Università di Pavia

# Il modello di regressione lineare gaussiano

Eduardo Rossi

Il modello di regressione lineare gaussiano

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

si basa sulla seguente ipotesi

$$\mathbf{y}|\mathbf{X} \sim N(\mathbf{X}\beta, \sigma^2\mathbf{I}_N)$$

oppure

$$\epsilon|\mathbf{X} \sim N(\mathbf{0}, \sigma^2\mathbf{I}_N)$$

Poichè $\widehat{\beta}$ è una funzione lineare di $\epsilon$

$$\widehat{\beta}|\mathbf{X} \sim N(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1})$$

La distribuzione condizionale dipende da $\mathbf{X}$, così che nessuna conclusione può essere tratta sulla distribuzione di $\beta$.

Per costruire la stima per intervalli dobbiamo disporre di una quantità pivotale. Partiamo da:

$$\frac{(\widehat{\beta}-\beta)'(\mathbf{X}'\mathbf{X})(\widehat{\beta}-\beta)}{\sigma^2}$$

poichè

$$\widehat{\beta}-\beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon$$

$$\frac{\epsilon'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon}{\sigma^2} = \frac{\epsilon'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon}{\sigma^2}$$

Possiamo stabilire che

$$\frac{\epsilon'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon}{\sigma^2} = \frac{\epsilon'\mathbf{P}_X\epsilon}{\sigma^2} \sim \chi^2_K$$

Infatti, $\mathbf{P}_X$ è una matrice idempotente di rango $K$, fissato sotto la distribuzione condizionale. Per l'assunzione di gaussianità di $\epsilon|\mathbf{X}$:

$$\sigma^{-1}\epsilon|\mathbf{X} \sim N(\mathbf{0}, \mathbf{I}_N)$$

questa distribuzione non dipende da $\mathbf{X}$ è dunque anche la distribuzione non condizionale, cioè

$$\sigma^{-1}\epsilon \sim N(\mathbf{0}, \mathbf{I}_N)$$

Si può invocare il seguente risultato: **Teorema**: Dato $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$ e $\mathbf{M}$ simmetrica, idempotente con rango $r$, allora

$$\mathbf{z}'\mathbf{M}\mathbf{z} \sim \chi^2_r$$

Consideriamo

$$\frac{\epsilon' \mathbf{M}_X \epsilon}{\sigma^2} = \left(\frac{\epsilon}{\sigma}\right)' \mathbf{M}_X \left(\frac{\epsilon}{\sigma}\right)$$

dove

$$\left(\frac{\epsilon}{\sigma}\right) \sim N(\mathbf{0}, \mathbf{I})$$

anche questa è una forma quadratica in una matrice idempotente, $\mathbf{M}_X$, quindi

$$\frac{\epsilon' \mathbf{M}_X \epsilon}{\sigma^2} \sim \chi^2_{(N-K)}$$

Le due forme quadratiche

$$\frac{\epsilon' \mathbf{P}_X \epsilon}{\sigma^2}$$
$$\frac{\epsilon' \mathbf{M}_X \epsilon}{\sigma^2}$$

sono indipendenti perchè $\mathbf{M}_X \mathbf{P}_X = \mathbf{0}$. Infatti;

$$\mathbf{M}_X \mathbf{P}_X = (\mathbf{I}_N - \mathbf{P}_X)\mathbf{P}_X = \mathbf{P}_X - \mathbf{P}_X = \mathbf{0}$$

Inoltre,

$$s^2 = \frac{\widehat{\epsilon}'\widehat{\epsilon}}{N-K}$$

poichè

$$\widehat{\epsilon} = \mathbf{M}_X \epsilon$$

$$s^2 = \frac{\widehat{\epsilon}'\widehat{\epsilon}}{N-K} = \frac{\epsilon'\mathbf{M}_X \epsilon}{N-K}$$

$$\frac{s^2}{\sigma^2} = \frac{1}{(N-K)} \frac{\epsilon'\mathbf{M}_X \epsilon}{\sigma^2} \sim \chi^2_{N-K}/(N-K)$$

$s^2$ è indipendente da $\widehat{\beta}$ e da $\widehat{\mu} = \mathbf{X}\widehat{\beta}$. Perchè

$$s^2 = \frac{\widehat{\epsilon}'\widehat{\epsilon}}{N-K} \sim \sigma^2 \chi^2_{N-K}/(N-K)$$

$$\widehat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\epsilon$$

dove $\widehat{\epsilon} = \mathbf{M}_X\mathbf{y}$ è indipendente da $\widehat{\mu} = \mathbf{P}_X\mathbf{y}$, perchè

$$E[\widehat{\epsilon}'\widehat{\mu}] = E[\epsilon'\mathbf{M}_X\mathbf{P}_X\mathbf{y}] = 0$$

$s^2$ è funzione solo di $\widehat{\epsilon}$ quindi è indipendente da $\widehat{\mu}$.

**Teorema**: Se $\mathbf{x} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, $\mathbf{Q}$ e $\mathbf{P}$ simmetriche ed idempotenti di rango $r$ e $s$ e $\mathbf{QP} = \mathbf{0}$, allora

$$\frac{\mathbf{x}'\mathbf{P}\mathbf{x}}{\mathbf{x}'\mathbf{Q}\mathbf{x}} \frac{r}{s} \sim F_{(s,r)}$$

Nel nostro caso,

$$\frac{\epsilon'\mathbf{P}_X\epsilon/\sigma^2 K}{\epsilon'\mathbf{M}_X\epsilon/\sigma^2(N-K)} = \frac{\chi^2_{(K)}/K}{\chi^2_{(N-K)}/(N-K)} \sim F_{(K,N-K)}$$

Il rapporto:

$$\frac{\epsilon' \mathbf{P}_X \epsilon / \sigma^2 K}{\epsilon' \mathbf{M}_X \epsilon / \sigma^2 (N-K)}$$

può essere scritto come:

$$\frac{\epsilon' \mathbf{P}_X \epsilon / \sigma^2 K}{\epsilon' \mathbf{M}_X \epsilon / \sigma^2 (N-K)} =$$

$$\frac{(\widehat{\beta} - \beta)'(\mathbf{X}'\mathbf{X})(\widehat{\beta} - \beta)/\sigma^2 K}{s^2/\sigma^2} =$$

$$\frac{(\widehat{\beta} - \beta)'(\mathbf{X}'\mathbf{X})(\widehat{\beta} - \beta)/K}{s^2} =$$

$$\frac{(\widehat{\beta} - \beta)'(\mathbf{X}'\mathbf{X})(\widehat{\beta} - \beta)}{K s^2}$$

$$F^*_{\alpha(K,N-K)} \quad : \quad Pr[F_{(K,N-K)} > F^*_{\alpha(K,N-K)}] = \alpha$$

$0 \leq \alpha \leq 1.$

$$Pr\left[\frac{(\widehat{\beta}-\beta)'(\mathbf{X}'\mathbf{X})(\widehat{\beta}-\beta)}{Ks^2} \leq F^*_{\alpha(K,N-K)}\right] = 1-\alpha$$

$$Pr\left[(\widehat{\beta}-\beta)'(\mathbf{X}'\mathbf{X})(\widehat{\beta}-\beta) \leq Ks^2 F^*_{\alpha(K,N-K)}\right] = 1-\alpha$$

I valori di $\beta$ che soddisfano questa diseguaglianza per dato $\alpha$ sono gli ellissoidi di confidenza nello spazio $K$-dimensionale, centrati su $\widehat{\beta}$.

Nel caso $K = 2$, la forma quadratica

$$(\widehat{\beta}_1 - \beta_1)^2 \sum_{t=1}^{N} x_{1t}^2 + 2(\widehat{\beta}_1 - \beta_1)(\widehat{\beta}_2 - \beta_2) \sum_{t=1}^{N} x_{1t}x_{2t} + (\widehat{\beta}_2 - \beta_2)^2 \sum_{t}^{2} x_{2t}^2$$

Il contorno della funzione implicita

$$ax^2 + byx + cy^2 = K$$

è un'ellisse con centro $(x = 0, y = 0)$, inclinata positivamente quando $b < 0$.

- $\widehat{\beta}_1$ e $\widehat{\beta}_2$ sono positivamente correlati quando $\sum x_{1t}x_{2t} < 0$.
- $\widehat{\beta}_1$ e $\widehat{\beta}_2$ sono negativamente correlati quando $\sum x_{1t}x_{2t} > 0$.

$$H_0: \quad \beta = \beta_0$$
$$H_1: \quad \beta \neq \beta_0$$

Restrizioni lineari

$$H_0: \quad \mathbf{R}\beta = \mathbf{c}$$
$$H_1: \quad \mathbf{R}\beta \neq \mathbf{c}$$

$$\begin{aligned} \mathbf{c} & \qquad (r \times 1) \\ \mathbf{R} & \qquad (r \times K) \\ r(\mathbf{R}) &= r < K \end{aligned}$$

Se

$$\widehat{\beta}|\mathbf{X} \sim N\left(\beta, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right)$$

$$\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)|\mathbf{X} \sim N\left(\mathbf{R}\beta - \mathbf{c}, \sigma^2\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right)$$

Dobbiamo ricavare la distribuzione sotto $H_0 : \mathbf{R}\beta - \mathbf{c} = \mathbf{0}$ di

$$\frac{(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c})}{rs^2}$$

$$\begin{aligned}
\mathbf{R}\widehat{\beta} - \mathbf{c} &= \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} - \mathbf{c} \\
&= \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \epsilon) - \mathbf{c} \\
&= \mathbf{R}\beta + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon - \mathbf{c} \\
&= \mathbf{R}\beta - \mathbf{c} + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon \\
&= \mathbf{0} + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\epsilon
\end{aligned}$$

Sotto $H_0$,

$$(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c}) =$$
$$\epsilon' \underbrace{\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'}_{\mathbf{W}} \epsilon$$

$$(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c}) = \epsilon'\mathbf{W}\epsilon$$

$\mathbf{W}$ simmetrica e idempotente, con

$$\mathbf{W}\mathbf{M}_X = \mathbf{0}$$

infatti

$$\left(\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)(\mathbf{I} - \mathbf{P}_X) =$$

$$\mathbf{W} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}\underbrace{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')}_{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'} =$$

$$\mathbf{W} - \mathbf{W} = \mathbf{0}$$

$$
\begin{aligned}
r(\mathbf{W}) &= tr(\mathbf{W}) \\
&= tr\left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] \\
&= tr\left[(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}\underbrace{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}}_{(\mathbf{X}'\mathbf{X})^{-1}}\mathbf{R}'\right] \\
&= tr\left[(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right] \\
&= tr\left[I_r\right] \\
&= r
\end{aligned}
$$

Le forme quadratiche

$$\frac{(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c})}{\sigma^2} = \frac{\epsilon'\mathbf{W}\epsilon}{\sigma^2} \sim \chi^2_r$$

$$\frac{s^2}{\sigma^2} = \frac{\epsilon'\mathbf{M}_X\epsilon}{(N-K)\sigma^2} \sim \chi^2_{N-K}/(N-K)$$

sono indipendenti. Il rapporto

$$\frac{(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c})}{rs^2} = \frac{\epsilon'\mathbf{W}\epsilon/\sigma^2 r}{\epsilon'\mathbf{M}_X\epsilon/\sigma^2(N-K)} \sim F_{(r,N-K)}$$

è la statistica test per le restrizioni lineari $\mathbf{R}\beta = \mathbf{c}$.

La procedura di test è calcolare questa statistica e rifiutare $H_0$ se il suo valore cade nella regione critica, cioè ha un valore maggiore di $F^*_\alpha$, tale che abbia una probabilità minore di $\alpha$ di essere estratta dalla distribuzione $F_{(r,N-k)}$.

Stimiamo $\beta$ imponendo i vincoli lineari $\mathbf{R}\beta = \mathbf{c}$. Funzione Lagrangiana:

$$\begin{aligned} L &= \frac{1}{2} S(\beta) + \lambda'(\mathbf{R}\beta - \mathbf{c}) \\ &= \frac{1}{2}(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta) + \lambda'(\mathbf{R}\beta - \mathbf{c}) \end{aligned}$$

Condizione del primo ordine:

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= \frac{1}{2} \frac{\partial(\mathbf{y}'\mathbf{y} - 2\beta'\mathbf{X}'\mathbf{y} + \beta'\mathbf{X}'\mathbf{X}\beta)}{\partial \beta} + \frac{\partial[\lambda'(\mathbf{R}\beta - \mathbf{c})]}{\partial \beta} \\ &= -\mathbf{X}'\mathbf{y} + \mathbf{X}'\mathbf{X}\widetilde{\beta} + \mathbf{R}'\widetilde{\lambda} = \mathbf{0} \\ &= -\mathbf{X}'(\mathbf{y} - \mathbf{X}\widetilde{\beta}) + \mathbf{R}'\widetilde{\lambda} = \mathbf{0} \\ \frac{\partial L}{\partial \lambda} &= (\mathbf{R}\widetilde{\beta} - \mathbf{c}) = \mathbf{0} \end{aligned}$$

$$-\mathbf{X}'(\mathbf{y} - \mathbf{X}\widetilde{\beta}) + \mathbf{R}'\widetilde{\lambda} = \mathbf{0} \quad (1)$$

$$(\mathbf{R}\widetilde{\beta} - \mathbf{c}) = \mathbf{0} \quad (2)$$

Premoltiplichiamo per $(\mathbf{X}'\mathbf{X})^{-1}$ la (1):

$$-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} - \mathbf{X}\widetilde{\beta}) + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\widetilde{\lambda} = \mathbf{0}$$

$$-(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}\widetilde{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\widetilde{\lambda} = \mathbf{0}$$

$$-\widehat{\beta} + \widetilde{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\widetilde{\lambda} = \mathbf{0}$$

$$\widetilde{\beta} = \widehat{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\widetilde{\lambda} \quad (3)$$

$$\mathbf{R}\widetilde{\beta} - \mathbf{R}\widehat{\beta} = -\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]\widetilde{\lambda}$$

$$\widetilde{\lambda} = \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{R}\widetilde{\beta})$$

sotto $H_0 : \mathbf{R}\widetilde{\beta} = \mathbf{c}$

$$\widetilde{\lambda} = \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c})$$

sostituendo nella (3)

$$\widetilde{\beta} = \widehat{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)$$

Premoltiplichiamo per $\mathbf{R}$

$$\mathbf{R}\widetilde{\beta} = \mathbf{R}\widehat{\beta} - \underbrace{\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}}_{\mathbf{I}_r}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) = \mathbf{R}\widehat{\beta} - \mathbf{R}\widehat{\beta} + \mathbf{c}$$

$$\mathbf{R}\widetilde{\beta} = \mathbf{c}$$

Notiamo che

$$\widetilde{\beta} - \widehat{\beta} = -(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)$$

$$(\widetilde{\beta} - \widehat{\beta})'(\mathbf{X}'\mathbf{X})(\widetilde{\beta} - \widehat{\beta}) =$$

$$\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \mathbf{R}\, \underbrace{(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})}_{\mathbf{I}_K}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'$$

$$\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) =$$

$$\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) =$$

$$\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \underbrace{\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'}_{\mathbf{I}_r}$$

$$\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) =$$

$$\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)$$

I residui del modello vincolato sono:

$$\begin{aligned}
\widetilde{\epsilon} &= \mathbf{y} - \mathbf{X}\widetilde{\beta} \\
&= \mathbf{y} - \mathbf{X}[\widehat{\beta} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)] \\
&= (\mathbf{y} - \mathbf{X}\widehat{\beta}) + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) \\
&= \widehat{\epsilon} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)
\end{aligned}$$

La somma dei quadrati dei residui del modello vincolato:

$$\begin{aligned}
\widetilde{\epsilon}'\widetilde{\epsilon} &= \left\{\widehat{\epsilon} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)\right\}' \\
&\quad \left\{\widehat{\epsilon} + \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)\right\} \\
\widetilde{\epsilon}'\widetilde{\epsilon} &= \widehat{\epsilon}'\widehat{\epsilon} + \left\{\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right\} \times \\
&\quad \left\{\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)\right\} \\
&\quad +2\left\{\left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right\}\widehat{\epsilon}
\end{aligned}$$

Ma

$$\left\{ \left( \mathbf{R}\widehat{\beta} - \mathbf{c} \right)' \left[ \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \right]^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right\} \widehat{\epsilon} = \mathbf{0}$$

perchè

$$\mathbf{X}'\widehat{\epsilon} = \mathbf{0}$$

quindi

$$
\begin{aligned}
\widetilde{\epsilon}'\widetilde{\epsilon} &= \widehat{\epsilon}'\widehat{\epsilon} + \left\{ \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \right\} \times \\
&\quad \left\{ \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) \right\} \\
&= \widehat{\epsilon}'\widehat{\epsilon} + \\
&\quad \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \mathbf{R} \underbrace{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{X}}_{\mathbf{I}_K} (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' \\
&\quad \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) \\
&= \widehat{\epsilon}'\widehat{\epsilon} + \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \underbrace{\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'}_{\mathbf{I}_r} \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right) \\
&= \widehat{\epsilon}'\widehat{\epsilon} + \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)
\end{aligned}
$$

la differenza tra le somme dei quadrati dei residui del modello vincolato e non è uguale al numeratore della statistica test F:

$$\widetilde{\epsilon}'\widetilde{\epsilon} - \widehat{\epsilon}'\widehat{\epsilon} = \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)' \left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1} \left(\mathbf{R}\widehat{\beta} - \mathbf{c}\right)$$

la statistica test F può essere espressa come

$$\frac{(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c})}{rs^2} = \frac{\widetilde{\epsilon}'\widetilde{\epsilon} - \widehat{\epsilon}'\widehat{\epsilon}}{\widehat{\epsilon}'\widehat{\epsilon}} \; \frac{N-K}{r} \sim F_{(r,N-K)}$$

$$\mathbf{y} = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2 + \epsilon$$

$$\mathbf{X}_1 \quad (N \times K_1)$$

$$\mathbf{X}_2 \quad (N \times K_2)$$

$$\beta_1 \quad (K_1 \times 1)$$

$$\beta_2 \quad (K_2 \times 1)$$

$$K = K_1 + K_2$$

$$H_0 : \beta_1 = \mathbf{0}$$

$$H_0 : \mathbf{R}\beta = \mathbf{0}$$

dove

$$\mathbf{R} = \begin{bmatrix} \mathbf{I}_r & \vdots & \mathbf{0}_{(r \times K_2)} \end{bmatrix}$$

$$\mathbf{R}\beta = \begin{bmatrix} \mathbf{I}_r & \vdots & \mathbf{0}_{(r \times K_2)} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \beta_1$$

sotto $H_0$ il modello è

$$\mathbf{y} = \mathbf{X}_2\beta_2 + \epsilon$$

$$\epsilon|\mathbf{X}_1, \mathbf{X}_2 \sim N(\mathbf{0}, \sigma^2\mathbf{I}_N)$$

$$\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}' = \begin{bmatrix} \mathbf{I}_r & \mathbf{0}_{(r\times K_2)} \end{bmatrix} \begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I}_r \\ \mathbf{0}' \end{bmatrix} = (\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1)^{-1}$$

$$\frac{(\mathbf{R}\widehat{\beta} - \mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta} - \mathbf{c})}{rs^2} = \frac{(\widehat{\beta}_1 - \mathbf{0})'(\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1)(\widehat{\beta}_1 - \mathbf{0})}{rs^2}$$

la statistica test F

$$\frac{(\widehat{\beta}_1)'(\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1)(\widehat{\beta}_1)}{rs^2}$$

ma

$$\widehat{\beta}_1 = [\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1]^{-1}\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{y}$$

$$\frac{(\mathbf{y}'\mathbf{M}_{X_2}\mathbf{X}_1)(\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1)^{-1}(\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{y})}{rs^2}$$

$$\mathbf{M}_X = \mathbf{M}_{X_2} - \mathbf{M}_{X_2}\mathbf{X}_1(\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_{X_2}$$

$$\mathbf{y}'\mathbf{M}_{X_2}\mathbf{X}_1(\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{M}_{X_2}\mathbf{y} = \mathbf{y}'(\mathbf{M}_{X_2} - \mathbf{M}_X)\mathbf{y} = \mathbf{y}'\mathbf{M}_{X_2}\mathbf{y} - \mathbf{y}'\mathbf{M}_X\mathbf{y}$$

$\mathbf{y}'\mathbf{M}_{X_2}\mathbf{y}$ somma dei quadrati del modello vincolato, $\mathbf{y}'\mathbf{M}_X\mathbf{y}$ somma dei quadrati del modello non vincolato.

$$\begin{aligned}
\mathbf{y} &= \mathbf{x}_1\beta_1 + \mathbf{X}_2\beta_2 + \epsilon \\
&= \begin{bmatrix} \mathbf{x}_1 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \epsilon \\
&= \mathbf{X}\beta + \epsilon
\end{aligned}$$

$$\epsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_N)$$

$$\begin{aligned}
\mathbf{x}_1 &: (N \times 1) \\
\mathbf{X}_2 &: (N \times (K-1))
\end{aligned}$$

$$
\begin{aligned}
(\mathbf{X}'\mathbf{X})^{-1} &= \left\{ \begin{bmatrix} \mathbf{x}_1' \\ \mathbf{X}_2' \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 & \mathbf{X}_2 \end{bmatrix} \right\}^{-1} \\
&= \begin{bmatrix} \mathbf{x}_1'\mathbf{x}_1 & \mathbf{x}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{x}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} (\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1} & * \\ * & * \end{bmatrix}
\end{aligned}
$$

Ipotesi nulla

$$
\begin{aligned}
H_0 :& \quad \beta_1 = c \\
H_1 :& \quad \beta_1 \neq c
\end{aligned}
$$

$$
\mathbf{R} = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix} \qquad r = 1
$$

Restrizioni lineari

$$H_0 : \quad \mathbf{R}\beta = c$$

$$H_1 : \quad \mathbf{R}\beta \neq c$$

$$\frac{(\mathbf{R}\widehat{\beta}-\mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta}-\mathbf{c})}{rs^2} = \frac{(\widehat{\beta}_1 - c)'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\widehat{\beta}_1-\mathbf{c})}{s^2}$$

$$\begin{aligned}(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}') &= \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} (\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1} & * \\ * & * \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= (\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}\end{aligned}$$

Ricordando che il rapporto

$$\frac{(\mathbf{R}\widehat{\beta}-\mathbf{c})'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\mathbf{R}\widehat{\beta}-\mathbf{c})}{rs^2} = \frac{\epsilon'\mathbf{W}\epsilon/\sigma^2 r}{\epsilon'\mathbf{M}_X\epsilon/\sigma^2(N-K)} \sim F_{(r,N-K)}$$

è la statistica test per le restrizioni lineari $\mathbf{R}\beta = \mathbf{c}$.

Sotto l'ipotesi di normalità di $\epsilon$ lo stimatore OLS $\widehat{\beta}$ è statisticamente indipendente dal vettore dei residui $\widehat{\epsilon}$ e da tutte le funzioni di $\widehat{\epsilon}$, come $s^2$:

$$\frac{(\widehat{\beta}_1 - c)'(\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}')^{-1}(\widehat{\beta}_1 - \mathbf{c})}{s^2} = \frac{(\widehat{\beta}_1 - c)^2}{s^2(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}} \sim F_{(1,N-K)}$$

$$\frac{\widehat{\beta}_1 - c}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}} \sim t_{(N-K)}$$

perchè $t^2_{(N-K)} = F_{(1,N-K)}$.

La strategia è rifiutare l'ipotesi nulla se la statistica test assume un valore che è inverosimile se l'ipotesi nulla fosse vera. Si rifiuta l'ipotesi nulla se

$$Pr\left\{\left|\frac{\widehat{\beta}_1 - c}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}}\right| > t^*_{\alpha/2,N-K}\right\} < \alpha$$

cioè se

$$\left|\frac{\widehat{\beta}_1 - c}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}}\right| > t^*_{\alpha/2,N-K}$$

dove

$$t^*_{\alpha/2,N-K}: \quad Pr\{|t_{(N-K)}| > t^*_{\alpha/2,N-K}\} = \alpha$$

Ipotesi nulla

$$H_0: \quad \beta_1 = 0$$
$$H_1: \quad \beta_1 \neq 0$$

$$\frac{\widehat{\beta}_1}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}} \sim t_{(N-K)}$$

rifiutiamo l'ipotesi nulla se

$$\left|\frac{\widehat{\beta}_1}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}}\right| > t^*_{\alpha/2,N-K}$$

Ipotesi nulla

$$H_0: \quad \beta_1 < c$$
$$H_1: \quad \beta_1 \geq c$$

La statistica test rimane

$$\frac{\widehat{\beta}_1 - c}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}}$$

rifiutiamo l'ipotesi nulla se

$$\frac{\widehat{\beta}_1 - c}{s\sqrt{(\mathbf{x}_1'\mathbf{M}_{X_2}\mathbf{x}_1)^{-1}}} > t^*_{\alpha,N-K}$$

$$t^*_{\alpha,N-K}: \quad Pr\{t_{(N-K)} > t^*_{\alpha,N-K}\} = \alpha$$

L'intervallo di confidenza per $\beta_k$

$$Pr\left\{\widehat{\beta}_k - t^*_{\alpha/2}\widehat{se(\widehat{\beta}_k)} \leq \beta_k \leq \widehat{\beta}_k + t^*_{\alpha/2}\widehat{se(\widehat{\beta}_k)}\right\} = 1 - \alpha$$

Test congiunto che tutti i coefficienti eccetto il termine costante siano uguali a zero:

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{(K-1)\times 1} \vdots & \mathbf{I}_{K-1} \end{bmatrix} \qquad r = K-1$$

$$H_0 : \quad \mathbf{R}\beta = \mathbf{0}$$

$$H_1 : \quad \mathbf{R}\beta \neq \mathbf{0}$$

Modello ristretto:

$$\mathbf{y} = \iota\beta_1 + \epsilon$$

$$\iota = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\widehat{\beta}_1 = (\iota'\iota)^{-1}\iota'\mathbf{y} = \frac{1}{N}\sum_t y_t$$

$$\widehat{\mathbf{y}} = \iota\widehat{\beta}_1 = \iota\overline{y}$$

$$\widetilde{\epsilon} = \mathbf{y} - \widehat{\mathbf{y}} = \mathbf{y} - \iota\overline{y} = \mathbf{M}_\iota\mathbf{y}$$

$$\widetilde{\epsilon}'\widetilde{\epsilon} = \mathbf{y}'\mathbf{M}_{\iota}\mathbf{y}$$

Il modello sotto l'ipotesi alternativa è:

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

La somma dei quadrati dei residui del modello non ristretto:

$$\widehat{\epsilon}'\widehat{\epsilon} = (\mathbf{y} - \mathbf{X}\widehat{\beta})'(\mathbf{y} - \mathbf{X}\widehat{\beta})$$

$$\begin{aligned}
\widehat{\epsilon}'\widehat{\epsilon} &= \mathbf{y}'\mathbf{y} - 2\widehat{\beta}'\mathbf{X}'\mathbf{y} + \widehat{\beta}'\mathbf{X}'\mathbf{X}\widehat{\beta} \\
&= \mathbf{y}'\mathbf{y} - 2\widehat{\beta}'\mathbf{X}'(\mathbf{X}\widehat{\beta} + \widehat{\epsilon}) + \widehat{\beta}'\mathbf{X}'\mathbf{X}\widehat{\beta} \\
&= \mathbf{y}'\mathbf{y} - 2\widehat{\beta}'\mathbf{X}'\mathbf{X}\widehat{\beta} - 2\widehat{\beta}'\underbrace{\mathbf{X}'\widehat{\epsilon}}_{\mathbf{0}} + \widehat{\beta}'\mathbf{X}'\mathbf{X}\widehat{\beta} \\
&= \mathbf{y}'\mathbf{y} - \widehat{\beta}'\mathbf{X}'\mathbf{X}\widehat{\beta}
\end{aligned}$$

La statistica test $F$:

$$F = \frac{\widetilde{\epsilon}'\widetilde{\epsilon} - \widehat{\epsilon}'\widehat{\epsilon}}{\widehat{\epsilon}'\widehat{\epsilon}} \; \frac{N-K}{K-1}$$

è uguale a

$$\frac{\widetilde{\epsilon}'\widetilde{\epsilon}/(\mathbf{y}'\mathbf{M}^0\mathbf{y}) - \widehat{\epsilon}'\widehat{\epsilon}/(\mathbf{y}'\mathbf{M}_\iota\mathbf{y})}{\widehat{\epsilon}'\widehat{\epsilon}/(\mathbf{y}'\mathbf{M}_\iota\mathbf{y})} \; \frac{N-K}{K-1}$$

ma

$$R^2 = 1 - \frac{\widehat{\epsilon}'\widehat{\epsilon}}{\mathbf{y}'\mathbf{M}_\iota\mathbf{y}}$$

quindi

$$F = \frac{1-(1-R^2)}{1-R^2} \; \frac{N-K}{K-1} = \frac{R^2}{1-R^2} \; \frac{N-K}{K-1} \sim F_{(K-1,N-K)}$$

In una regressione multipla, l'$R^2$ aumenta con l'aumentare del numero dei regressori.

$\overline{R}^2$ aggiustato per i gradi di libertà:

$$\overline{R}^2 = 1 - \frac{\widehat{\epsilon}'\widehat{\epsilon}/(N-K)}{\mathbf{y}'\mathbf{M}_\iota\mathbf{y}/(N-1)}$$

$$\overline{R}^2 = 1 - \frac{(N-1)}{(N-K)}(1-R^2)$$

L'$\overline{R}^2$ diminuisce se la variabile eliminata ha una statistica test t maggiore di 1 e viceversa.