MS 20/11/24

INFERENZA SU UNA TRASFORMAZIONE LINEARE $C\beta$ DEI PARAMETRI β

Sia $\underset{q\times p}{C}\,\underset{p\times 1}{\beta}$ una matrice di rango

rango$(C) = q \leq p$ = numero di righe linearmente indipendenti a cui siamo interessati.

CASI NOTEVOLI

1) $C = I$ $\quad C\beta = \beta$
siamo interessati a tutti i parametri, globalmente

2) $C = (0\ 0 \ldots 1 \ldots 0\ 0)$
i-esima posizione

$C\beta = \beta_i$
siamo interessati a un singolo β

3) $C = (0\ 1\ 0 \ldots -1\ 0\ 0)$
$C\beta = \beta_J - \beta_k$ per qualche $J \neq k$
siamo interessati a questa differenza, che per es. nel caso di predittore qualitativo (fattore) e' una differenza tra due gruppi (corrispondenti a due livelli del fattore)

4) $C = \left(0 \; 1 \; \ldots \; -\frac{1}{2} \; \ldots \; 0 \; -\frac{1}{2} \; 0 \; 0\right)$

$$C\beta = \beta_J - \frac{\beta_k + \beta_e}{2} \qquad \text{contrasto}$$

x es. nel caso di un solo fattore corrisponde alla differenza tra la media del gruppo J e la media delle medie dei gruppi k e e.
Per esempio, reddito medio italiano $_J$ meno "media europea", cioe' il reddito medio di francesi $_k$ e tedeschi $_e$.

$\square$

Naturalmente $C\beta$ e' stimata da $C\hat{\beta}$, e dal teorema precedente possiamo dedurre

$$C\hat{\beta} \sim N_q(\,?\,,\,?\,)$$

$$E(C\hat{\beta}) = C\,E(\hat{\beta}) = C\beta \qquad \text{stimatore non distorto}$$

$$\begin{aligned} \mathrm{VarCov}(C\hat{\beta}) &= C\,\mathrm{VarCov}(\hat{\beta})\,C' \\ &= \sigma^2 C(X'X)^{-1}C' \end{aligned}$$

quindi

$$C\hat{\beta} \sim N_q(C\beta,\ \sigma^2 C(X'X)^{-1}C')$$

e standardizzando

$$(C\hat{\beta}-C\beta)'(\sigma^2 C(X'X)^{-1}C')^{-1}(C\hat{\beta}-C\beta) \sim \chi^2(q)$$

notiamo che questa variabile aleatoria funzione del solo vettore aleatorio $\hat{\beta}$ (le altre cose, C, β, σ^2, X sono costanti).

D'altra parte, dal teorema della volta scorsa, sappiamo che

$$\frac{E'E}{\sigma^2} = \frac{(n-p)s^2}{\sigma^2} \sim \chi^2(n-p)$$

che dipende solo da s^2, che a sua volta e' indipendente da $\hat{\beta}$ (risultato 3 della volta scorsa)

Otteniamo quindi l'importante

TEO

$$\frac{\overbrace{(C\hat{\beta}-C\beta)'(\cancel{\sigma^2}C(X'X)^{-1}C')^{-1}(C\hat{\beta}-C\beta)}^{\chi^2(q)}}{q \; \underbrace{\frac{\cancel{(n-p)}s^2}{\cancel{\sigma^2}}}_{\chi^2(n-p)} \; \frac{1}{\cancel{n-p}}} \sim F(q, n-p)$$

gdl del NUM

importante risultato distributivo su cui costruiremo:

1) test di ipotesi su $C\beta$

2) regioni di confidenza per $C\beta$

TEST F

ESEMPIO: modello di regressione multipla (polinomiale in x)

$$Y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 z_i + \varepsilon_i$$

$$\begin{pmatrix} Y_1 \\ \vdots \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & x_1^2 & z_1 \\ 1 & x_2 & x_2^2 & z_2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 & z_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

diverse ipotesi di interesse:

$H_0^A : \beta_3 = 0$ $\quad c = (0\,0\,0\,1)$

(e' z un predittore importante?)

$H_0^B : \beta_2 = 0$ $\quad c = (0\,0\,1\,0)$

(e' necessario un termine quadratico in x?)

$H_0^C : \beta_1 = \beta_2 = 0$

(e' necessario x, in ogni caso?)

$$c = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

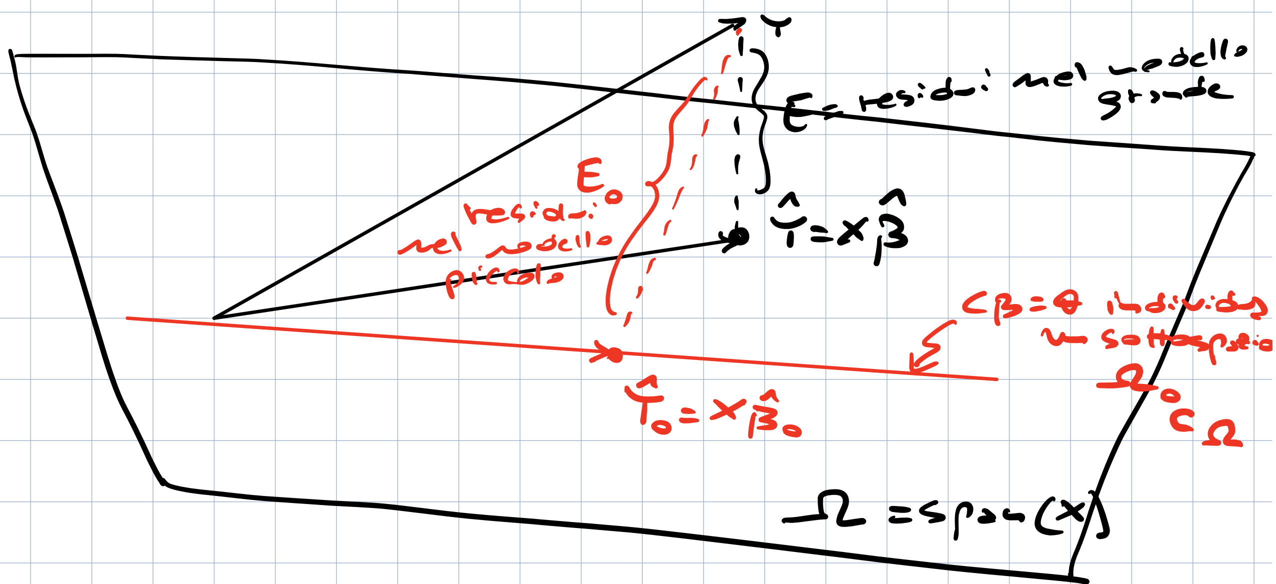
che scriveremo $H_0 : c\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Cioè in generale affrontiamo il test dell'ipotesi lineare generale

$$H_0 : C\beta = \theta$$

(spesso $\theta = 0$, vettore di zeri)

Il vincolo $C\beta = \theta$ identifica un nuovo sottospazio di span(X):



(qualche dettaglio omesso, vedi: Seber Linear Regression Analysis 1977)

Quindi chiedersi se $C\beta = \theta$ è come chiedersi se possiamo restringerci da Ω a Ω_0.

spazio grande

spazio piccolo innestato su Ω

La regola di decisione del sistema di ipotesi

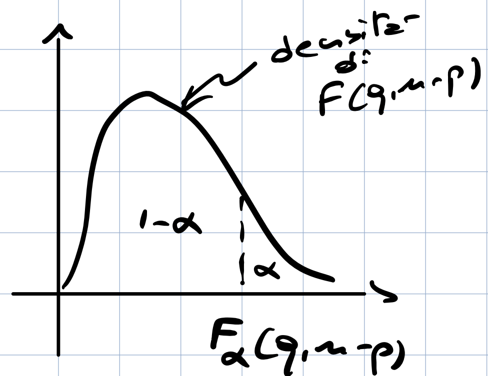
$$H_0: \; C\beta = \theta$$

contro $\quad H_a: \; C\beta \neq \theta$

è fornita dai risultati precedenti:

"Rifiuto H_0 se

$$\frac{(C\hat{\beta}-\theta)'(C(X'X)^{-1}C')^{-1}(C\hat{\beta}-\theta)}{q\, s^2} > F_\alpha(q, n-p)"$$



punto estremo di ordine α, cioè $(1-\alpha)$ quantile, di una v.a. F con q e $n-p$ g.d.l.

Infatti se H_0 è vera, la quantità a sinistra tende a essere piccola, ed è maggiore di $F_\alpha(q, n-p)$ solo con probabilità α (prob. di errore di prima specie).

Se invece H_0 non è vera ($C\beta \neq \theta$) tende a essere più grande.

Si dimostra (Seber cap. 3.9)

$$(C\hat{\beta}-\theta)'(C(X'X)^{-1}C')^{-1}(C\hat{\beta}-\theta)$$

$$= E_0'E_0 - E'E$$

Somma dei quadrati dei residui nel modello piccolo

$E(Y)=X\beta \in \Omega_0$

Restricted
R
S
S

Somma dei quadrati dei residui nel modello grande

$E(Y) \in \Omega$

Unrestricted
Residual
Sum of
Squares

$E_0'E_0$ tende a essere più grande perché norma di una proiezione su uno spazio più ristretto.

Inoltre, poiché $\hat{Y}=X\hat{\beta}$ e $\hat{Y}_0=X\hat{\beta}_0$

$$(Y-\hat{Y}_0)'(Y-\hat{Y}_0) = (Y \pm \hat{Y} - \hat{Y}_0)'(Y \pm \hat{Y} - \hat{Y}_0)$$

$$= (Y-X\hat{\beta})'(Y-X\hat{\beta}) + (X\hat{\beta}_0 - X\hat{\beta})'(X\hat{\beta}_0 - X\hat{\beta})$$

e quindi

$$E_0' E_0 - E'E = RRSS - URSS$$

$$= (X\hat{\beta}_0 - X\hat{\beta})'(X\hat{\beta}_0 - X\hat{\beta})$$

$$= \underbrace{\| \underset{\substack{\text{valori}\\\text{fittati}\\\text{vincolati}}}{\hat{Y}_0} - \underset{\substack{\text{valori}\\\text{fittati}\\\text{liberi}}}{\hat{Y}} \|^2}_{\text{norma al quadrato}}$$

(teo di Pitagora in $\mathbb{R}^n$)

Morale: possiamo riscrivere
" Rifiuto H_0 se

$$\frac{RRSS - URSS}{q \frac{URSS}{n-p}} = \frac{\| \hat{Y}_0 - \hat{Y} \|^2}{q \frac{\| Y - \hat{Y} \|^2}{n-p}} \geq F_\alpha(q, n-p)$$

ESEMPIO (caso notevole 1 modificato)

$$H_0: \beta_1 = \beta_2 = \dots \beta_{p-1} = 0$$

$$\cancel{C\beta = 0}$$
$$\cancel{I\beta = \beta}$$

test del tutto o niente,
minimo sindacale
$q = p-1$ vincoli

e $\qquad \hat{Y}_0 = \begin{pmatrix} \bar{Y} \\ \bar{Y} \\ \vdots \\ \bar{Y} \end{pmatrix}$

e rifiuto H_0 se

$$\frac{RRSS - URSS}{(p-1)\,\frac{URSS}{n-p}} = \frac{\|\hat{Y} - \bar{Y}\cdot\mathbf{1}\|}{(p-1)\,s^2} > F_\alpha(p-1, n-p)$$

che vedete in qualsiasi output di regressione.