STATISTICA, anno LXII, n. 2, 2002

UN TEOREMA DI RAPPRESENTAZIONE PER LO STIMATORE VINCOLATO (*)

Maria Grazia Zoia

1. INTRODUZIONE

Il tema della stima vincolata – dopo la soluzione classica offerta da Theil (1961) e la sua sistematizzazione nella teoria unificata della stima lineare di Rao (1971, 1972, 1973), con la parallela trattazione di Faliva (1972, 1987) – ha recentemente trovato una occasione di rivisitazione ad opera dell'autore (cfr. Faliva e Zoia, 2000) con una rilettura in chiave duale dei risultati classici.

Il presente saggio sviluppa questo approccio alla stima vincolata pervenendo ad un elegante teorema di rappresentazione.

2. IL MODELLO LINEARE VINCOLATO ED IL PROBLEMA DELLA STIMA DEI PARAMETRI

Il modello lineare vincolato – con restrizioni parametriche lineari esatte sui parametri – può essere opportunamente specificato come qui di seguito indicato:

a) Ipotesi relative alle informazioni campionarie

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

$$\mathrm{E}\,(\varepsilon) = \mathbf{0} \tag{2}$$

$$\begin{cases} \mathrm{E}(\boldsymbol{\varepsilon\varepsilon}') = \boldsymbol{\sigma}^2\boldsymbol{\Omega}, \quad \mathrm{tr}(\boldsymbol{\Omega}) = \mathrm{r}(\boldsymbol{\Omega}) = N \\ \boldsymbol{\Omega} \ \text{matrice nota} \end{cases} \tag{3}$$

$\boldsymbol{X}$ matrice di elementi non stocastici (4)

(*) Il presente articolo si inserisce nell'ambito di un progetto condotto con il parziale supporto del finanziamento alla ricerca dell'Università Cattolica (D1.1, 2000-2001).

b) Ipotesi relative alle informazioni a priori

$$\underset{(H,1)}{s} = \underset{(H,K)}{R} \underset{(K,1)}{\beta} \tag{5}$$

***R** e s sono, rispettivamente, una matrice ed un vettore dati di elementi non stocastici, con **R** tale che*:

$$r\,(\boldsymbol{R}) = H < K \tag{6}$$

c) Ipotesi di collegamento

$$r\begin{pmatrix} X \\ R \end{pmatrix} = K < N + H \tag{7}$$

per la cui interpretazione si rinvia ai manuali di uso corrente (Faliva, 1987, pp. 412-416). In termini dei vettori e delle matrici a blocchi:

$$\underset{(N+H,1)}{\tilde{y}} \equiv \begin{bmatrix} y \\ s \end{bmatrix} \tag{8}$$

$$\underset{(N+H,K)}{\tilde{X}} \equiv \begin{bmatrix} X \\ \boldsymbol{R} \end{bmatrix} \tag{9}$$

$$\underset{(N+H,1)}{\tilde{\varepsilon}} \equiv \begin{bmatrix} \varepsilon \\ \eta \end{bmatrix} = \begin{bmatrix} \varepsilon \\ \mathbf{0}_H \end{bmatrix} \tag{10}$$

$$\underset{(N+H,N+H)}{\tilde{\Omega}} \equiv \begin{bmatrix} \Omega & \boldsymbol{O}_{N,H} \\ \boldsymbol{O}_{H,N} & \boldsymbol{O}_H \end{bmatrix} \tag{11}$$

il modello può essere riformulato come segue:

$$\tilde{y} = \tilde{X}\beta + \tilde{\varepsilon} \tag{12}$$

$$\mathrm{E}\,(\tilde{\varepsilon}) = \mathbf{0} \tag{13}$$

$$\begin{cases} \mathrm{E}\,(\tilde{\varepsilon}\tilde{\varepsilon}') = \sigma^2\tilde{\Omega}, \quad \mathrm{tr}(\tilde{\Omega}) = \mathrm{r}(\tilde{\Omega}) = N \\ \tilde{\Omega} \quad \text{matrice nota orlata di zeri} \end{cases} \tag{14}$$

$$r\,(\tilde{X}) = K < N + H \tag{15}$$

$$r\,(\tilde{X}, \tilde{\Omega}) = N + H \tag{16}$$

ovvero nella forma standard di un modello lineare (consistente) con matrice di dispersione singolare (orlata di zeri, nella fattispecie), per il quale valgono i seguenti risultati classici per la stima efficiente dei parametri (Rao 1971 e 1973, p. 293 e successive; Faliva, 1972 e 1987, pp. 417-421):

$$\underset{(K,1)}{\boldsymbol{b}} = \begin{bmatrix} O_{K,N+H} & I_K \end{bmatrix} \begin{bmatrix} \tilde{\Omega} & \tilde{X} \\ \tilde{X}' & O_K \end{bmatrix}^{-1} \begin{bmatrix} \tilde{y} \\ \mathbf{0}_K \end{bmatrix} \tag{17}$$

$$\underset{(K,K)}{V(\boldsymbol{b})} = -\sigma^2 \begin{bmatrix} O_{K,N+H} & I_K \end{bmatrix} \begin{bmatrix} \tilde{\Omega} & \tilde{X} \\ \tilde{X}' & O_K \end{bmatrix}^{-1} \begin{bmatrix} O_{N+H,K} \\ I_K \end{bmatrix} \tag{18}$$

dove $\boldsymbol{b}$ e $V(\boldsymbol{b})$ stanno ad indicare, rispettivamente, lo stimatore dei parametri $\boldsymbol{\beta}$ e la relativa matrice di dispersione.

Riscrivendo – ai sensi della (9) e della (11) sopra – la matrice composta che figura nelle espressioni (17) e (18), nella forma:

$$\begin{bmatrix} \Omega & O_{N,H} & X \\ O_{H,N} & O_H & R \\ X' & R' & O_K \end{bmatrix}$$

e ricorrendo alle usuali formule di inversioni per parti (cfr. Goldberger, 1964, pp. 27-28), con semplici computi si perviene alle seguenti espressioni risolutive per $\boldsymbol{b}$ e $V(\boldsymbol{b})$:

$$\boldsymbol{b} = \begin{bmatrix} I_K & O_{K,H} \end{bmatrix} \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & O_H \end{bmatrix}^{-1} \begin{bmatrix} X'\Omega^{-1}y \\ s \end{bmatrix} \tag{19}$$

$$V(\boldsymbol{b}) = \sigma^2 \begin{bmatrix} I_K & O_{K,H} \end{bmatrix} \begin{bmatrix} X'\Omega^{-1}X & R' \\ R & O_H \end{bmatrix}^{-1} \begin{bmatrix} I_K \\ O_{H,K} \end{bmatrix} \tag{20}$$

Allo stimatore (19) si può altresì pervenire guardando al problema di stima nell'ottica dei metodi di adattamento e, più specificatamente, ai sensi del metodo dei minimi quadrati condizionati, formalmente esprimibile come il problema di programmazione classica qui di seguito indicato:

$$\begin{cases} \min\limits_{\beta}(y - X\beta)'\,\Omega^{-1}(y - X\beta) \\ \text{condizionato da } s = R\beta \end{cases} \tag{21}$$

3. IL TEOREMA DI RAPPRESENTAZIONE PER LO STIMATORE VINCOLATO

Sulla base delle considerazioni svolte nel paragrafo 2, al problema della stima vincolata può essere data una elegante soluzione ai sensi del seguente teorema di rappresentazione:

Teorema. Nel modello lineare vincolato – specificato ai sensi delle ipotesi (1)÷(7) del paragrafo 2 – per lo stimatore lineare efficiente $\boldsymbol{b}$ dei parametri $\boldsymbol{\beta}$ e per la corrispondente matrice di dispersione $V(\boldsymbol{b})$ sussistono le seguenti rappresentazioni:

$$\boldsymbol{b} = \boldsymbol{R}'_{\perp}(\boldsymbol{R}_{\perp}\boldsymbol{A}\boldsymbol{R}'_{\perp})^{-1}\boldsymbol{R}_{\perp}\boldsymbol{X}'\Omega^{-1}\boldsymbol{y} + \left[\boldsymbol{I} - \boldsymbol{R}'_{\perp}(\boldsymbol{R}_{\perp}\boldsymbol{A}\boldsymbol{R}'_{\perp})^{-1}\boldsymbol{R}_{\perp}\boldsymbol{A}\right]\boldsymbol{R}^{g}\boldsymbol{s} \tag{22}$$

$$V(\boldsymbol{b}) = \sigma^{2}\boldsymbol{R}'_{\perp}(\boldsymbol{R}_{\perp}\boldsymbol{A}\boldsymbol{R}'_{\perp})^{-1}\boldsymbol{R}_{\perp} \tag{23}$$

avendo indicato con $\boldsymbol{A}$ la matrice:

$$\boldsymbol{A} = \boldsymbol{X}'\Omega^{-1}\boldsymbol{X} \tag{24}$$

La (1) può essere altresì riformulata come:

$$\boldsymbol{b} = \tilde{\boldsymbol{b}} + \Phi\,(\boldsymbol{y} - \boldsymbol{X}\tilde{\boldsymbol{b}}) \tag{25}$$

ai sensi delle posizioni:

$$\tilde{\boldsymbol{b}} = \boldsymbol{R}^{g}\boldsymbol{s} \tag{26}$$

$$\Phi = \boldsymbol{R}'_{\perp}(\boldsymbol{R}_{\perp}\boldsymbol{A}\boldsymbol{R}'_{\perp})^{-1}\boldsymbol{R}_{\perp}\boldsymbol{X}'\Omega^{-1} \tag{27}$$

con $\boldsymbol{R}'_{\perp}$ e $\boldsymbol{R}^{g}$ che stanno ad indicare, rispettivamente, il complemento ortogonale di $\boldsymbol{R}'$ e l'inversa generalizzata di Moore-Penrose di $\boldsymbol{R}$.

Subordinatamente al verificarsi della condizione di rango:

$$\mathrm{r}\,(\boldsymbol{X}) = K \tag{28}$$

che implica la (7) (ma non viceversa), sussistono le rappresentazioni:

$$\boldsymbol{b} = \{\boldsymbol{I} - \boldsymbol{A}^{-1}\boldsymbol{R}'(\boldsymbol{R}\boldsymbol{A}^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}\}\boldsymbol{A}^{-1}\boldsymbol{X}'\Omega^{-1}\boldsymbol{y} + \boldsymbol{A}^{-1}\boldsymbol{R}'(\boldsymbol{R}\boldsymbol{A}^{-1}\boldsymbol{R}')^{-1}\boldsymbol{s} \tag{29}$$

$$V(\boldsymbol{b}) = \sigma^{2}\boldsymbol{A}^{-1}\{\boldsymbol{A} - \boldsymbol{R}'(\boldsymbol{R}\boldsymbol{A}^{-1}\boldsymbol{R}')^{-1}\boldsymbol{R}\}\boldsymbol{A}^{-1} \tag{30}$$

La (29) può essere altresì riformulata come:

$$\boldsymbol{b} = \hat{\boldsymbol{b}} + \Psi\,(\boldsymbol{s} - \boldsymbol{R}\hat{\boldsymbol{b}}) \tag{31}$$

ai sensi delle posizioni:

$$\hat{\boldsymbol{b}} = \boldsymbol{A}^{-1}\boldsymbol{X}'\Omega^{-1}\boldsymbol{y} \tag{32}$$

$$\Psi = A^{-1}\boldsymbol{R}'(RA^{-1}\boldsymbol{R}')^{-1} \tag{33}$$

Dalle rappresentazioni suddette scaturisce infine l'elegante formula:

$$\boldsymbol{b} = \Phi X\hat{\boldsymbol{b}} + \boldsymbol{\Psi}R\tilde{\boldsymbol{b}} \tag{34}$$

che esprime lo stimatore vincolato $\boldsymbol{b}$ in funzione dello stimatore $\hat{\boldsymbol{b}}$ basato sulle sole informazioni campionarie e dello stimatore "virtuale" $\tilde{\boldsymbol{b}}$ basato sulle informazioni a priori, per il tramite delle matrici di proiezione ΦX e $\Psi \boldsymbol{R}$.

Dimostrazione. Alla luce delle posizioni:

$$\begin{bmatrix} A & R' \\ R & O_H \end{bmatrix}^{-1} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix} \tag{35}$$

le espressioni dello stimatore dei parametri $\boldsymbol{b}$ e della pertinente matrice di dispersione $V(\boldsymbol{b})$ di cui alle formule (19) e (20) divengono:

$$\boldsymbol{b} = C_1 X'\Omega^{-1}\boldsymbol{y} + C_2\boldsymbol{s} \tag{36}$$

$$V(\boldsymbol{b}) = \sigma^2 C_1 \tag{37}$$

La (29) e la (30) (cfr. Theil 1961, 1971; Goldberger 1964), discendono immediatamente dalla classica formula di inversione per parti:

$$\begin{bmatrix} A & R' \\ R & O_H \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} - A^{-1}R'(RA^{-1}R')^{-1}RA^{-1} & A^{-1}R'(RA^{-1}R')^{-1} \\ (RA^{-1}R')^{-1}RA^{-1} & -(RA^{-1}R')^{-1} \end{bmatrix} \tag{38}$$

applicabile subordinatamente al verificarsi della condizione di rango:

$$\mathrm{r}\,(\mathbf{X}) = K \Rightarrow \mathrm{r}\,(X'\Omega^{-1}X) = K \tag{39}$$

La (22) e la (23) discendono per converso dalla formula di inversione per parti (Faliva e Zoia, 2000):

$$\begin{bmatrix} A & R' \\ R & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} R'_\perp(R_\perp AR'_\perp)^{-1}R_\perp & (I - C_1A)R^g \\ (R')^g(I - AC_1) & (R')^g(AC_1A - A)R^g \end{bmatrix} \tag{40}$$

applicabile in generale.

Per quanto concerne la (34) la relativa dimostrazione poggia sulle argomentazioni che seguono.

Sotto l'ipotesi di rango (28) si constata agevolmente che sussiste la relazione:

$$\Phi\boldsymbol{y} = \Phi\, X\hat{\boldsymbol{b}} \tag{41}$$

Uguagliando i blocchi in alto a sinistra delle inverse partizionate (38) e (40), postmoltiplicando per $\boldsymbol{A}$ e ricorrendo alle (27) e (33) sopra, risulta essere:

$$\Psi \boldsymbol{R} = (\boldsymbol{I} - \Phi \boldsymbol{X}) \tag{42}$$

Alla luce della (41) e della (42) si passa dalla (25) alla (34) con computi elementari.

Infine, è agevole verificare le proprietà di idempotenza ed ortogonalità:

$$(\Phi \boldsymbol{X})^2 = \Phi \boldsymbol{X}$$
$$(\Psi \boldsymbol{R})^2 = \Psi \boldsymbol{R}$$
$$\Phi \boldsymbol{X} \Psi \boldsymbol{R} = \boldsymbol{0}$$

Di riflesso le matrici $\Phi \boldsymbol{X}$ e $\Psi \boldsymbol{R}$ sono interpretabili come matrici di proiezione (cfr. Rao e Mitra, 1971).

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

MARIA GRAZIA ZOIA

RIFERIMENTI BIBLIOGRAFICI

M. FALIVA (1972), *Stimatori lineari efficienti dei parametri nel modello di regressione lineare*, "Statistica", XXII, pp. 415-452.

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

M. FALIVA, M.G. ZOIA (2000), *On a partitioned inversion formula having useful applications in econometrics*, "Cahiers du Département d'Econométrie, Université de Genève", n° 2000-1 (pubblicato su *Econometric Theory*, 18, 2002 pp. 525-530).

A.S. GOLDBERGER (1964), *Econometric Theory*, Wiley, N.Y.

R.C. RAO (1971), *Unified theory of linear estimation*, "Sankhya", series A, pp. 371-394.

R.C. RAO (1972), *Unified theory of least squares*, "Communications in Statistics", pp. 1-8.

R.C. RAO (1973), *Linear Statistical Inference and its Applications* (2° ed.) Wiley, N.Y.

R.C. RAO, S.K. MITRA (1971), *Generalized Inverse of Matrices and Its Applications*, Wiley, N.Y.

H. THEIL (1961), *Economic Forecasts and Policy*, North Holland, Amsterdam.

H. THEIL (1971), *Principles of Econometrics*, North Holland, Amsterdam.

RIASSUNTO

Un teorema di rappresentazione per lo stimatore vincolato

L'articolo rivisita il problema della stima vincolata nel modello lineare avvalendosi di un recente risultato (Faliva e Zoia, 2000) sull'inversione per parti, che consente di ottenere una nuova espressione risolutiva per lo stimatore dei parametri, più generale ed al tempo stesso speculare rispetto alla nota formula di Theil (1961). Ne risulta un elegante teorema di rappresentazione in cui risultati originali e classici si compongono in un quadro unitario.

SUMMARY

A representation theorem for the restricted least-squares estimator

A recent result in partitioned inversion (Faliva and Zoia, 2000) paves the way to shed some more light on constrained estimation , as the paper shows for the linear model.

Indeed, a novel closed-form expression for the restricted least-squares (RLS) estimator is derived, which reads as a mirror image of the well-known Theil's (1961) solution and actually turns out to cover the latter as a special case. This leads eventually to an elegant representation theorem for the RLS estimator which combines classical and new results as well.