0}_{n \times 1}$

$$\Rightarrow E[\underline{y}|X] \overset{\text{LIN}}{=} E[X\underline{\beta} + \underline{u} | X] = X\underline{\beta} + 0 = X\underline{\beta} \quad \text{c.v.d.}$$

OSS Modello $y_i = \underline{X}_i^T \underline{\beta} + u_i$

IP 1 $E[u_i | \underline{X}_i] = 0 \Rightarrow$ valgono alcune PROPRIETÀ SULL'ERRORE:

(a) $E[E[u_i | \underline{X}_i]] =$ (legge valore atteso iterato) $= E[u_i] = 0$

(b) $COV[u_i, X_{ji}] = 0$ $\forall j = 1, \dots, K$

DIM per def e poiché $E(u_i) = 0 \Rightarrow COV(u_i, X_{ji}) = E[u_i \cdot X_{ji}] =$ (legge valore atteso iterato) $= E[E(u_i \cdot X_{ji} | \underline{X}_i)] = E[X_{ji} \cdot E[u_i | \underline{X}_i]] = E[0] = 0$ c.v.d.

OSS In generale $E[A | B] = 0 \Rightarrow COV(A, B) = 0$ (A, B) incorrelate

- Modello aggregato $\underline{y} = X \cdot \underline{\beta} + \underline{u}$

$\Rightarrow$ Riscrivo IP in termini dell'errore (equivalente in termini $\underline{y}$)

(A) $E[\underline{u} | X] = \underline{0}$ $\{ \Rightarrow \underline{u}$ e X sono incorrelati: $COV(u_i, X_{jk}) = 0$ $\forall i, k = 1, \dots, n$, $\forall j = 1, \dots, K$

$(\Rightarrow E[\underline{u}] = \underline{0})$

(B) $VAR[\underline{u} | X]$ = (OSS $X\underline{\beta} = E[\underline{y}|X]$) $= VAR[\underline{y} | X] = E[(\underline{y} - E(\underline{y}|X))(\underline{y} - E(\underline{y}|X))^T | X]$

DIM $= E[\underline{u} \cdot \underline{u}^T | X]$ (poiché $E[\underline{u}] = 0$).

$\Rightarrow \underline{y} - E(\underline{y}|X) = \underline{u}$

OSS $VAR(\underline{u}) \neq VAR(\underline{y})$ (si spera che il modello spieghi qualcosa) $\Rightarrow$ Solo $|X$

$$\underset{n \times n}{VAR}[\underset{n \times 1}{\underline{u}} | X] = \begin{bmatrix} VAR[u_1|X] & & COV(u_1, u_n|X) \\ & VAR[u_2|X] & \\ COV(u_n, u_1|X) & & VAR[u_n|X] \end{bmatrix}$$

i cui elementi sono $VAR(u_i | X) = VAR(y_i | X)$ e $COV(u_i, u_j | X) = COV(y_i, y_j | X)$

OSS DIM che $COV(y_i, y_j | X) = 0$ $\forall i \neq j$.

$COV(u_i, u_j | X) = 0$ $\forall i \neq j$

Scanned by CamScanner

(III) $COV(u_i, u_j | X) = E[(u_i - E[u_i|X]) \cdot (u_j - E[u_j|X]) | X] =$ ($i \neq j$)

$= E[u_i \cdot u_j | X]$ (oss con $u_i = \varepsilon(y_i, \underline{x}_i)$ con funzione y_i ...

(condizionando a X ... $\underline{x}_i$ è noto ... noto ... y_i)

(ora oss $y_i \perp\!\!\!\perp \underline{x}_j$ $\forall j \neq i$ perché CCIS)

$= E[u_i \cdot u_j | X] =$ (perché $E[X|Z,W] = E[X|Z]$ se $X \perp\!\!\!\perp W$)

$= E[u_i \cdot u_j | \underline{X}_i, \underline{X}_j] = COV(y_i, y_j | \underline{X}_i, \underline{X}_j) =$ (perché CCIS $y_i \perp\!\!\!\perp y_j ...$)

(oss se $W \perp\!\!\!\perp Z \Rightarrow E[W \cdot Z] = E[W] \cdot E[Z]$ e

oss se $(W|X) \perp\!\!\!\perp (Z|X) \Rightarrow E[W \cdot Z|X] = E[W|X] \cdot E[Z|X]$)

$= E[u_i | \underline{X}_i, \underline{X}_j] \cdot E[u_j | \underline{X}_i, \underline{X}_j] = E[u_i | \underline{X}_i] \cdot E[u_j | \underline{X}_j] = 0 \cdot 0 = 0$

$\Rightarrow$ oss $COV(u_i, u_j | X) = COV(u_i, u_j | \underline{X}_i, \underline{X}_j) = COV(y_i, y_j | X) =$

$= COV(y_i, y_j | \underline{X}_i, \underline{X}_j) = 0$

$\Rightarrow$ MATRICE $VAR(\underline{u} | X)$ è MATRICE DIAGONALE

oss $VAR[u_i | X] =$ (perché $E[u_i|X] = 0$) $= E[u_i^2 | X] =$ (perché $u_i \perp\!\!\!\perp \underline{x}_j$ $\forall j \neq i$)

$= E[u_i^2 | \underline{X}_i] = VAR[u_i | \underline{X}_i] =$ (perché CCIS ...)

$= VAR[y_i | \underline{X}_i] = VAR(y_i | X)$

$\Rightarrow$ RIASSUMENDO:

$\underline{y} = X\underline{\beta} + \underline{u}$)

(*) $E[\underline{u} | X] = 0$

(B) $VAR[\underline{u} | X] = DIAG[VAR(u_1 | \underline{X}_1), \ldots, VAR(u_n | \underline{X}_n)]$ (perché $VAR(u_i|X) = VAR(u_i|\underline{X}_i)$)

oss sotto 4 HP non si può dire di più in termini di FREQ

(per dire di più: $VAR(u_i | \underline{X}_i) = cost$ serve 5 HP)

SOTTO 4 (H) ⇒ PROPRIETÀ STIMATORE: $\underset{k\times 1}{\hat{\underline{\beta}}_{OLS}} \doteq (X^T X)^{-1} X^T \underset{n\times 1}{\underline{y}}$

$= (X^T X)^{-1} X^T (X \cdot \underline{\beta} + \underline{u}) = I \cdot \underline{\beta} + \underbrace{(X^T X)^{-1} X^T \underline{u}}_{\hat{\beta} - \beta \text{ = errore di stima}}$

[▲ STIMATORE CORRETTO o NON DISTORTO $\hat{\beta}$ ⇔ $E[\hat{\underline{\beta}}] = \underline{\beta}$
(cioè $E[\hat{\underline{\beta}} - \underline{\beta}] = 0$ = valore atteso dell'errore di stima)

1) $\hat{\underline{\beta}}_{OLS}$ È CORRETTO

DIM oss. nei nostri modelli X sono v.c. ⇒ la DIM è utile oss. prima che

I) STIMATORE È CORRETTO CONDIZIONATAMENTE A X:
cioè $E[\hat{\underline{\beta}} | X] = \underline{\beta}$

DIM
$E[\hat{\underline{\beta}} | X] = E[\underline{\beta} + (X^T X)^{-1} X^T \underline{u} | X] \overset{\text{LIN}}{=} \overset{\text{COST}}{E[\underline{\beta} | X]} +$
$+ E[(X^T X)^{-1} X^T \underline{u} | X] = \underline{\beta} + (X^T X)^{-1} X^T \cdot E[\underline{u} | X] = \underline{\beta} + 0 = \underline{\beta}$ c.v.d.

(OSS se $E[u_i | \underline{X}_i] = 0$, ma $E[u_i | X] \neq 0$ ⇒ $\hat{\beta}_{OLS}$ non è corretto)

II) $E[\hat{\underline{\beta}}] = E[\underbrace{E[\hat{\underline{\beta}} | X]}_{\beta}] = \underline{\beta}$ c.v.d.

[▲ STIMATORE $\hat{\beta}$ È CONSISTENTE ⇔ $\hat{\beta} \xrightarrow[n\to+\infty]{P} \beta$ (p-lim $\hat{\beta} = \beta$)

▲ CONVERGENZA IN PROBABILITÀ (X C.A.)

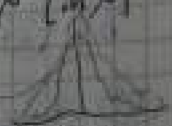
OSS LIMITE IN PROB DI UN VETTORE È IL VETTORE DEI LIMITI IN PROB DELLE SUE COMPONENTI:

$$\text{p-lim}(\hat{\underline{\beta}}) = \begin{bmatrix} \text{p-lim } \hat{\beta}_1 \\ \vdots \\ \text{p-lim } \hat{\beta}_k \end{bmatrix} \qquad \Rightarrow \hat{\underline{\beta}} \xrightarrow[n\to+\infty]{P} \underline{\beta} \iff \hat{\underline{\beta}} \text{ è consistente}$$

▲ OPERATORE P-LIM: $\underset{n\to+\infty}{\text{p-lim}} (\hat{\beta}_{(n)}) = \beta \iff \lim_{n\to+\infty} \Pr\{ |\hat{\beta}_{(n)} - \beta| > \delta \} = 0 \quad \forall \delta > 0$

cioè $\hat{\beta}_{(n)} \xrightarrow[n\to+\infty]{P} \beta$

OSS CONVERGENZA IN PROB. significa intuitivamente che crescendo n (cioè $n\uparrow$)
⇒ ~~distribuzione~~ distribuzione è sempre più concentrata nel parametro.



OSS COND. SUFF. PER LA CONSISTENZA DI $\hat{\beta}$: (non nec.)

① ASINTOTICA CORRETTEZZA: $\lim_{n\to+\infty} E[\hat{\beta}_n] = \beta$ (⇐ correttezza)
② $\lim_{n\to+\infty} VAR(\hat{\beta}_n) = 0$

Scanned by CamScanner

20/10/15

OSS $p\text{-}\lim(\hat{\beta}_{OLS}) = \beta$ consistenza stimatore $\hat{\beta}$

OSS analogo consistenza (e conv. in prob.) per qualsiasi vettori aleatori e matrici aleatorie (per tutte le componenti)

$\hat{\Sigma} \xrightarrow{p} \Sigma$

• $p\text{-}\lim(\hat{\beta}_{OLS}) = p\text{-}\lim\left(\beta + (X^TX)^{-1} X^T \underline{u}\right) =$ (p-lim somma = somma p-lim)

$= p\text{-}\lim(\beta) + p\text{-}\lim\left((\tfrac{1}{n}X^TX)^{-1}(\tfrac{1}{n}X^T\underline{u})\right) = \beta + p\text{-}\lim\left[(\tfrac{1}{n}X^TX)^{-1}\right] \cdot$

$\cdot\, p\text{-}\lim\left[(\tfrac{1}{n}X^T\underline{u})\right]$ (perché p-lim prodotto è prodotto p-lim)

$=$ (matrice inversa è funz. continua e quindi p-lim conserva)

$\beta + \left[p\text{-}\lim(\tfrac{1}{n}X^TX)\right]^{-1} \cdot p\text{-}\lim(\tfrac{1}{n}X^T\underline{u}) =$ (moltiplico e divido per n)

OSS $\frac{1}{n}X^T\underline{u} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^n (X_1)_i u_i \\ \frac{1}{n}\sum_{i=1}^n X_{2i} u_i \\ \vdots \\ \frac{1}{n}\sum_{i=1}^n (X_k)_i u_i \end{bmatrix}$ | $\bar{u} = \frac{1}{n}\sum_i u_i \xrightarrow{p} E[u_i] = 0$ (condizioni)

$\frac{1}{n}\sum X_{ji} \cdot u_i \xrightarrow{p} E[X_{ji} \cdot u_i] = cov(X_{ji}, u_i) = 0 \quad \forall j = 2, \dots, k$

$\Rightarrow p\text{-}\lim(\frac{1}{n}X^T\underline{u}) = \underline{0}$ (OSS momenti 4° momento finito $\Rightarrow$ var $< +\infty$)

$\Rightarrow \frac{1}{n}(X^TX) = \frac{1}{n}\sum_{i=1}^n \underline{X}_i \cdot \underline{X}_i^T = \hat{Q}_X = \left[\hat{q}_{\ell m} = \frac{1}{n}\sum_{i=1}^n (X_\ell)_i \cdot (X_m)_i\right] \longrightarrow$

$\xrightarrow[n\to\infty]{p} \left[E[(X_\ell)_i \cdot (X_m)_i]\right]$ (convergenza momenti empirici a teorici $\forall \ell, m$)

$\Rightarrow \hat{Q}_X \xrightarrow[n\to\infty]{p} Q_X = E[\underline{X}_i \cdot \underline{X}_i^T]$

$= \beta + [Q_X]^{-1} \cdot 0 = \beta$ C.V.D.

$\hat{\beta}_{OLS} \xrightarrow[n\to\infty]{p} \beta$

OSS abbiamo dim. consistenza stimatore OLS sotto (1):

1) momenti 4° finiti

2) $(X_j)_{j=1..k}$ L.I.

3) errore incorrelato coi regressori

OSS per la consistenza di $\hat{\beta}_{OLS}$ bastano:

$\begin{cases} E[u_i] = 0 & j=1 \\ cov((X_j)_i, u_i) = 0 & j = 2, \dots, k \end{cases}$ oltre (11) su Q_X

(2.15) MOLTO MENO DI QUELLE PER LA CORRETTEZZA $E[\underline{u}|X]=0$

$\left(E[\underline{u}|X]=0 \Rightarrow \begin{cases} E[u_i]=0 \\ COV(x_{ji}, u_i)=0 \quad j=1,\dots,k \end{cases} \right)$

OSS se $E[u_i|\underline{X}_i]=0 \wedge E[u_i|X]\neq 0$
$\Rightarrow$ STIMATORE È DISTORTO, MA CONSISTENTE

▲ PROPRIETÀ ASINTOTICA NORMALITÀ DI $\hat{\beta}_{OLS}$

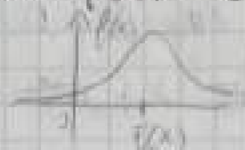
RICHIAMO: NORMALE MULTIVARIATA

• NORMALE UNIVARIATA $X \sim N(\mu, \sigma^2)$ con $\mu \triangleq E(X)$ $\sigma^2 \triangleq VAR(X)$

SUPPORTO $= \mathbb{R}$ DENSITÀ $= f(x) = (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \geq 0$

KERNEL (CARATTERIZZA DISTR.)

$f(x) \propto$ KERNEL



• NORMALE MULTIVARIATA:

SUPPORTO $= \mathbb{R}^m$ (con $E[\underline{X}]=\underline{\mu}$ $VAR[\underline{X}]=\Sigma$ DEF. POSITIVA)

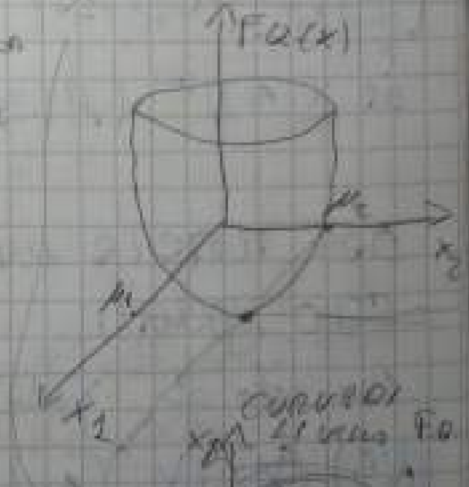
DENSITÀ CONGIUNTA $= f(\underline{x}) \propto \exp\left\{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \Sigma^{-1}(\underline{x}-\underline{\mu})\right\} \triangleq$ KERNEL

OSS $\Sigma \triangleq$ MATRICE VAR-COV $\Sigma^{-1} =$ MATRICE DI PRECISIONE

(OSS se Σ DEF. POS. $\Rightarrow \Sigma^{-1}$ DEF. POS. $\Rightarrow$ F.Q. DEF. POS)

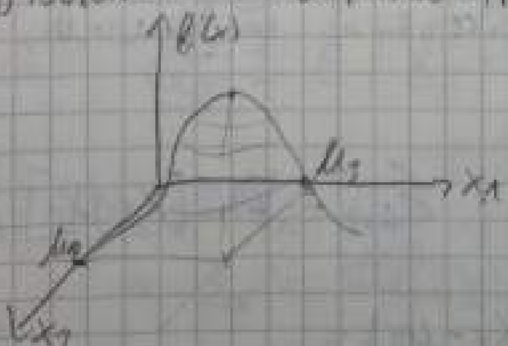
OSS F.Q. DEF. POS $\Rightarrow \underline{x}^T Q \underline{x} > 0 \quad \forall \underline{x} \neq \underline{0}$ pongo $\underline{v} \triangleq (\underline{x}-\underline{\mu})$

$\Rightarrow \underline{v}^T \Sigma^{-1} \underline{v} > 0 \quad \forall \underline{v} \neq \underline{0} \Rightarrow \forall \underline{x} \neq \underline{\mu}$



$\Rightarrow$ se moltiplico F.Q. $\cdot -\frac{1}{2} \Rightarrow$ ribalto (ottengo praticamente il paraboloide è girato) poi lo metto a esponente

$\Rightarrow$ NORMALE MULTIVARIATA



Scanned by CamScanner

- PROPRIETÀ NORMALE MULTIVARIATA:

▲ $\underset{n\times 1}{\underline{X}} \sim N^{(n)}(\underline{\mu}, \Sigma)$ $\underline{X}$ è CONGIUNTAMENTE NORMALE n-DIMENSIONALE definita da μ e Σ.

(N1) TRASF. LINEARI AFFINI DI $\underline{X}$ SONO ANCORA NORMALI MULTIVARIATE $\underset{m\times 1}{\underline{Y}} = \underset{m\times n}{A}\,\underset{n\times 1}{\underline{X}} + \underset{m\times 1}{\underline{b}} \Rightarrow \underline{Y} \sim N^{(m)}(A\cdot\underline{\mu}+\underline{b};\ A\Sigma A^T)$

OSS I MOMENTI LI CONOSCEVAMO GIÀ (N1) CI DICE IN + CHE È $N(\cdot)$

OSS LA NORMALE MULTIVARIATA È CHIUSA PER TRASF. AFFINI

(N3) SE $\underset{n\times 1}{\underline{X}} \sim N^{(n)}(\mu, \Sigma)$ CON $\Sigma = DIAG(\sigma_1^2, \dots, \sigma_n^2)$ (CIOÈ $\sigma_{ij}=0$ $i\neq j$) (CIOÈ X_j NON SONO CORRELATE TRA LORO)
$\Rightarrow X_j \perp\!\!\!\perp X_i$, $i\neq j$ (SE X_j CONGIUNTAMENTE NORMALI)

(N4) $\underset{n\times 1}{\underline{X}} \doteq \begin{bmatrix} \underset{k\times 1}{\underline{X}_1} \\ \underset{(n-k)\times 1}{\underline{X}_2} \end{bmatrix}$ $\underline{\mu} \doteq \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ $\underset{n\times n}{\Sigma} = \begin{bmatrix} \underset{k\times k}{\Sigma_{11}} & \underset{k\times(n-k)}{\Sigma_{12}} \\ \underset{(n-k)\times k}{\Sigma_{21}} & \underset{(n-k)\times(n-k)}{\Sigma_{22}} \end{bmatrix}$

(con $\underset{k\times k}{\Sigma_{11}} = VAR(\underline{X}_1)$ $\Sigma_{22} = VAR(\underline{X}_2)$ $\underset{k\times(n-k)}{\Sigma_{12}} = COV(\underline{X}_1, \underline{X}_2) = \Sigma_{21}^T$
$\underset{(n-k)\times k}{\Sigma_{21}} = COV(\underline{X}_2, \underline{X}_1)$)

SE $\underline{X} \sim N^{(n)}(\underline{\mu}, \Sigma)$ $\Rightarrow X_1 \sim N^{(k)}(\mu_1, \Sigma_{11})$
$X_2 \sim N^{(n-k)}(\mu_2, \Sigma_{22})$

CIOÈ TUTTE LE MARGINALI DI UNA NORMALE MULTIVARIATA SONO NORMALI. (OSS Non vale il viceversa in generale, ma solo se $X_1 \perp\!\!\!\perp X_2$)

(N5) $(\underset{k\times 1}{\underline{X}_1} | \underset{(n-k)\times 1}{\underline{X}_2}) \sim N^{(k)}\Big(\underset{\text{F. REG VERA}}{\overset{\mu_{1|2}}{E[\underline{X}_1|\underline{X}_2]}}\ ;\ \underset{k\times k}{\Sigma_{1|2}} \doteq VAR[\underline{X}_1|\underline{X}_2]\Big)$

con $\mu_{1|2} = (\text{LINEARE}) = \underline{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}_2 - \underline{\mu}_2) =$
$= \underbrace{(\underline{\mu}_1 - \Sigma_{12}\Sigma_{22}^{-1}\underline{\mu}_2)}_{\substack{COST \\ k\times 1}} + \underbrace{\Sigma_{12}\Sigma_{22}^{-1}}_{k\times k}\underset{k\times 1}{\underline{X}_2}$ TRASF. AFFINE $\underline{X}_2$

Scanned by CamScanner

$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ (NO STUDIO MEMORIA)

22/10/15

OSS Sulle k_i (HP)

OLS È CORRETTO, CONSISTENTE e ASINTOTICAMENTE NORMALE

OSS ASINTOTICAMENTE NORMALE.

(N.B.) All'aumentare del campione all'infinito $\hat{\beta}_{\infty}$ si concentra in una costante $\Rightarrow \hat{\beta}_{\infty}$ è DISTR. DEGENERE.

$\Rightarrow$ COSA VUOL DIRE ASINTOTICA NORMALITÀ?

$\hat{\beta}_n \not\xrightarrow{n\to\infty}$ A NORMALE, ma a costante $\Rightarrow$

$\Rightarrow$ IDEA: trovo trasf. di $\hat{\beta}_n$ t.c. $VAR(\hat{\beta}_n) \not\to 0 \Rightarrow$ distr. non più degenere. $E(\hat{\beta}_n)$ perché corretto

- $\sqrt{n}\,(\hat{\underline{\beta}}_n - \underline{\beta}) \xrightarrow[n\to\infty]{d} N(\underline{0}, \Sigma_{\sqrt{n}(\hat{\beta}-\beta)})$

NORMALITÀ ASINTOTICA

(con $\Sigma_{\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta})} \triangleq AVAR$ = Matrice Var Asintotica di $\sqrt{n}(\hat{\beta}-\beta)$ o $\sqrt{n}\,\hat{\underline{\beta}}$)

PROPRIETÀ DI ASINTOTICA NORMALITÀ

OSS Noi lavoreremo sempre con campioni finiti

$\Rightarrow$ usiamo risultato approssimato:

- $\hat{\underline{\beta}}_n \underset{(n>>0)}{\approx} N(\underline{\beta}, \Sigma_{\hat{\underline{\beta}}})$

(con $\Sigma_{\hat{\underline{\beta}}} = AVAR(\hat{\underline{\beta}}) \underset{(n>>0)}{\simeq} \frac{1}{n}\Sigma_{\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta})}$)

Th $\Sigma_{\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta})} = Q_X^{-1}\Sigma_V Q_X^{-1}$

$\Rightarrow \Sigma_{\hat{\underline{\beta}}} = \frac{1}{n} Q_X^{-1}\Sigma_V Q_X^{-1}$

con $\underset{k\times k}{Q_X} \triangleq E[\underline{X}_i \cdot \underline{X}_i^T]$ il cui generico elemento è $q_{\ell m} = E[X_{\ell i} \cdot X_{m i}]$

$\underline{V}_i = \underline{X}_i \cdot u_i$ (regr. del modello $Y = X^T\beta + u_i$)

$\Rightarrow \Sigma_V = VAR(\underline{V}_i)$ uguale $\forall i$ (NO DIM)

● PROPRIETÀ Σ_V SUB ... MODELLO

- $\underline{V}_i \doteq \underline{X}_i \cdot u_i$ I.I.D. $\forall i$
- $E[\underline{V}_i] = \begin{bmatrix} E[u_i] \\ E[X_{2i} \cdot u_i] \\ \vdots \\ E[X_{Ki} \cdot u_i] \end{bmatrix} = \underline{0}$
- $\Sigma_V = VAR(V_i) = E[\underline{V}_i \cdot \underline{V}_i^T] = \quad (= \forall i)$
  $= E[\underline{X}_i \cdot \underline{X}_i^T \cdot u_i^2]$
- $Q_X = E[\underline{X}_i \cdot \underline{X}_i^T]$

OSS In generale Σ_V e Q_X IGNOTE
$\Rightarrow \Sigma_{\sqrt{n}(\hat{\beta}-\beta)}$ IGNOTA
$\Rightarrow$ Nella pratica va stimata
$\Rightarrow$ Cerchiamo stimatore CONSISTENTE (anche distorto) della matrice $\Sigma_{\sqrt{n}(\hat{\beta}-\beta)}$ cioè:

$$\hat{\Sigma}_{\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta})} \xrightarrow{P} \Sigma_{\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta})}$$

{PER (Th) se stimo con stimatori consistenti i fattori
{ $\Rightarrow$ è stimatore consistente anche il prodotto

OSS IDEA: Sostituisco i corrispondenti momenti campionari

- $\hat{Q}_X \doteq \frac{1}{n} \sum_{i=1}^{n} \underline{X}_i \cdot \underline{X}_i^T \xrightarrow{P} Q_X \doteq E[\underline{X}_i \cdot \underline{X}_i^T]$
  È STIMATORE CONSISTENTE
- $\hat{\Sigma}_V \doteq \frac{1}{n} \sum_{i=1}^{n} \underline{X}_i \cdot \underline{X}_i^T \cdot \hat{u}_i^2 \longrightarrow \Sigma_V \doteq E[\underline{X}_i \cdot \underline{X}_i^T \cdot u_i^2]$ ← ERRORI IGNOTI ⇒ STIMATI

(OSS con $\hat{u}_i \doteq y_i - \underline{x}_i^T \hat{\underline{\beta}} \xrightarrow{P} u_i \doteq y_i - \underline{X}_i^T \underline{\beta}$
(perché OLS È STIMATORE CONSISTENTE)

È STIMATORE CONSISTENTE

Scanned by CamScanner

OSS Posso sostituire a $\sum_{i=1}^{n} \underline{X_i} \cdot \underline{X_i}^T \triangleq X^T X$

$$\Rightarrow \hat{\Sigma}_{\hat{\beta}} = \frac{1}{n}\left(\frac{1}{n} X^T X\right)^{-1} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} \underline{X_i} \underline{X_i}^T \cdot \hat{u}_i^2\right) \cdot \left(\frac{1}{n} X^T X\right)^{-1}$$

▲ FORMULA di WHITE $= (X^T X)^{-1} \underbrace{\left(\sum_{i=1}^{n} \underline{X_i} \underline{X_i}^T \hat{u}_i^2\right)}_{X^T \operatorname{diag}(\hat{u}_i^2) X} (X^T X)^{-1}$

$\bar{e}$ STIMATORE CONSISTENTE DI $\Sigma_{\hat{\beta}} = AVAR(\hat{\beta}_{OLS})$

(OSS in campioni piccoli al posto di $\frac{1}{n}$ metto $\frac{1}{n-k}$)

OSS LA FORMULA DI WHITE = STIMATORE $\bar{e}$ ABBIAMO SOTTO 4 HP MODELLO, MA NON SERVE (5) OMOSCHEDASTICITÀ

E ANCHE PER (DIM) $\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Sigma_{\sqrt{n}(\hat{\beta}-\beta)})$

(con $\Sigma_{\sqrt{n}(\hat{\beta}-\beta)} = Q_X^{-1} \Sigma_V Q_X^{-1}$) NON SERVE HP (5)

OSS STIMATORE DI WHITE $\bar{e}$ STIMATORE CONSISTENTE DELLA MAT VAR ASINTOTICA DI OLS (SOTTO 4 HP) SIA IN CASO DI ETEROSCHEDASTICITÀ E OMOSCHEDASTICITÀ $\Rightarrow$ Si dice STIMATORE DI WHITE $\bar{e}$ ROBUSTO ALL'ETEROSCHEDASTICITÀ

Th SOTTO HP 1÷4 ∪ (5) OMOSCHEDASTICITÀ

$\Rightarrow \Sigma_V = \sigma_u^2 Q_X$

$\Rightarrow AVAR(\hat{\beta}) = \sigma_u^2 \cdot Q_X^{-1}$ $\left(Q_X^{-1} \sigma_u^2 \cancel{Q_X} \cancel{Q_X^{-1}}\right)$

OSS Se c'$\bar{e}$ HP (5) OMOSCHEDASTICITÀ posso trovare FORMULA PIÙ SEMPLICE DI FORMULA DI WHITE

$\Rightarrow$ ▲ FORMULA CLASSICA $\hat{\Sigma}_{\hat{\beta}_{classica}} = \frac{1}{n}\left(\hat{\sigma}_u^2 \left(\frac{1}{n} X^T X\right)^{-1}\right)$

$\bar{e}$ STIMATORE CONSISTENTE POICHÉ RAPPORTO DI STIMATORI CONSISTENTI (SOLO SE C'$\bar{e}$ OMOSCHEDASTICITÀ)

Scanned by CamScanner

▲ $SCR = \left(\sqrt{\frac{\sum \hat{u}_i^2}{n-k}}\right)^2 \xrightarrow{d/} S^2 \xrightarrow{P} \sigma_u^2$

(SUB HP 1÷4 ∧ ⑤ OMOSCHEDASTICITÀ)

VARIANZA CAMPIONARIA CORRETTA e STIMATORE CONSISTENTE (per eccesso)

DI σ_u^2

OSS ⇒ FORMULA DI WHITE SOTTO HP ⑤ OMOSCHEDASTICITÀ PORTERÀ AGLI STESSI RISULTATI ASINTOTICI DELLA FORMULA CLASSICA (in quanto entrambi stimatori consistenti)

OSS formula classica va bene solo sotto HP ⑤

Formula di White va bene sempre.

Th (GAUSS-MARKOV) (VERSIONE $|X$ perché X = v.c. in generale ignota)

SUB HP ①÷④ ∧ (⑤ OMOSCHEDASTICITÀ) CIOÈ.

$$\begin{cases} \underline{Y} = X\underline{\beta} + \underline{u} \quad \text{Modello con} \\ (a)\ E[\underline{u} \mid X] = 0 \\ (b)\ VAR[\underline{u} \mid X] = DIAG(\sigma_u^2, \dots, \sigma_u^2) = \sigma_u^2 I \end{cases}$$

(b) HP 5

⇒ CONDIZIONATAMENTE A X

LO STIMATORE OLS È BLUE = BEST LINEAR UNBIASED ESTIMATOR.

CIOÈ TRA TUTTI GLI STIMATORI DI $\underline{\beta}$ CHE CONDIZIONATAMENTE ALLE X SONO LINEARI, CORRETTI $\hat{\beta}_{OLS}$ È IL PIÙ EFFICIENTE (CON VARIANZA PIÙ PICCOLA)

CIOÈ $\forall$ STIMATORE $\underline{\tilde{\beta}}$ DI $\underline{\beta}$ t.c.

MATRICE DETERMINISTICA o più comunque FUNZIONE di X, FISSATE COND. A X COSTANTE.

$E[\underline{\tilde{\beta}} \mid X] = \underline{\beta}$ e LINEARE (NON AFFINE) : $\underline{\tilde{\beta}} = D\underline{y}$

⇒ $\underline{\hat{\beta}}_{OLS}$ È IL PIÙ EFFICIENTE $\hat{\beta}_{OLS} = (X^T X)^{-1} X^T \underline{y}$

Scanned by CamScanner

(DIM) • $\hat{\beta}_{OLS}$ è CORRETTO $^{COND AX}$ (Vedo Dim precedente): $E[\hat{\beta}_{OLS}|X]=\beta$
• $\hat{\beta}_{OLS}$ è LINEARE CONDIZIONATAMENTE ALLE X
in quanto $\hat{\beta}_{OLS} = \underbrace{[(X^T X)^{-1} X^T]}^{\text{COND AX NOTE}} \underline{Y} = D\,Y$

$\Rightarrow$ in generale $\hat{\beta}$ STIMATORE CORRETTO E LIN COND AX (DI β)
ha varianza $\geq \hat{\beta}_{OLS}$ (NO DIM)

OSS Th G-M parla di efficienza RELATIVA cioè (SOTTO HP)
è IL + EFFICIENTE NELLA CLASSE DEI SOLI LINEARI E CORRETTI.

$\Rightarrow VAR(\hat{\beta}_{OLS}|X) =$ SOTTO HP (G-M / OLS...) (5) OMOSCHEDASTICITÀ $\Rightarrow VAR(\underline{u}|X) = \sigma^2_u I$
(X CASA)

26/10/15

• TESTARE L'HP DEL MODELLO

▲ CHI-QUADRATO CON M G.d.L. $= \chi^2_m$ è UNA TRASFORMAZIONE DI M V.A. $Z_1, \dots, Z_m \overset{IID}{\sim} N(0,1)$ (N.B. $Z_i \perp\!\!\!\perp Z_j$ $i \neq j$)

$\chi^2(m) = \sum_{i=1}^m Z_i^2$ (ES) $Z_1^2 = \chi^2(1)$

OSS $\underline{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_m \end{bmatrix} \sim N^{(m)}(\underline{0}, I_m)$ $\Rightarrow \chi^2_m = \sum_{i=1}^m Z_i^2 = \underline{Z}_i^T \underline{Z}_i$
(NOTAZIONE VETTORIALE)

■ Th (HP) • $\underset{m \times 1}{\underline{X}} \sim N^{(m)}(\underline{\mu}; \Sigma)$

• Σ SIMM. DEF. POSITIVA ($\Rightarrow$ RANK$(\Sigma) = m \Rightarrow \exists!\, \Sigma^{-1}$ SIMM. DEF. POS.) MATRICE PRECISIONE

$\Rightarrow \underset{1 \times 1}{W} = (\underline{X} - \underline{\mu})^T \Sigma^{-1} (\underline{X} - \underline{\mu})$ = V.C. [F.Q. ALEATORIA, trasf. X]

$\Rightarrow W \geq 0$ ($W = 0 \Leftrightarrow \underline{X} = \underline{\mu}$) $\Rightarrow$ W è F.Q. DEF. POSITIVA (perché Σ^{-1} DEF. POSITIVA), ma valori aleatori(?)

$W \sim \chi^2(m)$

OSS W è trasf. che generalizza nel caso multivariato l'operazione di: 1) STANDARDIZZAZIONE $\frac{X-\mu}{\sigma} \sim N(0,1)$
e poi 2) SOMMARE LE VARIABILI STANDARDIZZATE AL QUADRATO

OSS Passo 1) oltre a standardizzare rende $\perp\!\!\!\perp$ le V.C.

OSS $\Sigma = A^T A \Rightarrow A$ è MATRICE RADICE QUADRATA DI Σ
$\Rightarrow$ NOTAZIONE $A \doteq \Sigma^{1/2}$ (è solo notazione)

- F-TEST testa restrizioni lineari sui parametri del modello

MODELLO: $Y_i = X_i^T \cdot \underset{k \times 1}{\beta} + u_i$

se (1) $H_0: \beta_2 = 0 \Rightarrow$ TEST T (caso particolare TEST F)

se (2) $H_0: \beta_2 = -\beta_3 \Leftrightarrow \beta_2 + \beta_3 = 0 \Rightarrow$ TEST T (caso particolare TEST F)

se (3) $H_0: \beta_2 = \beta_3 = \dots = \beta_k = 0 \Rightarrow$ TEST F ($k-1$ vincoli lineari)

(OSS $\beta_2 \beta_3 = \frac{1}{\beta_4}$ è VINCOLO NON LINEARE (si può testare, ma + complicato))

$\Rightarrow$ VINCOLI LINEARI SUI PARAMETRI è H_0. (sistema di vincoli)

parametri nulla (R $\neq \beta$) vincoli intor. (numeri ignoti)

$H_0: \underset{\substack{m \times k \\ \text{matrice nota}}}{R} \cdot \underset{k \times 1}{\beta} = \underset{m \times 1}{r}$ (noti) con $m \leq k$

OSS R DI RANGO PIENO $= m \Rightarrow$ NON ABBIAMO VINCOLI LINEARMENTE DIPENDENTI TRA DI LORO. (vincoli inutili che non dicono nulla in +)

Esempio con $k=3$ $y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

$H_0: \beta_2 = \beta_3 = 0$ (tipico F-test per vedere se almeno una variabile ha effetto sig.)

$\Rightarrow H_0 = \begin{cases} \beta_2 = 0 \\ \beta_3 = 0 \end{cases}$

$\Rightarrow \underset{2 \times 3}{R} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = r$ con $R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Esempio (x casa) $H_0: \beta_2 + \beta_3 = 0$

$\Rightarrow \overset{R}{[0, 1, 1]} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = 0 = r$

$\Rightarrow$ STATISTICA TEST CON m VINCOLI

STAT TEST $F_0 \underset{\text{sotto } H_0}{\overset{\text{asintoticamente}}{\sim}} F_{m, +\infty}$

- CALCOLIAMO STAT TEST F_0

(oss sfrutteremo asintotica normalità: $\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta}) \xrightarrow[n\to+\infty]{d} N(0, \Sigma_{\sqrt{n}(\hat{\beta}-\beta)})$
e $\hat{\Sigma}_{\sqrt{n}(\hat{\beta}-\underline{\beta})}$ di White è stimatore consistente
(cioè $\xrightarrow[n\to+\infty]{P}$) a $\Sigma_{\sqrt{n}(\hat{\underline{\beta}}-\underline{\beta})}$
$\Rightarrow$ Approx con $n>>0$ $\hat{\underline{\beta}} \approx N(\underline{\beta}, \Sigma_{\hat{\underline{\beta}}})$ $\Rightarrow$ (perché trasf. lin. di $N(\cdot)$ è $N(\cdot)$)
$\Rightarrow$ per $n>>0$ $\underset{m\times 1}{\underline{\delta}} \doteq \underset{m\times k}{R}\,\underset{k\times 1}{\hat{\underline{\beta}}} \approx N(R\underline{\beta}, R\Sigma_{\hat{\underline{\beta}}}R^T)$
($\Rightarrow$ per ottenere $\chi^2_{(m)}$)
$\Rightarrow \underset{1\times 1}{W} \doteq (R\hat{\underline{\beta}}-R\beta)^T(R\Sigma_{\hat{\beta}}R^T)^{-1}(R\hat{\underline{\beta}}-R\underline{\beta}) \underset{n>>0}{\approx} \chi^2_{(m)}$ (Pr. Th)
$\Rightarrow$ v.c. scalare $W \xrightarrow{d} \chi^2(m)$

[def/oss F di Snedecor-Fisher $\doteq F_{m,n} \doteq \dfrac{\chi^2(m)/m}{\chi^2(n)/n}$ rapporto tra 2 χ^2/g.d.l. con $\chi^2_{(m)} \perp\!\!\!\perp \chi^2_{(n)}$]
con $\chi^2(m) \perp\!\!\!\perp \chi^2(n)$

[oss se $n\to+\infty$ $\dfrac{\chi^2(n)}{n} \xrightarrow{P} 1$ $\Rightarrow$ rimane solo numeratore

$\Rightarrow$ def $F_{m,+\infty} \doteq \dfrac{\chi^2(m)}{m}$

$\Rightarrow \dfrac{W}{m} \xrightarrow{d} F_{m,+\infty}$ $\Rightarrow$ $F_0 \doteq$ stat test $\doteq \dfrac{W}{m}$ e $F_{m,+\infty}\cdot m = \chi^2_{(m)}$

(oss in alcuni pacchetti statistici c'è test F in altri χ^2)

$$F = \frac{1}{m}(R\hat{\underline{\beta}}-R\underline{\beta})^T(R\Sigma_{\hat{\beta}}R^T)^{-1}(R\hat{\underline{\beta}}-R\underline{\beta})$$

$$\Rightarrow F \underset{n>>0}{\approx} F_{m,+\infty}$$

ma non conosco $R\underline{\beta}$ (ma sotto H_0: $R\underline{\beta}=r$) e
$\Sigma_{\hat{\beta}}$, ma poiché approx è asintotica $\Rightarrow$ posso usare stimatore (asintotico) consistente, es. White $\hat{\Sigma}_{\sqrt{n}(\hat{\beta}-\underline{\beta})} \xrightarrow{P} \Sigma_{\sqrt{n}(\hat{\underline{\beta}}-\beta)}$.

$\Rightarrow (R\hat{\underline{\beta}}-R\underline{\beta})^T(R\underbrace{\hat{\Sigma}_{\hat{\beta}}}_{\text{stimatore consistente}}R^T)^{-1}(R\hat{\underline{\beta}}-R\underline{\beta}) \doteq \hat{W}$

$\Rightarrow \hat{W} \underset{n>>0}{\approx} \chi^2_{(m)}$

Scanned by CamScanner

$F_0 \triangleq \frac{1}{m}(R\hat{\beta}-r)^T (R\hat{\Sigma}_{\beta}R^T)^{-1}(R\hat{\beta}-r)$ = STAT TEST SUB H_0

$\Rightarrow \underset{\text{White}}{F_0} \overset{\text{sub } H_0}{\underset{n>>0}{\approx}} F_{m,+\infty}$ ∼ $m \cdot F_0 \overset{\text{sub } H_0}{\underset{n>>0}{\approx}} \chi^2(m)$

OSS In GRETL (o altri software) STAT TEST $= F_0 \underset{H_0}{\sim} F_{m,(n-k)}$ cioè si calcola p-value test non con distr. asintotica, ma con $(n-k)$ = g.d.l. modello con $n>>0$. ($\Rightarrow$ se n molto grande $\Rightarrow$ poca differenza)

N.B. DAL PUNTO DI VISTA TEORICO È SBAGLIATO, MA È PIÙ PRUDENTE CONSIDERARE $F_{m,(n-k)}$ $\Rightarrow$ SERVE MAGGIORE EVIDENZA CONTRO H_0 PER RIFIUTARLA.

OSS Se $H_0: \beta_2 = \beta_{2,0}$ valore fissato sub H_0 (es. $\beta_2 = 0$)

$\Rightarrow$ T-TEST

$R = [0, 1, 0, \dots]$ $\Rightarrow R\hat{\beta} = \hat{\beta}_2 \underset{n>>0}{\approx} N(\beta_2, \sigma^2_{\beta_2})$

$\Rightarrow \frac{\hat{\beta}_2-\beta_2}{\sigma^2_{\beta_2}} \underset{n>>0}{\approx} N(0,1)$ ma $\sigma^2_{\beta_2}$ è ignota $\Rightarrow$ la stimo con WHITE.

$\Rightarrow$ STAT. TEST T $\Rightarrow t = \frac{\hat{\beta}_2-\beta_2}{\hat{\sigma}_{\beta_2 \, \text{White}}}$ (perché $\hat{\sigma}_{\beta_2}$ è stimatore consistente)

$\Rightarrow t \overset{H_0}{\underset{n>>0}{\approx}} N(0,1)$ (Versione asintotica del test)

OSS T-STUDENT CON m G.d.L. $\triangleq t_m \triangleq \frac{Z}{\sqrt{\chi^2(m)/m}}$ con $Z \perp\!\!\!\perp \chi^2(m)$

$\Rightarrow t_m \xrightarrow{d} Z \sim N(0,1)$

OSS $t_m^2 = F_{1,m}$

$\Rightarrow$ VALORE STAT t SUB H_0: $t_0 = \frac{\hat{\beta}_2-\beta_{2,0}}{\hat{\sigma}_{\hat{\beta}_2}} \overset{H_0}{\underset{n>>0}{\approx}} N(0,1)$

$\Rightarrow$ Nei software (es. GRETL) IN MANIERA PIÙ CONSERVATIVA si calcolano p-value e soglia da $T_{(n-k)}$

OSS Test con (t_{n-k}) se faccio quadrato $\Rightarrow F_{1,(n-k)}$
e $N(0,1)$ al quadrato $\Rightarrow \chi^2_1$
$\Rightarrow$ ugualmente faccio test F.

Leggo l'articolo di un modello di Solow

(A.G. 27/10/15)

PROPONGONO ESTENSIONE DEL MODELLO DI SOLOW PER SPIEGARE CRESCITA PAESI (INTRODUCONO UNA NUOVA GRANDEZZA: IL CAPITALE UMANO oltre al capitale fisico)

• MODELLO di SOLOW ('60)

È un modello di crescita (...)

Si parte da f. produzione:

$$Y(t) = F(\underbrace{K(t)}_{\text{CAPITALE FISICO}}, \underbrace{L(t)}_{\text{LAVORO}} \cdot \underbrace{A(t)}_{\text{STATO TECNOLOGIA}})$$

OSS $L(t) \cdot A(t)$ = UNITÀ DI LAVORO EFFETTIVE (anche a parità di lavoratori se aumenta la tecnologia è come averne di più.)

OSS F ha forma COBB-DOUGLAS:

$$Y_t = K(t)^{\alpha} \, (L(t) \cdot A(t))^{1-\alpha} \qquad 0<\alpha<1$$

È modello DI CRESCITA ESOGENA, perché la tecnologia cresce in modo esogeno (non è spiegato perché) al tasso costante $g = \frac{\dot{A}(t)}{A(t)}$ con $\dot{A}(t)$

$A(t) = A(0) \cdot e^{gt}$ $\Rightarrow$ g è tasso di crescita di A

$L(t) = L(0) \cdot e^{nt}$ $\Rightarrow$ n è tasso di crescita lavoratori (= popolazione)

POPOLAZIONE = LAVORATORI (tutti i lavoratori sono occupati)

OSS il capitale fisico $K(t)$ è VARIABILE ENDOGENA è spiegato come varia dal modello,

(HP) RISPARMIO = INVESTIMENTO (tutto il risparmio è investito) e il risparmio è frazione (= s) dell'OUTPUT ma nel tempo il capitale si deprezza al tasso δ COSTANTE

Scanned by CamScanner

IN OGNI PERIODO:

$$\dot{K}(t) = s \cdot Y(t) - \delta \cdot K(t)$$

varia nel tempo MODELLO DINAMICO, ma DETERMINISTICO.

OSS qui $L(t)$ e $A(t)$ continuano a crescere a tasso $\Rightarrow$ $Y(t)$?

ma noi vogliamo che dopo un po' non cresca + il tasso,

pongo:
$$\frac{Y(t)}{A(t)L(t)} \doteq y(t) \qquad k(t) \doteq \frac{K(t)}{A(t) \cdot L(t)}$$
unità effettive

ho trovato output per unità effettive.

e capitale per unità effettive.

$\Rightarrow$ non cresce sempre, ma si dim che $\exists !$ STABILE PUNTO DI EQUILIBRIO: $y(t) = y^*$ punto di STEADY STATE.

$\Rightarrow$ Risolta eq. differenziale $\Rightarrow$ soluzione:

$$K^* = \left(\frac{s}{n+g+\delta}\right)^{\frac{1}{1-\alpha}}$$

$$\Rightarrow y(t) = k(t)^{\alpha}$$
$$y^* = \left(\frac{s}{n+g+\delta}\right)^{\frac{\alpha}{1-\alpha}}$$

$$\frac{Y(t)}{L(t)} = A(t) \cdot y(t)$$ in termini di lavoratori pro capite

$\Rightarrow$ L'OUTPUT PRO CAPITE (per lavoratore) CRESCE IN STEADY STATE DINAMICO AL TASSO g. (Con Modello Solow)

$\Rightarrow$ IL STEADY-STATE DINAMICO:

$$\left(\frac{Y(t)}{L(t)}\right)^* = A(t) \cdot y^*.$$
cresce nel tempo = tasso di crescita $\frac{K(t)}{L(t)}$

= CRESCITA OUTPUT PROCAPITE (per lavoro) È LEGATO SOLO ALLA CRESCITA DELLA TECNOLOGIA (per questo è modello di crescita ESOGENA non spiega la crescita)

LABORATORIO

GRETL Stima Modello di Solow 27/11/15

DATASET MRW

GDP = GDP per adulto (per [illegible])

Obs GDP è a valori [illegible], noi vogliamo i valori reali, annuali.

⇒ altri confronti si utilizzano dati PPP a parità potere d'acquisto

Obs n_i = tasso di crescita pop., [illegible] confr.
un solo valore trend, [illegible] ⇒ n_i è $\bar{n}_i$ [illegible]

[illegible] ⇒ $\frac{INV}{PIL} = s$ ⇒ facendo sempre media

Obs C'è variabilità [illegible] GDP [illegible] per [illegible]

Obs le variabili sono già misurate in termini %, ma noi [illegible] ⇒ le trasformiamo per avere la stessa scala, per le altre variabili non trasf., perché con log().

• NONOIL (significa non oil), 121 paesi

Obs dall'appendice dell'articolo trova tutti i dati.

Loro considerano 3 campioni ≠ 1) OECD (22 paesi)

2) Campione intermedio (75 paesi)

3) Non OIL (98)

⇒ [illegible] 3 gruppi, noi usiamo tutti i 121 paesi. (3) Nonoil: paesi che [illegible] di petrolio ⇒ è quello che spiega la crescita non altro ⇒ li eliminano.

(2) Tolgo paesi (troppo poveri) in cui dati non affidabili) e OIL e paesi troppo piccoli

Obs trasform.

[illegible]

[illegible]

Es. $t=25$ ⇒ Modello

29/11/…

$$\ln\left(\left(\frac{Y}{L}\right)_i\right) = \beta_0 + \beta_1 \ln(s_i) + \beta_2 \ln(n_i + g + \delta) + u_i$$

con $u_i \sim IID(0, \sigma^2_u)$

oss $\beta_0, \beta_1, \beta_2$ sono parametri del modello da stimare, non strutturali come nel modello $\beta_1 = \frac{\alpha}{1-\alpha}$ $\beta_2 = -\beta_1$

oss F. PRODUZIONE È COBB DOUGLAS PER TUTTI I PAESI

$Y(t) = K(t)^{\alpha} (A(t) L(t))^{1-\alpha}$ (Cobb Douglas)

in cui α rappresenta ELASTICITÀ (parziale) DELL'OUTPUT RISPETTO AL CAPITALE FISICO (parziale perché ceteris paribus)

⇒ $\ln[Y(t)] = \alpha \cdot \ln[K(t)] + (1-\alpha) \ln[A(t) \cdot L(t)]$

⇒ Interpretiamo α = ELASTICITÀ OUTPUT rispetto a CAPITALE

α può anche essere interpretato come frazione dell'output che remunera il capitale

(⇒ su base di dati economici stimiamo $\alpha \simeq 1/3$
⇒ $\beta_1 \simeq 1/2$ ⇒ $\beta_2 \simeq -1/2$ ⇒ vediamo a posteriori se stime hanno senso con questi dati)

oss Nel modello β_1, β_2 sono liberi di variare (senza vincolo $\beta_1 = -\beta_2$) infatti stimiamo β e non parametro strutturale α.

▲ α = parametro strutturale

▲ β_1, β_2 = parametri della forma ridotta.

(NON INCLUDIAMO VINCOLO $\beta_1 = -\beta_2$ della teoria economica
ma sarà $H_0: \beta_1 = -\beta_2$ che testeremo
come potremo poi verificare che $H_0: \beta_1 = 1/2$
Test per capire se modello teorico è in accordo con dati empirici)

Modello in GRETL se seleziono standard error Robusti
⇒ (cross-sect data HC0 = formula di WHITE ROBUSTA

OBS se penso (come adesso) che errori OMOSCHEDASTICI
⇒ non clicco Robust standard errors ⇒ (stima stessa
ERRORI SENZA ETEROSCHEDASTICITÀ altrimenti test è
sbagliato.

⇒ VEDIAMO TEST DI DEFAULT:

1) T-TEST con $H_0: COEF_j = 0$ $j = 0, 1, 2$ (rifiuto ... [illegible])

2) F TEST $H_0: \beta_1 = \beta_2 = 0$ (nessuna capacità predittiva dei regressori) $F(2, \frac{n-3}{n})$

F-VALUE alto ⇒ P-VALUE basso $\leq \alpha$ ⇒ rifiuto [illegible]
⇒ Modello più meglio del nullo

OBS 1) e 2) ci dicono che coefficienti sono rilevanti

OBS RIFACCIO LE TEST CON ERRORI ROBUSTI e
RISULTATI ANALOGHI.

⇒ TEST WHITE H_0: OMOSCHEDASTICITÀ degli errori

5

Scanned by CamScanner

21/12/16
5/11/16

Modello: $Y_t = \beta_0 + \beta_1 \cdot DIFF1 + \varepsilon_t$

$\beta_0, \beta_1, \beta_2$

$\beta_1 = \frac{\alpha}{1-\alpha}$

$\beta_2 = -\frac{\gamma}{1-\alpha}$

$\beta_1 = \frac{\alpha}{1-\alpha} = f(\alpha)$

$\alpha = \frac{\beta_1}{1+\beta_1} = g(\beta_1)$

$\hat{\alpha} = g(\hat{\beta}_1)$ e $\hat{\alpha}$ è consistente e asintoticamente normale perché $\hat{\beta}_1$ è consistente e asintoticamente normale

oss Avrei potuto analogamente stimare $\hat{\alpha}$ da $\hat{\beta}_2$ ma avrei $\neq$ risultati → α e stime $\neq$ dai parametri della forma ridotta ⇒

⇒ PROBLEMI: LA FORMA STRUTTURALE NON È IDENTIFICATA
non è possibile trovare in modo univoco α.

N.B. Non posso considerarne solo uno dei 2 a caso ⇒ SBAGLIATO!

⇒ Sfrutto info Teoria Economica.

$Y_t = \beta_0 + \beta_1 \cdot DIFF1 + \varepsilon_1$

⇒ IL NUOVO MODELLO HA ESATTA IDENTIFICAZIONE

oss $\alpha = g(\beta_1)$ $\hat{\beta}_1 \xrightarrow{P} \beta_1$ $\hat{\beta}_1 \overset{a}{\sim} N(\beta_1, \sigma^2_{\beta_1})$

$\hat{\alpha} = g(\hat{\beta}_1)$ è STIMATORE CONSISTENTE e

$\hat{\alpha} \approx N(g(\beta_1), \hat{\sigma}^2_{\alpha})$

Perché

• METODO DELTA

- METODO DELTA

$$\hat{\sigma}_{\hat{\alpha}}^2 = \hat{\sigma}^2_{\hat{\beta}_1} \cdot \left[g'(\beta_1) \Big|_{\beta_1 = \hat{\beta}_1} \right]^2$$

⇒ ... conti ...

con $g'(\beta_1) = \dfrac{1}{(1+\beta_1)^2}$

$$\Rightarrow \hat{\sigma}^2_{\hat{\alpha}} = \hat{\sigma}^2_{\hat{\beta}_1} \cdot \left[\frac{1}{(1+\hat{\beta}_1)} \right]^2$$

$$\Rightarrow s.e.(\hat{\alpha}) = s.e.(\hat{\beta}_1) \cdot \left| \frac{1}{(1+\hat{\beta}_1)^2} \right|$$

⇒ con: $\hat{\alpha} \simeq 0{,}59$ $\hat{\beta}_1 \simeq 1{,}43$ $se(\hat{\alpha}) \simeq 0{,}02$
(come da tabella paper)

⇒ SI PUÒ FARE TEST con $H_0: \alpha = \frac{1}{3} = 0{,}33$

⇒ t-test $t_0 = \dfrac{\hat{\alpha} - 0{,}33}{se(\hat{\alpha})} = \ldots \Rightarrow$ RIFIUTO H_0

OSS + VELOCEM

OSS Alternativa al METODO DELTA nel nostro caso

$$Y_i = \beta_0 + \left(\frac{\alpha}{1-\alpha}\right) \cdot DIFF_{1i} + \varepsilon_i$$ MODELLO NON LINEARE NEI PARAMETRI

→ "Metodo dei MINIMI QUADRATI NON LINEARI"

File gretl

OSS Questi TEST vanno fatti dopo aver accettato il modello con i TEST CORRETTA SPECIFICAZIONE DEL MODELLO.

RIPRENDO MODELLO DI SOLOW [illegible] DA [illegible] 3/12/15

$\Rightarrow \frac{\alpha}{1-\alpha-\beta}$ [illegible] 1 [illegible] $= 1 \neq \frac{1}{2}$

$\Rightarrow$ [illegible] modello di Solow [illegible]

$\Rightarrow$ MODELLO IN CUI NON CONTROLLIAMO PER UN REGRESSORE RILEVANTE e MODELLO ~~[illegible]~~ REGRESSORE RILEVANTE.

$\Rightarrow$ LE STIME ~~[illegible]~~ SARANNO BIAS SE IL REGRESSORE RILEVANTE CORRELA CON IL REGR. $\Rightarrow$ SE NO [illegible] OMESSO

$\Rightarrow$ OLS NON CONSISTENTE E DISTORTO.

OSS VEDIAMO CHE È (correlato) CORRETTO NEL CAPITALE E RILEVANTE, [illegible] rilevante $\Rightarrow$ OMISSIONE REGRESSORE ($\ln k_i$) $\Rightarrow$ DISTORSIONE E INCONSISTENZA DELLE STIME.

$\Rightarrow$ MRW dicono che per questo modello di Solow è [illegible]

$\Rightarrow$ MRW dicono che il loro modello è meglio.

OSS se aggiungiamo OECD tra variabili esplicative

$\Rightarrow$ vediamo che [illegible] non è + sign.

MODELLO DI SOLOW $Y_i = \beta_0 + \beta_1 \ln(s_i) + \beta_2 \ln(n_i + g + \delta) + \varepsilon_i$

$\ln(A_{i,0}) = a_i + \varepsilon_i$

OSS se al posto di per molti intercette $\beta_{0i} \neq$ tipo paese

$\Rightarrow$ non posso stimare i parametri perché [illegible] modello [illegible]. $\Rightarrow$ SI POTRÀ STIMARE [illegible] per DATI PANEL) $\Rightarrow$ su dati CROSS-SECTION possiamo stimare [illegible] ottima modello con 2 intercette

una per paesi OECD e una per paesi non OECD.

$\Rightarrow$ [illegible] aggiungendo una dummy OECD, o togliere la cost e mettere le due dummy OECD e non OECD.

Scanned by CamScanner

oss Le determinanti di y^* = livello dell'output per unità effettiva di lavoro: dipende da s, n, g, δ

$$\ln\left(\frac{Y}{L}\right)_t = \ln(A(t)) + \ln(y^*) = \underbrace{\ln(A_0) + g \cdot t}_{\text{TECNO LINEARE}} + \underbrace{\frac{\alpha}{1-\alpha}}_{>0}\ln(s) - \underbrace{\frac{\alpha}{1-\alpha}}_{>0}\ln(n+g+\delta)$$

effetto neg su output procapite

effetto positivo su output procapite

abbiamo relazione lineare sui log

oss Questo è un modello deterministico, ma noi vogliamo poter stimare ⇒ introduciamo un errore.
⇒ IDEA STIMARE α. sfruttando la variabilità tra paesi.
Noi poniamo relazione cross-section per l'economia i-esima valore $\forall$ paese.

$$\Rightarrow \ln\left(\frac{Y}{L}\right)_{ti} = \ln(A_0)_i + gt + \frac{\alpha}{1-\alpha}\ln(s_i) - \frac{\alpha}{1-\alpha}\ln(n_i+g+\delta)$$

(con $g+\delta = 0{,}05$ per HP (vedere che è costante su tutti i paesi))

HP $\ln(A_{0,i})$ livello tecnologico può essere $\neq$ tra paesi e inoltre paesi $\neq$ hanno risorse $\neq$, ... ⇒ non è uguale per tutti ⇒ fisso a = valore medio + u_i = errore IID rispetto a paesi i. $u_i \sim iid(0, \sigma_i^2)$

$$\Rightarrow \text{HP } \ln(A_{0,i}) = a + u_i$$

$$\Rightarrow \text{Modello } \ln\left(\frac{Y}{L}\right)_{ti} = a + gt + \frac{\alpha}{1-\alpha}\ln(s_i) - \frac{\alpha}{1-\alpha}\ln(n_i+g+\delta) + u_i$$

oss loro modello su dati annuali '60-'85, noi stimiamo nell'ultimo disponibile t = '85. (fisso t)
se $t_0 = 0$ (nel 1960) ⇒ $t = 25$ (nel 1985)

$$\Rightarrow \text{Modello: } \ln\left(\frac{Y_{1985}}{L_{1985}}\right)_i = \beta_0 + \beta_1\ln(s_i) + \beta_2\ln(n_i+g+\delta) + u_i$$

dove $\beta_0 = a + gt$ = costante $\forall i$ = paese.

$\beta_1 = \frac{\alpha}{1-\alpha}$ $\beta_2 = -\beta_1$

Scanned by CamScanner

⇒ obiettivo è stimare dai dati CROSS-SECTION i β.
Per poter usare OLS serve ulteriore (H):
(H) $COV(\ln(S_i), u_i) = 0 \wedge COV(u_i, \ln(M_i+g+\delta)) = 0$

⇒ OBIETTIVO STUDIO SU ARTIC. [illegible] MODELLO.

• CAP 4 VERBEEK TEST DI WHITE 31/10/16

Sotto (H) Modello $E[u_i|\underline{X}_i] = E[u_i|X] = 0$

⇒ $VAR[u_i|\underline{X}_i] = E[u_i^2|\underline{X}_i] = h(\underline{X}_i) \geq 0$

OSS Breusch-Pagan dimostrano che $h(\cdot)$ può essere f. qualunque sotto vincolo $h(\underline{X}_i) \geq 0$.

Ma Th IN TERMINI ASINTOTICI: il modello va bene comunque se usiamo per $h(\cdot)$ una forma quadratica

⇒ TEST DI WHITE:

⇒ $VAR[u_i|\underline{X}_i] = E[u_i^2|\underline{X}_i] = h(\underline{X}_i) \overset{(H)}{=}$

$\overset{(H)}{=} \delta_0 + \delta_1 X_{1i} + \delta_2 X_{2i} + \delta_3 X_{1i}^2 + \delta_4 X_{2i}^2 + \delta_5 (X_{1i} \cdot X_{2i})$

(funzione abbastanza completa che va bene per test $\forall h(\cdot)$) in cui c'è eteroschedasticità

H0: OMOSCHEDASTICITÀ (corretta specificazione modello classico)

⇒ $H_0: h(\underline{X}_i) = \delta_0 > 0$ (con $\delta_0 = \sigma_u^2 \ \forall u$)

⇒ $H_0: \delta_1 = \delta_2 = \delta_3 = \dots = \delta_5 = 0$ (5 RESTRIZIONI)

OSS $u_i^2 = \underbrace{E[u_i^2|\underline{X}_i]}_{h(X_i) = F. REG.} + v_i$, t.c. $E[v_i|\underline{X}_i] = 0$

solita decomposizione modello con F. REG.

OSS Sotto H0: $\hat{u}_i^2 \xrightarrow{P} u_i^2$ (vale in senso asintotico sotto H0, usare residui o errori è uguale) ⇒ ALLORA USO

• REGRESSIONE AUSILIARIA (NON HA SIGNIFICATO ECONOMICO MA UTILE PER TEST)

OLS $\hat{u}_i^2 = \delta_0 + \dots + \delta_5 (X_{1i} \cdot X_{2i}) + v_i$

utile per test perché $\underbrace{nR^2}_{\text{STAT-TEST ASINTOTICA DEL TEST DI WHITE}} \overset{H_0}{\underset{n>30}{\sim}} \underbrace{\chi^2_{(5)}}_{\text{N. VINCOLI DI H0}}$

OSS GRETL TEST WHITE $R^2 = 0{,}058239$

$\Rightarrow \pi^{oss} \cdot R^2 \Rightarrow$ NON RIF. H_0 (p value alto) so non abbiamo ETEROSCHEDASTICITÀ.

OSS Problema Test: se tante variabili $\Rightarrow$ il numero di parametri e vincoli di H_0 ↑ $\Rightarrow$ (Tolgo prodotti misti). se no Breusch test χ^2

OSS Se pensiamo che conti solo una variabile $\Rightarrow$ restringo H_1 $\Rightarrow$ Test + potente se è vera la mia idea, se no può essere fuorviante.

... standard errori classici (... usa di default quelli robusti)

● 2 STEP dopo aver "accettato" OMOSCHEDASTICITÀ

$\Rightarrow$ Modello $\underline{Y} = X\underline{\beta} + \underline{u}$

• $E[\underline{u}\cdot\underline{u}^T|X] = Var[\underline{u}|X] = \sigma_u^2 I_n$ (omoschedasticità)

$\Rightarrow \hat{\beta}$ è CORRETTO, CONSIST., ASINT. $N(\cdot)$ + BLUE (GAUSS-MARKOV solo per OMOSCHEDASTICITÀ)

OSS Se HP 6 $(\underline{u}|X) \sim N(\underline{0}, \sigma_u^2 I_n)$ (HP IID + NORM.)

$\Rightarrow u_i \overset{IID}{\sim} N(0, \sigma_u^2)$

OSS Se aggiungiamo HP 5 OMOSCHEDASTICITÀ

$\Rightarrow$ "Modello Classico" + HP 6 "Modello Classico Normale"

Th SUB HP (1-6) $\Rightarrow \hat{\underline{\beta}}_{OLS} = \hat{\underline{\beta}}_{ML}$ Stimatore OLS coincide con stimatore di MAX VEROSIMIGLIANZA che è il nostro stimatore ideale (+ efficiente fra tutti gli stimatori (LIN e NON))

$\Rightarrow (\underline{Y}|X) \sim N(X\underline{\beta}, \sigma^2 I_n)$

$(\hat{\underline{\beta}}_{OLS}|X) \sim N(\underline{\beta}, \sigma_u^2 (X^T X)^{-1})$

Scanned by CamScanner

⇒ anche in campioni piccoli sotto (6) conosciamo distr. prob. stimatore ⇒ ripercussioni su Stat Test:

⇒ $H_0: \beta_j = \beta_{j,0}$ ⇒ $t_0 = \frac{\hat{\beta}_j - \beta_{j,0}}{se(\hat{\beta}_j)} \overset{H_0}{\sim} t_{(n-k)}$ anche in piccoli

(anche se n piccolo è stat esatta sotto (6))

$F_{0,\text{classica}} \overset{H_0}{\sim} F_{m,(n-k)}$ (anche con n piccolo)

OSS: Altro motivo (oltre ad essere conservativa) per usare distr. non asintotiche. (se senza normalità ⇒ risultati valgono solo per campioni grandi)

⇒ ■ TEST NORMALITÀ DEI RESIDUI OLS

(la verifica HP (6))

$H_0: u_i \sim N(\cdot)$

(1) TEST JARQUE-BERA TEST (J-B)

è un test generale, semplice.

OSS Per la Normale INDICE DI SIMMETRIA $= S = 0$

INDICE DI KURTOSI $= (K-3) = 0$

con $S \triangleq \frac{E[(X - E(X))^3]}{\sigma_X^3}$

$K \triangleq \frac{E[(X - E(X))^4]}{\sigma^4}$

⇒ STIMATORI CAMPIONARI: $\hat{S} = \frac{\frac{1}{n}\sum(X_i - \bar{X})^3}{\hat{\sigma}_X^3} \xrightarrow[P]{\text{IID } H_0 N(\cdot)} 0$

$\hat{K} = \frac{\frac{1}{n}\sum(X_i - \bar{X})^4}{\hat{\sigma}^4} \xrightarrow[P]{\text{IID } H_0 N(\cdot)} 3$

⇒ $(\hat{K} - 3) = \frac{\frac{1}{n}\sum \hat{u}_i^4}{\hat{\sigma}_{\hat{u}}^4} - 3 \xrightarrow[\text{sotto } H_0]{P} 0$

e $\hat{S} = \frac{\frac{1}{n}\sum \hat{u}_i^3}{\hat{\sigma}_{\hat{u}}^3} \xrightarrow[\text{sotto } H_0]{P} 0$

OSS $\Rightarrow$ $\left[\begin{array}{l} \hat{S} \underset{n>>0}{\overset{H_0}{\sim}} N\left(0, \frac{6}{n}\right) \qquad \hat{K} \underset{n>>0}{\overset{H_0}{\sim}} N\left(3, \frac{24}{n}\right) \\ e \ \hat{S} \perp\!\!\!\perp \hat{K} \text{ ASINTOTICAMENTE} \end{array}\right.$

$\Rightarrow$ Costruisco STAT TEST ASINTOTICA

$$\left(\frac{\hat{S}-0}{\sqrt{\frac{6}{n}}}\right)^2 + \left(\frac{(\hat{K}-3)}{\sqrt{\frac{24}{n}}}\right)^2 \underset{n>>0}{\overset{H_0}{\sim}} \chi^2(1) \underset{\perp\!\!\!\perp}{+} \chi^2(1) = \chi^2(2)$$

$\Rightarrow$ STAT TEST $JB_0 = n\left(\frac{\hat{S}^2}{6} + \frac{(\hat{K}-3)^2}{24}\right) \underset{n>>0}{\overset{H_0}{\sim}} \chi^2(2)$

STAT TEST per testare $H_0: u \overset{iid}{\sim} N(0, \sigma^2_u)$

- GRETL TEST NORMALITY OF RESIDUALS:
  lui dà JB_0 e PVALUE $\Rightarrow$ NON RIFIUTIAMO $H_0: \sim N(\cdot)$

TEST VINCOLI SU PARAMETRI

$H_0: \beta_1 = -\beta_2$ (Predetto da teoria modello solare)

β_0	β_1	β_2
b[1]	b[2]	b[3]

$\Rightarrow$ H_0: b[2]+b[3]=0. $\Rightarrow$ $R = (0,1,1)$ $r = 0$
in notazione matriciale.

GRETL $\rightarrow$ TEST $\rightarrow$ LINEAR RESTRICTIONS $\rightarrow$ b[2]+b[3]=0

F_0 = STAT TEST $\overset{H_0}{\sim} F_{1,72}$

P VALUE $> \alpha$ $\Rightarrow$ NON RIF. H_0: VINCOLO

$\Rightarrow$ SEMBRA ESSERE IN LINEA CON TEORIA.

$\Rightarrow$ Se vincolo è "accettato" dai dati $\Rightarrow$ posso usare OLS con vincolo $\Rightarrow$ Stimo un parametro in meno $\Rightarrow$ $\nearrow$ G.d.L $\Rightarrow$ $\nearrow$ GUADAGNO IN EFFICIENZA STIMATORI (STD. ERRORS)

OSS Se avessi messo il vincolo, quando non è accettato dai dati $\Rightarrow$ ERRORE! STD ERROR. $\searrow$, ma STIMATORE DISTORTO e NON CONSISTENTE ⚠

Scanned by CamScanner

⇒ MODELLO VINCOLATO:

$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$ con $\overset{\text{vincolo}}{\beta_2 = -\beta_1}$

⇒ Modello $Y_i = \beta_0 + \beta_1 (X_{1i} - X_{2i}) + u_i$

STIMO IN GRETL MODELLO VINCOLATO.

5/12/15

• OSS PRIMA DI FARE TEST SU SIGNIFICATIVITÀ DEI PARAMETRI BISOGNA FARE TEST SU (FE) MODELLO.

In particolare ne mancano 2:

① RESET TEST DI RAMSEY.

TEST PER LA LINEARITÀ DEL MODELLO DI REGRESSIONE.

② CHOW TEST: TEST PER LA STABILITÀ DEI PARAMETRI cioè i parametri sono uguali H_0 per tutti gli individui vs. H_1: eterogeneità nei parametri (NASCE IN AMBITO S.S. per stare parametri)

① RESET TEST

test inserisce nei regressori Y_{HAT}^2, Y_{HAT}^3, ...

e vede se modello è migliore

LO FA IN MODO PARSIMONIOSO.

modello stimato è $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$ (= H_0)

⇒ e se penso che forse è meglio regressione quadratica dovrebbe fare modello. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + [\beta_3 X_1^2 + \beta_4 X_2^2 + \beta_5 X_1 X_2] + v$ (= H_1)

⇒ TEST H_0: $\beta_3 = \beta_4 = \beta_5 = 0$ è test LINEARITÀ.

OSS Test di RAMSEY fa + o - la stessa cosa, ma per evitare l'esplodere del n° di parametri ⇒ ↓ gdl ⇒ ↓ potenza test ⇒

⇒ aggiungo termini aggiuntivi [ ] con $\hat{Y}^2$

Scanned by CamScanner

dove $\hat{Y}^2$ è valore predetto del Modello LIN H_0 (al quadrato) ⇒ dove $\hat{Y}^2 = (\hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2)^2$
⇒ effettivamente avremmo quadrato di X_1 e X_2 e prodotto misto come volevamo, ma c'è differenza che i coef. non sono + liberi ⇒ è un modo parsimonioso per inserire termini non lineari.

⇒ Modello DI REGRESSIONE AUSILIARIA utile solo per calcolare stat test. $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \delta \hat{Y}^2 + \varepsilon$

⇒ $H_0: \delta = 0$

⇒ $nR^2 \underset{n>>0}{\overset{H_0}{\approx}} \chi^2(1)$ ∃ anche versione Test F per piccoli campioni.

(Non rifiuto ... ⇒ ... H_0)

⇒ Provo con il cubo

REG. AUSILIARIA $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \delta_1 \hat{Y}^2 + \delta_2 \hat{Y}^3 + U_2$

⇒ $H_0: \delta_1 = \delta_2 = 0$. ⇒ $nR^2 \underset{n>>0}{\overset{H_0}{\approx}} \chi^2(2)$

posso anche fare in GRETL test completo con tutte le varianti

② TEST CHOW

Nasce per S.S. divide dati in più campioni e vede se coef. = (con test.)

⇒ Es. Vedo se relazioni per Stessi POVERI e RICCHI è ≠

⇒ GRETL CHOW TEST con dummy RICCO

è come fare Modello e stime su campione RICCHI e poi stime con modello ≠ su campione POVERI

noi troviamo valori molto ≠.

• TEST DI CHOW fa unico modello su tutto introducendo dummies t.c. coef. (intercetta e slope) ≠ nei gruppi

OSS BENA DUMMY · X_5 TRA REGRESSORI ⇒ (LE REGRESSIONI OCCD; vediamo che coef ≠ (⊕ intercetta = const per entrambe, mentre è const − cambia OCCD per OCCD)

OSS POSSO ANCHE STIMARE 2 MODELLI ≠ SU OCCD e NON OCCD

TEST CHOW DI FATTO È F TEST SU H_0: COEFF IN PIÙ = 0 RIF ⇒ RIF. H_0

OSS (CHOW) IN ALTERNATIVA FACCIO F TEST TRA MODELLO COMPLETO E INIZIALE.

⇒ PROBLEMA: I PARAMETRI SONO STATISTICAMENTE SIG. ≠ ⇒ i parametri stimati come media tra i stime nei sottogruppi ⇒ STIME DISTORTE E INCONSISTENTI

⇒ STIMATORE DISTORTO e INCONSISTENTE (problema grave)
le stime saranno media pesata delle stime nei sottogruppi ⇒ leggo articolo.

⇒ ARTICOLO PONE DUBBI SU VALIDITÀ DEL MODELLO DI SOLOW ⇒ Autori aggiungono altro regressore: FATTORE UMANO ≃ CONOSCENZA, ISTRUZIONE LAVORATORI per spiegare differenza paesi ricchi vs poveri.

⇒ NUOVA F. PRODUZIONE
OUTPUT = $Y = F(\underset{\text{Osk}}{\overset{\text{CAPITALE FISICO}}{K}}, L \cdot A, \underset{\text{Osh}}{\overset{\text{CAPITALE UMANO}}{H}})$

ma c'è altra equazione che spiega come aumenta il capitale umano
$$\dot{H}(t) = s_h \cdot Y(t) - \delta \cdot H(t)$$
$$\dot{K}(t) = s_K \cdot Y(t) - \delta K(t)$$

⇒ valore STEADY-STATE $\frac{Y}{AL} = y$ (leggo paper)

→ NUOVO MODELLO: DI SOLOW AUMENTATO DEL CAPITALE UMANO

$$\log\left(\frac{Y}{L}\right) = \beta_0 + \left(\frac{\alpha}{1-\alpha-\beta}\right) \log(s_K) + \frac{\beta}{1-\alpha-\beta} \log(s_H) - \frac{\alpha+\beta}{1-\alpha-\beta} \log(n+g+\delta) + \varepsilon$$

poiché $Y = K^{\alpha} H^{\beta} (AL)^{1-\alpha-\beta}$ (LEGGO PAPER)

cioè aggiungo un nuovo regressore $\log(s_H)$ al modello di prima. (loro usano come misura % pop. attiva iscritta a scuola secondaria)

X CASA RIFACCIO DATI CON PAPER E STIME

3/11/15

MODELLI PER SERIE TEMPORALI (vedi manuale di [illegible])

- DAL PUNTO DI VISTA CLASSICO LA S.S. È COMPOSTA DA PIÙ COMPONENTI:

$$Y_t = \underbrace{T_t}_{\text{TREND LUNGO PERIODO}} + \underbrace{C_t}_{\text{CICLICO}} + \overset{\text{STAGIONALITÀ}}{S_t} + \underbrace{\varepsilon_t}_{\text{ERRORE}}$$

(~ IID $(0, \sigma^2_\varepsilon)$ di solito)

OSS Spesso SS disponibili sono già destagionalizzate.

OSS L'approccio classico vede le componenti trend, ciclo, stagionalità deterministiche e l'errore aleatorio.

OSS L'approccio moderno prevede aleatorietà delle componenti. (approccio di Box e Jenkins)

▲ S.S. = UNA REALIZZAZIONE DEL PROC. STOCASTICO.
⇒ Y_t = realizzazione di un proc. stocastico.

(OSS Si può vedere anche $Y_t = T_t + X_t$ dove T_t = trend deterministico e tutto X_t è PROC. STOC. (noi studio proc. stoc.))

OSS Noi lavoreremo su S.S. già destagionalizzate.

PROC. STOCASTICO

Y_t è v.c. al tempo t

⇒ $Y_1, Y_2, \dots, Y_t, \dots$ è insieme di v.c. riferite tutte allo stesso fenomeno (es. PIL ITA nel tempo)

⇒ una volta osservate le v.c. abbiamo numeri ⇒

⇒ $y_1, \dots, y_T$ = S.S. = DATI = CAMPIONE

e v.c. $Y_{T+1}, Y_{T+2}, \dots$ valori futuri ignoti ⇒ che voglio prevedere.

▲ PROC. STOC = $\{\dots, Y_1, \dots, Y_t, \dots\}$ è SUCCESSIONE ORDINATA DI v.c. tutte riferite al medesimo fenomeno in istanti di tempo t ≠.

t = indice temporale

⇒ PROC. STOC. = $\{Y_t\}_{t=-\infty}^{t=+\infty} = \{Y_t\}_{t \in T}$

con $T = \{-\infty, \dots, +\infty\}$

OSS Quello definito è proc. stoc. a tempo discreto

T = discreto.

⇒ PROBLEMA: DESCRIZIONE PROC. STOC.

se abbiamo un numero finito di variabili (es $n=2$)

⇒ li descrivo con la LEGGE CONGIUNTA:

$(Y_{t_1}, Y_{t_2}) \to F_{t_1 t_2}(y_{t_1}, y_{t_2})$ con $t_1 \neq t_2$

⇒ LEGGE (TEMPORALE/PROBABILISTICA) DEL PROC. STOC.

$\mathcal{L}$ = sarebbe la distr. congiunta infinito dimensionale, ma sotto (th) CONDIZIONI DI KOLMOGOROV (per noi sempre vere) ⇒ basta pensare a tutte le distr. finite dimensionali

⇒ $\mathcal{L} = \{F_Y(y_{t_1}, y_{t_2}, \dots, y_{t_n}), \forall n \geq 0, \forall t_1, \dots, t_n \text{ distinti}\}$

OSS Dal punto di vista pratico è difficile da specificare in generale, si riesce in casi particolari.

es. $Y_t \sim$ i.i.d. $N(\mu, \sigma_y^2)$ $\forall t$

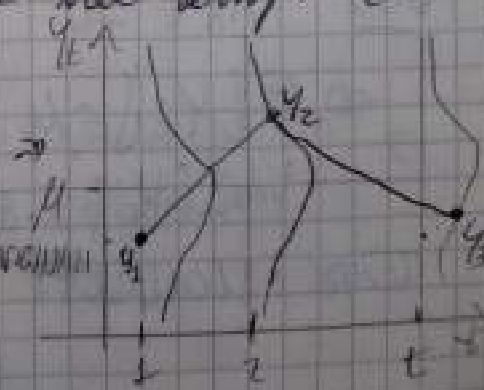
⇒ ∃ $\mathcal{L}$

OSS In generale le s.s. sono correlate nel tempo (es dati economici)

OSS Se proc. stoc. $\{Y_t\}_{t \in T}$ è Y_t i.i.d. $N(\mu, \sigma^2)$ ⇒

⇒ io osservo una parte di $\{Y_t\}_{t \in T}$ disegno unendoli

= REALIZZAZIONE DEL PROC. STOC.



• SCOPO ANALISI S.S. osservata S.S. andare a [illegible] inferenza su caratteristiche del PROC. STOC. che l'ha generata.

OSS Nel caso di dipendenza tra variabili ⇒ più difficile, ma sotto (H) si può fare.

10/11/15

PROC. RISALIRE DA S.S. OSSERVATA A PROC. STOC. CHE L'HA GENERATA È DIFFICILE E FACCIAMO APPROSSIMAZIONI E OSSERVIAMO IPOTESI (H) di STAZIONARIETÀ.

▲ PROC. STOC. STAZIONARIO IN SENSO STRETTO/FORTE.

(Stazionarietà è un fenomeno che evolve allo stesso modo nel tempo dal punto di vista probabilistico)

se $F(y_{t_1}, y_{t_2}, \dots, y_{t_m}) = F(y_{t_1+k}, \dots, y_{t_m+k})$

cioè $(Y_{t_1}, \dots, Y_{t_m}) \overset{d}{=} (Y_{t_1+k}, \dots, Y_{t_m+k})$

$\forall k \in \mathbb{Z} \quad \forall m > 0 \ (\in \mathbb{N}) \quad \forall t_1 \neq t_2 \neq \dots \neq t_m$

OSS con $m = 1 \Rightarrow F(y_{t_1}) = F(y_{t_1+k}) \ \forall k, \forall t_1$

$\Rightarrow (Y_t) \sim$ I.D. LE MARGINALI SONO I.D. (in generale $\cancel{\text{I}}$)

OSS LE DISTR. CONGIUNTE FINITE DIMENSIONALI SONO INVARIANTI RISPETTO A TRASLAZIONI

⇒ IL PROC. HA STESSA STRUTTURA PROBABILISTICA NEL TEMPO.

OSS STAZ ⇒ $F(y_t) = F(y_{t+k}) \ \forall k, \forall t_1$

$\Rightarrow E[Y_t] = \mu \ \forall t$

$VAR[Y_t] = \sigma_y^2 \ \forall t$

TUTTI I MOMENTI UNIVARIATI SONO COSTANTI NEL TEMPO.

Scanned by CamScanner

OSS con $n=2$ $\Rightarrow F(y_{t_1}, y_{t_2}) = F(y_{t_1+k}, y_{t_2+k})\ \forall k$

$\Rightarrow cov(Y_{t_1}, Y_{t_2}) = cov(Y_{t_1+k}, Y_{t_2+k})\ \forall k$

es $cov(Y_1, Y_3) = cov(Y_2, Y_4)$ traslazione rigida $\Rightarrow$ cov $=$

▲ FUNZIONE DI AUTOCOVARIANZA: $\gamma(\cdot,\cdot)$ DI $\{Y_t\}_t$

$cov(Y_{t_1}, Y_{t_2}) \doteq \gamma(t_1, t_2)$

(OSS AUTOCOVARIANZA perché riferita sempre allo stesso fenomeno Y_t es PIL nel tempo)

SE $\{Y_t\}_t$ È STAZ. IN SENSO STRETTO.

$\Rightarrow \gamma(t_1,t_2) = \gamma(t_1+k, t_2+k)\quad \forall k$

$\Rightarrow$ LA F AUTOCOVARIANZA DIPENDE SOLO DAL LAG.

es $\gamma(t_1,t_2) = \gamma(|t_1-t_2|) = \gamma(t_1+k, t_2+k)\quad \forall k$

$\gamma(h)\ , h>0\quad$ con $h \doteq |t_1-t_2|$

$$\Rightarrow \gamma(h) = \begin{cases} \sigma_Y^2 & \text{se } h=0 \\ cov(Y_t, Y_{t-h}) & \text{se } h\neq 0 \end{cases}$$

OSS $(Y_t)_t \overset{iid}{\sim} N(\mu, \sigma_Y^2)$ È STAZ. IN SENSO STRETTO.

$E[Y_t] = \mu$
$$\gamma(h) = \begin{cases} \sigma_Y^2 & , h=0 \\ 0 & , h\neq 0 \end{cases}$$

OSS $(Y_t)_t \sim$ STAZ. IN SENSO STRETTO È I.D. ma in generale $\gamma(h)\neq 0$ per $h\neq 0$ (v.a. correlate nel tempo)

OSS In realtà μ e VAR potrebbero essere tos e cmq essere PROC. STAZ. IN SENSO STRETTO

▲ STAZIONARIETÀ IN SENSO LATO (O DEBOLE O IN COV.)
(+ utilizzata nelle applicazioni considera solo i primi 2 momenti del proc.)

(1) $E[Y_t] = \mu \ \forall t$ con μ ∃ FINITO

(2) VAR $[Y_t] = \sigma^2_y \ \forall t$ con $\sigma^2_y > 0$ ∃ FINITO

(3) COV $(Y_t, Y_{t-h}) = \gamma(|h|) = \gamma(h)$ ∃ FINITE

(OSS (1) VAR < +∞ ⇒ COV FINITO)

OSS STAZ FORTE ⇏ STAZ. DEBOLE perché non è detto che ∃ finiti i primi 2 momenti.
ma se Y_t è STAZ. IN SENSO STRETTO CON ∃ FINITI MOMENTI PRIMI e SECONDI ⇒ Y_t STAZ. IN COV.

OSS STAZ DEBOLE ⇏ STAZ. FORTE
Ma se PROC. NORMALE ⇒ STAZ DEBOLE ⇔ STAZ. FORTE

▲ PROC. NORMALE (STAZ. o NON STAZ.)
SE TUTTE LE DISTR. ~~CONGIUNTE~~ FINITE DIMENSIONALI SONO NORMALI CONGIUNTE.

OSS Normale ⇒ MOMENTI PRIMI e SECONDI FINITI e LA DISTR. NORMALE MULTIVARIATA È NOTA ASSEGNATI MOMENTI PRIMI e SECONDI

■ Th SE $\{Y_t\}_t$ PROC. NORMALE STAZ IN COV ⇒ STAZ IN SENSO STRETTO.

OSS (In generale) PROC. $\{Y_t\} \sim N(\mu_t, \sigma^2_y)$
⇒ $E[Y_t] = \mu_t$ ⇒ LA MEDIA CAMBIA NEL TEMPO ⇒ NON STAZ.

⇒ C'È TREND LIN
(NO STAZ. IN MEDIA)

▲ FUNZIONE DI AUTOCORRELAZIONE $\rho(t, t-h) = \rho(h)$

(È F. DI AUTOCOV. STANDARDIZZATA)

$$\rho_{X,Y} = \frac{COV(X,Y)}{\sqrt{VAR(X) \cdot VAR(Y)}} = CORR(X,Y)$$

$$\Rightarrow \rho(t, t-h) \triangleq CORR(Y_t, Y_{t-h}) = \frac{COV(Y_t, Y_{t-h})}{\sqrt{VAR(Y_t)\, VAR(Y_{t-h})}}$$

in generale è F. AUTOCORR.

OSS SE PROC. È STAZ. IN COV.

$$\rho(t, t-h) \triangleq \rho(h) \triangleq \frac{\gamma(h)}{\sqrt{\gamma(0) \cdot \gamma(0)}} = \frac{\gamma(h)}{\gamma(0)}$$

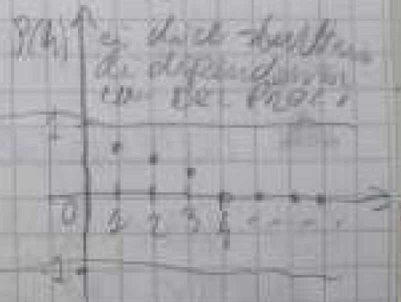
● PROPRIETÀ $\rho(\cdot, \cdot)$

① $\rho(0) = 1$

③ $-1 \leq \rho(h) \leq 1$

② $\rho(\cdot, \cdot)$ È SIMMETRICA

OSS Di solito si usa $\rho(\cdot)$ e non $\gamma(\cdot)$ perché elimina la scala (= unità di misura) della variabile.



Ⓔ $corr(Y_{t+1}, Y_t) = 0{,}6 \Rightarrow$ posso usare Y_t noto per prevedere Y_{t+1} futuro.

$\rho(h) = 0 \quad \forall h \geq 2 \Rightarrow$ non userò altri valori passati. (almeno lin.)

• MODELLI ARMA

$(\varepsilon_t)_t \longrightarrow \boxed{ARMA(p,q)} \longrightarrow (y_t)_t$

PROC. INPUT NON OSS. FILTRO PROC. OUTPUT OSSERVABILE

① $\varepsilon_t \sim IID(0,\sigma^2_\varepsilon)$ → PROC. ERRORE (STAZ. DEBOLE)

② $\varepsilon_t \overset{iid}{\sim} N(0,\sigma^2_\varepsilon)$

③ $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ PROC. WHITE NOISE (WN):

$E[\varepsilon_t]=0 \quad \forall t$

$VAR[\varepsilon_t]=\sigma^2_\varepsilon \quad \forall t$, FINITA

$COV(\varepsilon_t,\varepsilon_{t-h})=0 \quad \forall h\neq 0$

UN PROC. WN CON MEDIA NULLA È PROC. DI V.A. EQUI CON VAR COSTANTE e FINITA e F. AUTOCOV NULLA.

⇒ ε_t NON SONO CORRELATI NEL TEMPO

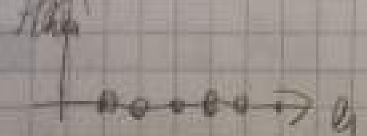
OSS $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ È PARTICOLARE PROC. STAZ IN COV.

$$\gamma(h) = \begin{cases}\sigma^2_\varepsilon & h=0\\ 0 & h\neq 0\end{cases}$$

OSS WN È PROC. INCORRELATO, MA NON PER FORZA IID

▲ CORRELOGRAMMA = GRAFICO DELLA F. AUTOCORR.

OSS PROC. $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ HA CORRELOGRAMMA PIATTO (e anche gli altri 2 ① + restrittivi)



OSS $\rho_{x,y}\neq 0 \Rightarrow y=\alpha+\beta x+u$

$\rho_{xy}=\pm 1 \Rightarrow y=\alpha+\beta x$

Se $\varepsilon_t \sim WN \Rightarrow \varepsilon_{t+1}=\alpha+\beta\,\varepsilon_{t+1-h}+u_{t+1}$

ma $\rho(h)=0 \quad \forall h\neq 0 \Rightarrow \beta=0$

⇒ IL PROC. WN NON È PREVEDIBILE LINEARMENTE DAL PASSATO.

Scanned by CamScanner

▲ MODELLO ARMA (= Autoregressivo a media mobile)

▲ MA(1) (modelli a media mobile di ordine 1)

$$Y_t = \varepsilon_t + a\cdot\varepsilon_{t-1} \quad \text{con } \varepsilon_t \sim WN(\ )$$

MA(2) $Y_t = \varepsilon_t + a_1\varepsilon_{t-1} + a_2\varepsilon_{t-2}$, con $\varepsilon_t \sim WN(\ )$

⇒ è media mobile di ordine 2, perché media del valore corrente e di 2 valori passati

▲ MA(q): $Y_t = \varepsilon_t + a_1\varepsilon_{t-1} + \dots + a_q\varepsilon_{t-q}$

è media mobile perché se ci spostiamo di un periodo anche la media trasla di un periodo.

OSS Se $Y_t \sim MA(1) \Rightarrow$

$$E[Y_t] = E[\varepsilon_t + a\varepsilon_{t-1}] = 0 \quad \forall t$$

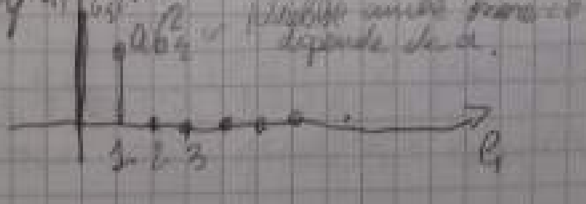
$$VAR[Y_t] = VAR[\varepsilon_t + a\varepsilon_{t-1}] = VAR(\varepsilon_t) + a^2 VAR(\varepsilon_{t-1}) + 2a\,COV(\varepsilon_t, \varepsilon_{t-1}) = \sigma^2_\varepsilon(1+a^2) \stackrel{\Delta}{=} \sigma^2_Y = \gamma_Y(0)$$

$$COV(Y_t, Y_{t-1}) = E[Y_t\cdot Y_{t-1}] - 0\cdot 0 =$$
$$= E[(\varepsilon_t + a\varepsilon_{t-1})\cdot(\varepsilon_{t-1} + a\varepsilon_{t-2})] =$$
$$= E(\varepsilon_t\cdot\varepsilon_{t-1}) + aE(\varepsilon_t\cdot\varepsilon_{t-2}) + aE(\varepsilon_{t-1}^2) + a^2E(\varepsilon_{t-1}\varepsilon_{t-2})$$
$$= a\sigma^2_\varepsilon = \gamma_Y(1)$$

$$COV(Y_t, Y_{t-2}) = E[Y_t\cdot Y_{t-2}] - 0\cdot 0 =$$
$$= E[(\varepsilon_t + a\varepsilon_{t-1})\cdot(\varepsilon_{t-2} + a\varepsilon_{t-3})] \stackrel{\text{conti}}{=} 0 = \gamma_Y(2)$$

OSS In generale $COV(Y_t, Y_{t-h}) = \gamma(h) = 0$ per $h > 1$ se $Y_t \sim MA(1)$

$Y_t \sim$ STAZ IN COV.



Scanned by CamScanner