STATISTICA, anno LXIV, n. 2, 2004

PROFILI ECONOMETRICI DEI TEST DI IPOTESI: LA VERIFICA DELLA SPECIFICAZIONE DEL MODELLO

M. Faliva, M.G. Zoia

1. INTRODUZIONE

Le procedure di convalida – preliminari ad ogni eventuale utilizzo dei modelli econometrici – hanno lo scopo di verificare, alla luce dell'evidenza empirica, l'attitudine del modello in esame a fornire una rappresentazione soddisfacente della realtà economica oggetto di indagine.

La convalida del modello si articola in diversi aspetti, fra cui la verifica della specificazione che è espressamente finalizzata all'esame dell'attendibilità delle ipotesi in cui si articola il modello. L'obiettivo è quello di pervenire ad individuare un modello correttamente specificato che rifletta gli aspetti qualificanti del processo generatore dei dati.

Nell'ambito dell'analisi della specificazione è consuetudine distinguere fra test di specificazione e test di cattiva specificazione. Qui di seguito focalizzeremo l'attenzione sui primi.

I test di specificazione poggiano sull'assunto che il modello sia correttamente specificato. Ed è questa presunzione che spiana la strada al ricorso alla teoria della verosimiglianza ai fini inferenziali.

Nel sistema di ipotesi sottostante la costruzione di un test di specificazione l'ipotesi nulla (H_0) consiste tipicamente in una asserzione di validità della specificazione assunta per il modello, laddove l'ipotesi alternativa (H_1) rispecchia un modello più generale.

L'oggetto dell'inferenza verte sulla bontà del modello parsimonioso corrispondente all'ipotesi nulla, che risulta "annidato" nel modello più articolato contemplato dall'ipotesi alternativa, e di riflesso risulta desumibile da quest'ultimo a seguito dell'imposizione di vincoli parametrici ad hoc. Di riflesso i test di specificazione risultano in concreto investire la significatività o meno di restrizioni parametriche che rappresentano, in senso figurato, una sorta di ponte mobile – che scorre sulle guide dell'evidenza empirica – fra ipotesi alternativa e nulla.

2. LA MASSIMA VEROSIMIGLIANZA E I TEST CONNESSI

I test suggeriti per la verifica della specificazione sono essenzialmente riconducibili nell'alveo dei principi noti come criterio di Wald (W), dei Moltiplicatori di Lagrange (LM) e del Rapporto di Verosimiglianza (LR) e si basano sulla teoria della verosimiglianza.

Al fine di illustrare la "ratio" sottostante le tre procedure all'oggetto, si consideri un modello così specificato:

$$\underset{(N,1)}{\mathbf{y}} = \mathbf{f}(\mathbf{X}, \boldsymbol{\beta}) + \boldsymbol{\varepsilon} \tag{2.1}$$

$$\boldsymbol{\varepsilon} \sim \mathcal{D}(\mathbf{o}, \boldsymbol{\Omega}(\boldsymbol{\varsigma})), \ \mathrm{r}(\boldsymbol{\Omega}) = N \tag{2.2}$$

$$\boldsymbol{\vartheta} = \begin{bmatrix} \underset{(K,1)}{\boldsymbol{\beta}} \\ \underset{(H,1)}{\boldsymbol{\varsigma}} \end{bmatrix} \in \Theta \subseteq R^{K+H}, K + H < N \tag{2.3}$$

dove $\mathbf{y}$ rappresenta il vettore di N osservazioni sulla variabile che si intende spiegare (endogena), $\mathbf{X}$ rappresenta la matrice delle osservazioni relative alle variabili esplicative, $\boldsymbol{\varepsilon}$ è un vettore di componenti stocastiche di disturbo non sistematiche, con matrice di dispersione $\boldsymbol{\Omega}(\boldsymbol{\varsigma})$, $\mathcal{D}$ è una distribuzione data, che di norma si identifica con la gaussiana, $\boldsymbol{\vartheta}$ è il vettore dei parametri del modello e Θ sta ad indicare un sottospazio di R^{K+H}: il sottospazio è proprio o improprio a seconda che nel modello operino o meno delle restrizioni parametriche.

Ai fini della successiva analisi faremo riferimento ad un sistema di ipotesi, nulla (H_0) ed alternativa (H_1), da sottoporre a verifica alla luce dell'evidenza empirica, così articolato:

$$\begin{cases} H_0 : \boldsymbol{\rho}(\boldsymbol{\vartheta}) = \mathbf{o}_M \Rightarrow \boldsymbol{\vartheta} \in \Theta \subset R^{K+H} \\ H_1 : \boldsymbol{\rho}(\boldsymbol{\vartheta}) \neq \mathbf{o}_M \Rightarrow \boldsymbol{\vartheta} \in R^{K+H} \end{cases} \tag{2.4}$$

dove $\boldsymbol{\rho}$ sta ad indicare un opportuno vettore di M ($M < K + H$) operatori funzionali indipendenti.

Le stime, vincolata $\tilde{\boldsymbol{\vartheta}}$ – sotto H_0 – e libera $\hat{\boldsymbol{\vartheta}}$ – sotto H_1 – dei parametri $\boldsymbol{\vartheta}$, ai sensi del criterio della massima verosimiglianza sono definite, rispettivamente, come segue:

$$\tilde{\boldsymbol{\vartheta}} = \arg\max_{\boldsymbol{\vartheta} \in \Theta} \ l(\boldsymbol{\vartheta}) \tag{2.5}$$

$$\hat{\boldsymbol{\vartheta}} = \arg\max_{\boldsymbol{\vartheta}} \ l(\boldsymbol{\vartheta}) \tag{2.6}$$

dove $l(\vartheta)$ sta ad indicare il logaritmo della funzione di verosimiglianza[1].

Lo stimatore $\tilde{\vartheta}$ è propriamente la soluzione del problema di programmazione classica:

$$\begin{cases} \max\limits_{\vartheta} \ l(\vartheta) \\ \text{condizionato da} \\ \rho(\vartheta) = \mathbf{o}_{\mathrm{M}} \end{cases} \tag{2.7}$$

Introducendo il vettore $\boldsymbol{\lambda}$ dei moltiplicatori di Lagrange e costruendo quindi il lagrangiano:

$$L(\vartheta, \boldsymbol{\lambda}) = l(\vartheta) + \boldsymbol{\lambda}'\rho(\vartheta) \tag{2.8}$$

le condizioni di primo ordine per un punto stazionario di $L(\vartheta, \boldsymbol{\lambda})$ e nella fattispecie per un massimo vincolato di $l(\vartheta)$ rispetto a ϑ sono date da:

$$\begin{cases} \dfrac{\partial L(\vartheta,\boldsymbol{\lambda})}{\partial \vartheta} = \mathbf{o}_{K+H} \\ \dfrac{\partial L(\vartheta,\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \mathbf{o}_{M} \end{cases} \Rightarrow \begin{cases} \mathbf{s}(\vartheta) + \mathbf{J}(\vartheta)\,\boldsymbol{\lambda} = \mathbf{o}_{K+H} \\ \rho(\vartheta) = \mathbf{o}_{\mathrm{M}} \end{cases} \tag{2.9}$$

avendo indicato con $\mathbf{s}(\vartheta)$ e con $\mathbf{J}(\vartheta)$ rispettivamente il gradiente del logaritmo della verosimiglianza[2] e la matrice jacobiana dei vincoli, i.e[3]:

---

[1] Per il modello in esame la funzione $l(\vartheta)$ assume la forma (cfr. Faliva, 1987, p. 203 ss.):

$$l(\vartheta) = l(\beta,\varsigma) = -\frac{\boldsymbol{N}}{2}\cdot \lg 2\pi + \lg \left|\det \tilde{\mathbf{J}}\right| - \frac{1}{2}\cdot \lg(\det \Omega) - \frac{1}{2}\cdot \boldsymbol{\varepsilon}'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon} =$$

$$= -\frac{\boldsymbol{N}}{2}\cdot \lg 2\pi + \lg \left|\det \tilde{\mathbf{J}}\right| - \frac{1}{2}\cdot \lg(\det \boldsymbol{\Omega}) - \frac{1}{2}\ (\mathbf{y}-\mathbf{f})'\cdot \boldsymbol{\Omega}^{-1}\cdot(\mathbf{y}-\mathbf{f}) \tag{o}$$

dove $\boldsymbol{\Omega}$ sta per $\boldsymbol{\Omega}(\varsigma)$ e $\mathbf{f}$ per $f(\boldsymbol{X}, \beta)$, per comodità notazionale, e $\tilde{\mathbf{J}}$ sta ad indicare la matrice jacobiana:

$$\tilde{\mathbf{J}} = \frac{\partial \varepsilon}{\partial \mathbf{y}'} = \frac{\partial(\mathbf{y} - f(\mathbf{X},\beta))}{\partial \mathbf{y}'} = I_N - \frac{\partial f(\mathbf{X},\beta)}{\partial \mathbf{y}'} \tag{oo}$$

[2] Per il modello in esame il gradiente all'oggetto è esprimibile come: cfr. Faliva, op. cit.:

$$\mathrm{s}(\vartheta) = \begin{cases} \dfrac{\partial \boldsymbol{l}(\beta,\varsigma)}{\partial \beta'} = \mathbf{y}'\boldsymbol{\Omega}^{-1}\dfrac{\partial \mathbf{f}}{\partial \beta'} - \mathbf{f}'\boldsymbol{\Omega}^{-1}\dfrac{\partial \mathbf{f}}{\partial \beta'} - (\boldsymbol{vec}\,\boldsymbol{J}^{-1})'\dfrac{\partial^2 \mathbf{f}}{\partial \mathbf{y}\partial \beta'} \\ \dfrac{\partial \boldsymbol{l}(\beta,\varsigma)}{\partial \varsigma'} = -\dfrac{1}{2}\{\boldsymbol{vec}\ (\boldsymbol{\Omega} - (\mathbf{y}-\mathbf{f})\cdot(\mathbf{y}-\mathbf{f})')\}'\cdot(\boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1})\cdot\dfrac{\partial \boldsymbol{vec}\ \boldsymbol{\Omega}}{\partial \varsigma'} \end{cases}$$

[3] Sotto condizioni molto generali (cfr. Crowder, 1976) vale il seguente risultato cruciale:

$$\underset{K+H,1}{\mathbf{s}(\vartheta)} = \frac{\partial l(\vartheta)}{l\vartheta} \tag{2.10}$$

$$\underset{K+H,M}{\mathbf{J}(\vartheta)} = \frac{\partial \boldsymbol{\rho}'(\vartheta)}{\partial \vartheta}, \; r\,(\mathbf{J}\,(\vartheta)) = M \tag{2.11}$$

La (2.9) è identicamente vera per $\vartheta = \tilde{\vartheta}$ mentre per $\vartheta = \hat{\vartheta}$ sussiste l'uguaglianza

$$\mathbf{s}(\hat{\vartheta}) = \boldsymbol{o} \tag{2.12}$$

Sussiste la relazione:

$$l(\hat{\vartheta}) \geq l(\tilde{\vartheta}) \tag{2.13}$$

che risulta verificata col segno di uguaglianza qualora $\boldsymbol{\lambda} = \boldsymbol{o}$, ovvero se $\hat{\vartheta}$ soddisfa i vincoli.

Alla luce di queste premesse, possibili misure di discrepanza fra enunciato dell'ipotesi nulla ed evidenza empirica sono rappresentate da:

$$i)\; m_1 : l\,(\hat{\vartheta}) - l\,(\hat{\vartheta}) \tag{2.14}$$

$$ii)\; m_2 : \mathbf{s}\,(\hat{\vartheta}) - \mathbf{s}\,(\hat{\vartheta}) \tag{2.15}$$

$$iii)\; m_3 : \boldsymbol{\rho}\,(\hat{\vartheta}) - \boldsymbol{\rho}\,(\hat{\vartheta}) \tag{2.16}$$

Quanto più le suddette misure risultano prossime a zero, e quindi non significative, tanto più plausibile sarà da ritenersi H_0.

Dalle *i)*, *ii)* e *iii)* sopra discendono i criteri

$$i)\; l\,(\hat{\vartheta}) \approx l\,(\hat{\vartheta}) \Rightarrow H_0 \text{ plausibile} \tag{2.17}$$

$$ii)\; \mathbf{s}\,(\hat{\vartheta}) \approx \mathbf{o}_{K+H} \Rightarrow H_0 \text{ plausibile,} \tag{2.18}$$

$$iii)\; \boldsymbol{\rho}\,(\hat{\vartheta}) \approx \mathbf{o}_M \Rightarrow H_0 \text{ plausibile} \tag{2.19}$$

$$\frac{1}{\sqrt{N}}\,\mathbf{s}(\hat{\vartheta}) \overset{N\to\infty}{\sim} \mathcal{N}\left(\mathbf{o}, \underset{N\to\infty}{\text{plim}}\left\{\frac{1}{N}\boldsymbol{H}(\hat{\vartheta})\right\}\right)$$

dove $\boldsymbol{H}(\hat{\vartheta})$ sta ad indicare la matrice (definita positiva):

$$\boldsymbol{H}(\hat{\vartheta}) = -\left.\frac{\partial^2 l(\vartheta)}{\partial\vartheta\partial\vartheta'}\right]_{\vartheta=\hat{\vartheta}}$$

Per quanto concerne la misura m_2 – con l'annesso criterio *ii)* –, da ispezione della prima delle equazioni di verosimiglianza vincolata, di cui alla (2.9) sopra, ovvero

$$\mathbf{s}\,(\vartheta) = -\,\mathbf{J}\,(\vartheta)\boldsymbol{\lambda} \tag{2.20}$$

si evince che l'esame della significatività o meno di $\mathbf{s}\,(\tilde{\vartheta})$ va di pari passo con l'esame della significatività o meno di $\tilde{\boldsymbol{\lambda}}$. Ne consegue che il criterio *ii)* sopra può essere sostituito dal criterio alternativo equivalente:

$$ii')\ \tilde{\boldsymbol{\lambda}} \approx \mathbf{o}_M \Rightarrow H_0 \text{ plausibile} \tag{2.21}$$

Le misure suddette, ed i relativi criteri, stanno alla base dei test noti come test del rapporto di verosimiglianza, *LR* (acronimo dell'inglese Likelihood Ratio), test dei moltiplicatori di Lagrange, *LM* (acronimo dell'inglese Lagrange Multipliers), e test di Wald, *W*, le cui statistiche sono asintoticamente distribuite come una χ^2 – centrale sotto l'ipotesi nulla – con un numero di gradi di libertà pari al numero (M) di vincoli postulati da H_0 (cfr. Engle, 1984).

Il test *LR* rifiuta H_0 per valori elevati della statistica:

$$LR \triangleq 2 \cdot \{l(\hat{\vartheta}) - l(\hat{\vartheta})\} \overset{N\to\infty}{\sim} \chi^2_M \tag{2.22}$$

Il test *LM* rifiuta H_0 per valori elevati della statistica:

$$LM \triangleq \mathbf{s}'\,(\tilde{\vartheta}) \cdot \mathbf{H}^{-1}\,(\tilde{\vartheta}) \cdot \mathbf{s}\,(\tilde{\vartheta}) \overset{N\to\infty}{\sim} \chi^2_M \tag{2.23}$$

dove $\mathbf{H}\,(\tilde{\vartheta})$ sta ad indicare la matrice hessiana della log-verosimiglianza cambiata di segno, i.e.:

$$\mathbf{H}\,(\vartheta) = -\,\frac{\partial^2 l(\vartheta)}{\partial\vartheta\partial\vartheta'} \tag{2.24}$$

valutata in corrispondenza di $\vartheta = \tilde{\vartheta}$ (cfr. nota 3). Il test all'oggetto, ai sensi della (2.20), può essere alternativamente formulato in termini del vettore dei moltiplicatori di Lagrange $\tilde{\boldsymbol{\lambda}}$ (valutato in corrispondenza della soluzione delle equazioni di verosimiglianza vincolata (2.9)), come qui di seguito indicato:

$$LM \triangleq \tilde{\boldsymbol{\lambda}}' \cdot \mathbf{J}\,(\tilde{\vartheta}) \cdot \mathbf{H}^{-1}\,(\tilde{\vartheta}) \cdot \mathbf{J}'(\tilde{\vartheta}) \cdot \tilde{\boldsymbol{\lambda}} \overset{N\to\infty}{\sim} \chi^2_M \tag{2.25}$$

Il test *W* rifiuta H_0 per valori elevati della statistica:

$$W \triangleq \boldsymbol{\rho}'(\hat{\vartheta}) \cdot [\mathbf{J}\,(\hat{\vartheta}) \cdot \mathbf{H}^{-1}(\hat{\vartheta}) \cdot \mathbf{J}'(\hat{\vartheta})]^{-1} \cdot \boldsymbol{\rho}\,(\hat{\vartheta}) \overset{N\to\infty}{\sim} \chi^2_M \tag{2.26}$$

La scelta del test ottimale – data l'equivalenza asintotica delle statistiche test in esame – va condotta o in base alle proprietà nei piccoli campioni, qualora queste siano note, o in base a considerazioni di carattere pragmatico, quali l'onere computazionale. Al riguardo si possono fare le seguenti considerazioni: di norma il test *LM* e il test *W* sono preferibili al test *LR*, in quanto quest'ultimo richiede per il suo computo sia le stime vincolate che quelle non vincolate. Il test *LM* si dimostra particolarmente vantaggioso in tanto in quanto richiede unicamente il calcolo delle stime del modello vincolato.

3. VERIFICHE DI IPOTESI DI RILEVANTE INTERESSE IN ECONOMETRIA

In questo paragrafo tratteremo di una coppia di test – speculari l'uno rispetto all'altro, alla stregua di un Giano bifronte – che sono emblematici delle questioni che si pongono nel corso di un processo, ancorché teoricamente ben fondato, di specificazione e che solo l'evidenza empirica – per il tramite di un ricorso oculato al metodo statistico – può consentire di dirimere.

Più specificamente ci occuperemo, da un lato, della verifica della validità globale della regressione, in presenza di intercetta, nel contesto del modello lineare classico, e, dall'altro, della verifica della sfericità degli errori nel contesto del modello lineare media-dispersione (Pukelsheim, 1976; Zoia, 1993).

Configurandosi il modello classico come un caso particolare del modello media-dispersione, procederemo – ai fini della successiva analisi – alla specificazione di quest'ultimo.

Il modello all'oggetto, caratterizzato da linearità nei coefficienti delle variabili esplicative e da una specificazione parametrica lineare dei secondi momenti degli errori, si articola nelle seguenti ipotesi[4]:

$$\underset{(N,1)}{\boldsymbol{y}} = \underset{(N,K)}{\boldsymbol{K}}\,\underset{(K,1)}{\beta} + \underset{(N,1)}{\varepsilon} \tag{3.1}$$

$$\mathrm{vec}\,\boldsymbol{\Omega}\,(\varsigma) = \underset{(N^2,H)}{\boldsymbol{A}}\,\underset{(H,1)}{\varsigma} \tag{3.2}$$

[4] Il modello all'oggetto discende dalla specificazione più generale di cui alle ipotesi (2.1), (2.2) e (2.3) del paragrafo 2, ai sensi delle posizioni:

$$\boldsymbol{f}(\boldsymbol{X}, \beta) = \boldsymbol{X}\beta \tag{o}$$

$$\mathrm{vec}\,\boldsymbol{\Omega}\,(\varsigma) = \boldsymbol{A}\varsigma \tag{oo}$$

come corollario si ha che (cfr. nota 1):

$$\frac{\partial \boldsymbol{f}}{\partial \beta'} = \boldsymbol{X} \tag{+}$$

$$\frac{\partial \varepsilon}{\partial y'} = \tilde{\boldsymbol{J}} = \boldsymbol{I}_N \tag{++}$$

$$\varepsilon \sim \mathcal{D}(\boldsymbol{o}, \boldsymbol{\Omega}\,(\varsigma)), \qquad r\,(\boldsymbol{\Omega}) = N \tag{3.3}$$

$$\boldsymbol{A}, \boldsymbol{X} \text{ matrici non stocastiche} \tag{3.4}$$

$$r\,(\boldsymbol{X}) = K \tag{3.5}$$

$$r\,(\boldsymbol{A}) = H \tag{3.6}$$

$$\underset{(K+H,\,1)}{\vartheta} = \begin{bmatrix} \beta \\ \varsigma \end{bmatrix} \in \Theta \subseteq R^{K+H},\ K+H < N \tag{3.7}$$

La log-verosimiglianza è data da[5]:

$$l\,(\vartheta) = l\,(\beta, \varsigma\,) = -\frac{N}{2}\lg 2\pi - \frac{1}{2}\,\lg(\det \boldsymbol{\Omega}) - \frac{1}{2}\,(\boldsymbol{y} - \boldsymbol{X}\beta)'\,\boldsymbol{\Omega}^{-1}\,(\boldsymbol{y} - \boldsymbol{X}\beta)' \tag{3.8}$$

dove $\boldsymbol{\Omega}$ sta per $\boldsymbol{\Omega}$ (ς) per comodità notazionale.

Per il gradiente e la matrice hessiana si dimostrano le eleganti espressioni in forma compatta[6]:

$$\underset{(K+H,1)}{\boldsymbol{s}(\vartheta)} = \begin{bmatrix} \dfrac{\partial \boldsymbol{l}(\beta, \varsigma)}{\partial \beta} \\ \dfrac{\partial \boldsymbol{l}(\beta, \varsigma)}{\partial \varsigma} \end{bmatrix} \tag{3.9}$$

$$\frac{\partial \boldsymbol{l}(\beta, \varsigma)}{\partial \beta} = \boldsymbol{X}'\boldsymbol{\Omega}^{-1} \cdot (\boldsymbol{y} - \boldsymbol{X}\beta) \tag{3.10}$$

$$\frac{\partial \boldsymbol{l}(\beta, \varsigma)}{\partial \varsigma} = -\frac{1}{2}\,\boldsymbol{A}' \text{vec}\,\boldsymbol{\Omega}^{-1} + \frac{1}{2}\,\boldsymbol{A}' \text{vec}\,[\boldsymbol{\Omega}^{-1} \cdot (\boldsymbol{y} - \boldsymbol{X}\beta) \cdot (\boldsymbol{y} - \boldsymbol{X}\beta)' \cdot \boldsymbol{\Omega}^{-1}] \tag{3.11}$$

$$\frac{\partial^2 l(\vartheta)}{\partial \vartheta \partial \vartheta'} = -\boldsymbol{H}\,(\vartheta) = \begin{bmatrix} \dfrac{\partial^2 \boldsymbol{l}(\beta, \varsigma)}{\partial \beta \partial \beta'} & \dfrac{\partial^2 \boldsymbol{l}(\beta, \varsigma)}{\partial \beta \partial \varsigma'} \\ \dfrac{\partial^2 \boldsymbol{l}(\beta, \varsigma)}{\partial \varsigma \partial \beta'} & \dfrac{\partial^2 \boldsymbol{l}(\beta, \varsigma)}{\partial \varsigma \partial \varsigma'} \end{bmatrix} \tag{3.12}$$

$$\frac{\partial^2 \boldsymbol{l}(\beta, \varsigma)}{\partial \beta \partial \beta'} = -\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X} \tag{3.13}$$

---

[5] L'espressione all'oggetto discende dalla (o) di Nota 1 come caso particolare corrispondente alle posizioni di Nota 4.

[6] La dimostrazione delle formule all'oggetto richiede il ricorso ad un'algebra piuttosto sofisticata per la derivazione in forma compatta. Le dimostrazioni sono reperibili in Faliva, 2004 (cfr. anche Faliva, 1975).

$$\frac{\partial^2 \boldsymbol{l}(\beta,\varsigma)}{\partial\beta\partial\varsigma'} = \left[\frac{\partial^2 \boldsymbol{l}(\beta,\varsigma)}{\partial\varsigma\partial\beta'}\right]' = -\,[(\boldsymbol{y}-\boldsymbol{X}\beta)'\otimes \boldsymbol{X}'] \cdot (\boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1})\cdot \boldsymbol{A} \tag{3.14}$$

$$\frac{\partial^2 \boldsymbol{l}(\beta,\varsigma)}{\partial\varsigma\partial\varsigma'} = -\frac{1}{2}\boldsymbol{A}' \cdot \{[\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{X}\beta)\cdot(\boldsymbol{y}-\boldsymbol{X}\beta)'\cdot\boldsymbol{\Omega}^{-1}]\otimes\boldsymbol{\Omega}^{-1} +$$

$$+\,\boldsymbol{\Omega}^{-1}\otimes[\boldsymbol{\Omega}^{-1}(\boldsymbol{y}-\boldsymbol{X}\beta)\cdot(\boldsymbol{y}-\boldsymbol{X}\beta)'\cdot\boldsymbol{\Omega}^{-1}] - \boldsymbol{\Omega}^{-1}\otimes\boldsymbol{\Omega}^{-1}\}\cdot\boldsymbol{A} \tag{3.15}$$

Stante che i test di ipotesi di cui ci occuperemo investiranno la significatività o meno di restrizioni parametriche lineari, le ipotesi nulla e alternativa, da sottoporre a verifica alla luce dell'evidenza empirica, assumeranno la forma[7]:

$$\begin{cases} H_0 : \boldsymbol{R}\vartheta = \mathbf{o}_M \\ H_1 : \boldsymbol{R}\vartheta \neq \mathbf{o}_M \end{cases} \tag{3.16}$$

Nel quadro di questa specificazione e dei risultati sopra esposti procediamo quindi ad operare le verifiche di interesse.

i) *Verifica della validità globale della regressione nel modello classico con intercetta*
Il modello di riferimento è il seguente:

$$\boldsymbol{y} = [\boldsymbol{u}, \boldsymbol{X}_1]\cdot\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} + \varepsilon \tag{3.17}$$

$$\varepsilon \sim \mathcal{N}\,(\mathbf{o}, \sigma^2\boldsymbol{I}_N) \tag{3.18}$$

$$\boldsymbol{X}_1 \text{ matrice non stocastica} \tag{3.19}$$

$$r([\boldsymbol{u}, \boldsymbol{X}_1]) = K < N \tag{3.20}$$

che discende dal modello di cui alle ipotesi (3.1) ÷ (3.7) sopra ai sensi delle posizioni:

$$\boldsymbol{\Omega}(\varsigma) = \sigma^2\boldsymbol{I}_N \Leftrightarrow \begin{cases} H = 1, \quad \varsigma = \sigma^2, \boldsymbol{A} = \text{vec } \boldsymbol{I}_N \\ \quad \text{vec } \Omega(\varsigma) = \text{vec } \boldsymbol{I}_N\,\sigma^2 \end{cases} \tag{3.21}$$

---

[7] La (3.16) si configura come un caso particolare della (2.4) del § 2 corrispondente alla posizione:

$$\rho(\vartheta) = \underset{(M,K+H)}{\boldsymbol{R}}\ \underset{(K+H,1)}{\vartheta},\ r(\boldsymbol{R}) = M < K + H$$

col corollario che la matrice jacobiana dei vincoli $\boldsymbol{J}(\vartheta)$ – di cui alla (2.11) del § 2 – è semplicemente $\boldsymbol{R}'$ ed ha pieno rango di colonna.

$$\boldsymbol{\beta} = \begin{bmatrix} \underset{(1,1)}{\beta_0} \\ \underset{(K-1,1)}{\boldsymbol{\beta}_1} \end{bmatrix} \tag{3.22}$$

$$\boldsymbol{X} = \begin{bmatrix} \underset{(N,1)}{\boldsymbol{u}} \; , \; \underset{(N,K-1)}{\boldsymbol{X}_1} \end{bmatrix} \tag{3.23}$$

avendo indicato con $\boldsymbol{u}$ il vettore i cui elementi sono tutti pari ad uno.

Ai fini della verifica della validità globale della regressione le ipotesi nulla ed alternativa da porre a raffronto si formulano come segue:

$$\begin{cases} H_0 : \boldsymbol{\beta}_1 = \boldsymbol{o}_{K-1} \Leftrightarrow \boldsymbol{R}_1\boldsymbol{\beta} = \boldsymbol{o}_{K-1} \Leftrightarrow \boldsymbol{R\vartheta} = \boldsymbol{o}_{K-1} \\ H_1 : \boldsymbol{\beta}_1 \neq \boldsymbol{o}_{K-1} \Leftrightarrow \boldsymbol{R}_1\boldsymbol{\beta} \neq \boldsymbol{o}_{K-1} \Leftrightarrow \boldsymbol{R\vartheta} \neq \boldsymbol{o}_{K-1} \end{cases} \tag{3.24}$$

dove $\boldsymbol{R}_1$ ed $\boldsymbol{R}$ sono le matrici di selezione:

$$\underset{(K-1,K)}{\boldsymbol{R}_1} = [\boldsymbol{o}_{K-1}, \boldsymbol{I}_{K-1}], \; \boldsymbol{R}_1' = \frac{\partial(\boldsymbol{R}_1\boldsymbol{\beta})'}{\partial\boldsymbol{\beta}} \tag{3.25}$$

$$\underset{(K-1,K+1)}{\boldsymbol{R}} = [\boldsymbol{R}_1, \boldsymbol{o}_{K-1}], \; \boldsymbol{R}' = \frac{\partial(\boldsymbol{R\vartheta})'}{\partial\boldsymbol{\vartheta}} \tag{3.26}$$

Le equazioni di verosimiglianza vincolata, sotto H_0, risultano essere[8]:

$$\begin{cases} \dfrac{1}{\sigma^2}\boldsymbol{X}' \cdot (\boldsymbol{y} - \boldsymbol{u}\beta_0) + \boldsymbol{R}_1' \boldsymbol{\lambda} = \boldsymbol{o}_K \\ -\dfrac{N}{2\sigma^2} + \dfrac{1}{2\sigma^4}(\boldsymbol{y} - \boldsymbol{u}\beta_0)' \cdot (\boldsymbol{y} - \boldsymbol{u}\beta_0) = 0 \\ \boldsymbol{\beta}_1 = \boldsymbol{o}_{K-1} \end{cases} \tag{3.27}$$

da cui è agevole pervenire alle stime:

---

[8] Le equazioni all'oggetto (cfr. la (2.9) del paragrafo 2) si giustificano alla luce delle premesse e del fatto che le derivate parziali prime della log-verosimiglianza di cui alle formule (3.10) ed (3.11) si specializzano in questo contesto come segue:

$$\begin{cases} \dfrac{\partial l}{\partial \boldsymbol{\beta}} = \dfrac{1}{\sigma^2}\mathbf{X}' \cdot (\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta}) \\ \dfrac{\partial l}{\partial \sigma^2} = -\dfrac{N}{2\sigma^2} + \dfrac{1}{2\sigma^4}\mathbf{X}' \cdot (\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta})' \cdot (\boldsymbol{y} - \mathbf{X}\boldsymbol{\beta}) \end{cases}$$

dove l sta per $l\,(\beta_0, \boldsymbol{\beta}_1, \sigma^2)$ per comodità notazionale.

$$\hat{\beta}_0 = \frac{1}{N}\boldsymbol{u}'\boldsymbol{y},\ \hat{\boldsymbol{\beta}}_1 = \boldsymbol{o}_{K-1} \tag{3.28}$$

$$\hat{\sigma}^2 = \frac{1}{N}\boldsymbol{y}'\boldsymbol{M}\boldsymbol{y},\ \boldsymbol{M} = \boldsymbol{M}' = \boldsymbol{M}^2 = \boldsymbol{I}_N - \frac{1}{N}\boldsymbol{u}\boldsymbol{u}' \tag{3.29}$$

$$\hat{\boldsymbol{\lambda}} = -\frac{1}{\hat{\sigma}^2}\boldsymbol{X}_1'\boldsymbol{M}\boldsymbol{y} \tag{3.30}$$

dei parametri e dei moltiplicatori.

Con qualche computo si dimostra che la matrice $\boldsymbol{H}(\vartheta)$ – ai sensi delle formule (3.12), (3.13), (3.14) e (3.15) sopra – valutata in corrispondenza alle soluzioni del sistema (3.27), sotto H_0, assume la forma:

$$\boldsymbol{H}\left(\begin{bmatrix}\hat{\beta}_0\\ \hat{\boldsymbol{\beta}}_1\\ \hat{\sigma}^2\end{bmatrix}\right) = \begin{bmatrix} \frac{N}{\hat{\sigma}^2} & \frac{1}{\hat{\sigma}^2}\boldsymbol{u}'\boldsymbol{X}_1 & 0 \\ \frac{1}{\hat{\sigma}^2}\boldsymbol{X}_1'\boldsymbol{u} & \frac{1}{\hat{\sigma}^2}\boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{o} \\ 0 & \boldsymbol{o}' & \frac{N}{2\hat{\sigma}^4}\end{bmatrix} \tag{3.31}$$

Il test dei moltiplicatori di Lagrange per la verifica della validità globale della regressione può essere formulato, alla luce di quanto esposto nel paragrafo 2, come segue:

$$LM \triangleq \hat{\boldsymbol{\lambda}}' \cdot \boldsymbol{R}' \cdot \boldsymbol{H}^{-1}\left(\begin{bmatrix}\hat{\beta}_0\\ \hat{\boldsymbol{\beta}}_1\\ \hat{\sigma}^2\end{bmatrix}\right) \cdot \boldsymbol{R}\cdot\hat{\boldsymbol{\lambda}} \overset{N\to\infty}{\sim} \chi^2_{K-1} \tag{3.32}$$

ed agevolmente valutato grazie ai risultati precedenti.

ii) Verifica della sfericità degli errori nel modello lineare media-dispersione

Il modello di riferimento è quello specificato dalle ipotesi (3.1) ÷ (3.7) sopra con le posizioni:

$$\underset{(N^2,H)}{\boldsymbol{A}} = [\,\underset{(N^2,1)}{\mathrm{vec}\,\boldsymbol{I}_N},\ \underset{(N^2,H-1)}{\boldsymbol{A}_1}\,] \tag{3.33}$$

$$\underset{(H,1)}{\boldsymbol{\varsigma}} = \begin{bmatrix}\underset{(1,1)}{\sigma^2}\\ \underset{(H-1,1)}{\boldsymbol{\varsigma}_1}\end{bmatrix} \tag{3.34}$$

Ai fini della verifica della sfericità degli errori le ipotesi nulla e alternativa da porre a rapporto si formulano come segue:

$$\begin{cases} H_0 : \varsigma_1 = \boldsymbol{o}_{H-1} \Leftrightarrow \bar{\boldsymbol{R}}_1 \varsigma = \boldsymbol{o}_{H-1} \Leftrightarrow \bar{\boldsymbol{R}} \vartheta = \boldsymbol{o}_{H-1} \\ H_1 : \varsigma_1 \neq \boldsymbol{o}_{H-1} \Leftrightarrow \bar{\boldsymbol{R}}_1 \varsigma \neq \boldsymbol{o}_{H-1} \Leftrightarrow \bar{\boldsymbol{R}} \vartheta \neq \boldsymbol{o}_{H-1} \end{cases} \tag{3.35}$$

dove $\bar{\boldsymbol{R}}_1$ ed $\bar{\boldsymbol{R}}$ sono le matrici di selezione:

$$\underset{(H-1,H)}{\bar{\boldsymbol{R}}_1} = [\boldsymbol{o}_{H-1}, \boldsymbol{I}_{H-1}], \; \bar{\boldsymbol{R}}_1' = \frac{\partial(\bar{\boldsymbol{R}}_1 \varsigma)'}{\partial \varsigma} \tag{3.36}$$

$$\underset{(H-1,K+H)}{\bar{\boldsymbol{R}}} = [\boldsymbol{o}_{H-1,K}, \bar{\boldsymbol{R}}_1], \; \bar{\boldsymbol{R}}' = \frac{\partial(\bar{\boldsymbol{R}} \vartheta)'}{\partial \vartheta} \tag{3.37}$$

Da ispezione delle terne (3.35) ÷ (3.37) e (3.24) ÷ (3.26) si evince come i due problemi della *verifica della validità globale* della regressione di cui al punto i) e della *verifica della sfericità* degli errori si presentino come speculari.

La procedura per pervenire alla statistica del test segue pertanto la falsariga di quella precedentemente illustrata.

Le equazioni di verosimiglianza vincolata, *sotto H_0*, risultano essere[9]:

$$\begin{cases} \dfrac{1}{\sigma^2} \boldsymbol{X}'(\boldsymbol{y} - \boldsymbol{X}\beta) = \boldsymbol{o}_K \\ \begin{bmatrix} -\dfrac{N}{2\sigma^2} + \dfrac{1}{2\sigma^4}(\boldsymbol{y} - \boldsymbol{X}\beta)'(\boldsymbol{y} - \boldsymbol{X}\beta) \\ -\dfrac{1}{2\sigma^2} \boldsymbol{A}'_1 \operatorname{vec} \boldsymbol{I}_N + \dfrac{1}{2\sigma^4} \boldsymbol{A}'_1 \operatorname{vec}\{(\boldsymbol{y} - \boldsymbol{X}\beta)(\boldsymbol{y} - \boldsymbol{X}\beta)'\} \end{bmatrix} + \bar{\boldsymbol{R}}_1' \lambda = \boldsymbol{o}_H \\ \varsigma_1 = \boldsymbol{o}_{H-1} \end{cases} \tag{3.38}$$

da cui, con qualche computo, si ottengono le stime:

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1} \cdot \boldsymbol{X}'\boldsymbol{y} \tag{3.39}$$

$$\hat{\sigma}^2 = \frac{1}{N} \mathbf{y}' \bar{\boldsymbol{M}} \boldsymbol{y}, \; \bar{\boldsymbol{M}} = \bar{\boldsymbol{M}}' = \bar{\boldsymbol{M}}^2 = \boldsymbol{I}_N - \boldsymbol{X} \cdot (\boldsymbol{X}'\boldsymbol{X})^{-1} \cdot \boldsymbol{X}' \tag{3.40}$$

$$\varsigma_1 = \boldsymbol{o}_{H-1} \tag{3.41}$$

$$\hat{\lambda} = -\frac{1}{2\hat{\sigma}^2} \boldsymbol{A}_1' \operatorname{vec} \boldsymbol{I}_N + \frac{1}{2\hat{\sigma}^4} \boldsymbol{A}_1' \cdot \{(\boldsymbol{M}\boldsymbol{y}) \otimes (\boldsymbol{M}\boldsymbol{y})\} \tag{3.42}$$

---

[9] Le equazioni all'oggetto si giustificano, alla luce delle premesse, dalle (3.10) e (3.11) sopra.

Con qualche ulteriore computo si dimostra che la matrice $\boldsymbol{H}(\vartheta)$ – ai sensi delle formule (3.12), (3.13), (3.14) e (3.15) sopra – valutata in corrispondenza alle soluzioni del sistema (3.38), sotto H_0, assume la forma:

$$\boldsymbol{H}\left(\begin{bmatrix} \hat{\beta} \\ \hat{\sigma}^2 \\ \hat{\varsigma}_1 \end{bmatrix}\right) = \begin{bmatrix} \boldsymbol{H}_{11} & \boldsymbol{H}_{12} \\ \boldsymbol{H}_{21} & \boldsymbol{H}_{22} \end{bmatrix} \tag{3.43}$$

dove:

$$\boldsymbol{H}_{11} = \frac{1}{\sigma^2}\boldsymbol{X}'\boldsymbol{X} \tag{3.44}$$

$$\boldsymbol{H}_{12} = \boldsymbol{H}'_{21} = [\boldsymbol{0}, \frac{1}{\sigma^4}\{(\boldsymbol{y}-\boldsymbol{X}\hat{\beta})' \otimes \boldsymbol{X}'\}\boldsymbol{A}_1] \tag{3.45}$$

$$\boldsymbol{H}_{22} = \begin{bmatrix} \dfrac{N}{2\hat{\sigma}^2} & \dfrac{1}{\hat{\sigma}^6}[(\bar{\boldsymbol{M}}\boldsymbol{y}) \otimes (\bar{\boldsymbol{M}}\boldsymbol{y})]'\boldsymbol{A}_1 - \dfrac{1}{2\hat{\sigma}^4}(\mathrm{vec}\boldsymbol{I}_N)'\boldsymbol{A} \\ \dfrac{1}{\hat{\sigma}^6}\boldsymbol{A}'_1[(\bar{\boldsymbol{M}}\boldsymbol{y}) \otimes (\bar{\boldsymbol{M}}\boldsymbol{y})] - \dfrac{1}{2\hat{\sigma}^4}\boldsymbol{A}'_1\,\mathrm{vec}\boldsymbol{I}_N & \dfrac{1}{2\hat{\sigma}^6}\boldsymbol{A}'_1[(\bar{\boldsymbol{M}}\boldsymbol{y}\boldsymbol{y}'\bar{\boldsymbol{M}})] \oplus \boldsymbol{I}_N + \boldsymbol{I}_N \oplus [(\bar{\boldsymbol{M}}\boldsymbol{y}\boldsymbol{y}'\bar{\boldsymbol{M}})]\boldsymbol{A}_1 - \dfrac{1}{2\hat{\sigma}^4}\boldsymbol{A}'_1\boldsymbol{A}_1 \end{bmatrix} \tag{3.46}$$

Il test dei moltiplicatori di Lagrange per la verifica della sfericità degli errori può essere formulato alla luce di quanto sopra, sulla falsariga di quanto esposto nel § 2, come segue:

$$\mathrm{LM} \triangleq \hat{\bar{\lambda}} \cdot \bar{\boldsymbol{R}}' \cdot \boldsymbol{H}^{-1}\left(\begin{bmatrix} \hat{\beta} \\ \hat{\sigma}^2 \\ \hat{\varsigma}_1 \end{bmatrix}\right) \cdot \bar{\boldsymbol{R}}\,\hat{\bar{\lambda}} \overset{N\to\infty}{\sim} \chi^2_{H-1} \tag{3.47}$$

e valutato grazie ai risultati precedenti.

Istituto di Econometria e Matematica per le applicazioni economiche, finanziarie ed attuariali
Università Cattolica di Milano

MARIO FALIVA
MARIA GRAZIA ZOIA

RIFERIMENTI BIBLIOGRAFICI

M.Y. CROWDER (1976), *Maximum likelihood estimation for dependent observations*, "J. Roy. Statist. Soc.", B, pp. 45-53.

R.F. ENGLE (1982), *A general approach to Lagrange multiplier model diagnostics*, "J. of Econometrics", pp. 83-104.

R.F. ENGLE (1984), *Wald, Likelihood Ratio and Lagrange Multiplier tests in econometrics*, in Z. Griliches e M.D. Intriligator (eds.), "Handbook of econometrics", vol. II, North Holland, New York.

G.B. EVANS e N.E. SAVIN (1982), *Conflict among the criteria revisited: the W, LR and LM tests*, "Econometrica", pp. 737-748.

M. FALIVA (1975), *Alcune regole di derivazione in termini matriciali*, "Statistica", pp. 329-340.

M. FALIVA (1987), *Econometria: Principi e Metodi*, Utet, Torino.

M. FALIVA (2004), *A dual of Nissen's theorem with applications to statistics*, "Contributi di Ricerca", 65 bis, Istituto di Econometria e Matematica, Università Cattolica, Milano.

A. GARDINI, G. CAVALIERE, M. COSTA e P. PARUOLO (2002), *Econometria*, F. Angeli, Milano.

J. KLEFFE (1977), *Optimal estimation of variance components. A Survey*, "Indian Journal of Statistics", pp. 211-244.

A. NADDEO (1963), *La teoria dei test statistici*, Giuffrè, Milano.

A. NADDEO (1987), *Inferenza statistica*, La Nuova Italia Scientifica, Roma.

F. PERACCHI (1995), *Econometria,* McGrawHill, Milano.

F. PUKELSHEIM (1976), *Estimating variance components in linear models*, "Journal of Multivariate Analysis", pp. 626-629.

M.G. ZOIA (1993), *Inverse generalizzate e stimatori dei parametri nel modello lineare*, "Statistica", pp. 73-85.

RIASSUNTO

Profili econometrici dei test di ipotesi: la verifica della specificazione del modello

Il saggio, nel quadro della verifica della specificazione del modello – previo un excursus sui criteri che poggiano sulla teoria della verosimiglianza – , tratta specificatamente di due aspetti emblematici e virtualmente speculari del processo di specificazione e convalida del modello. L'articolo, nel contesto del modello lineare media-dispersione, affronta, da un lato, il problema della verifica della validità globale della regressione e, dall'altro, il problema della verifica della sfericità degli errori e perviene, avvalendosi di un elegante strumentazione algebrica, ai pertinenti test dei moltiplicatori di Lagrange.

SUMMARY

Econometric profiles of testing of statistical hypotheses: model specification tests

The paper, within the frame of econometric specification testing – after an excursus on criteria based on likelihood theory –, focuses on a pair of emblematic and virtually specular facets of the model specification and validation process. Indeed, within a mean-dispersion linear specification framework, the paper faces the issues of testing on the one hand, the overall significance of the regression and on the other, the sphericalness of disturbances and, making use of an elegant algebraic toolkit, derives the ad hoc Lagrange multiplier test statistics.