Regressioni lineari con vincoli: uno studio di simulazione sugli stimatori misti a confronto con gli stimatori vincolati(*)

Agostino Tarsitano
Università degli studi della Calabria
Dipartimento di economia e Statistica
87030 Arcavacata di Rende (Cs)
agotar@unical.it

Riassunto.
Al centro di questo articolo stanno i vincoli di disuguaglianza sui parametri di un modello di regressione lineare uniequazionale. Relazioni stocastiche e funzionali sono uniformate in una trattazione unica attraverso l'approccio bayesiano. In particolare, gli stimatori vincolati si configurano come un valore di compromesso tra gli stimatori ordinari ai minimi quadrati e gli stimatori misti di Theil-Goldberger con prevalenza dell'uno o dell'altro estremo a seconda della efficacia dei vincoli.

Keywords:
regressione vincolate, procedura Theil-Goldberger, Stimatori di Stein, programmazione quadratica.

(*) Lavoro inserito in "Scritti in onore di Francesco Brambilla". Vol. 2, pp.831-842. Edizioni di "Bocconi Comunicazioni", Milano, 1986

1. Introduzione

Il modello studiato è il seguente:

$$\boldsymbol{y} = \boldsymbol{X}\beta + \boldsymbol{u} \tag{1.1}$$

$$\boldsymbol{R}\beta \geq \boldsymbol{r} \tag{1.2}$$

dove

$\boldsymbol{y}$= Vettore di osservazioni sulla variabile dipendente con *dim(**y**)=n*
$\boldsymbol{X}$= Matrice di ordine *(m,n)* di regressori non stocastici con *ran(**X**)=m*
β= Vettore dei coefficienti di regressione con *dim(β)=m*
$\boldsymbol{u}$= Vettore casuale con *dim(**y**)=n* e con distribuzione:

$$u \sim N_n\left(0, \sigma^2 \boldsymbol{I}_n\right) \tag{1.3}$$

Inoltre i b non sono matematicamente indipendenti, ma verificano la relazione (1.2) in cui

$\boldsymbol{y}$= Vettore non stocastico con *dim(**r**)=p*
$\boldsymbol{R}$= Matrice non stocastica di ordine *(p,m)* con *p>m* e con *ran(**R**)=m*

La disuguaglianza (1.2) indica che ogni elemento di $\boldsymbol{R}\beta$ è almeno uguale al corrispondente elemento di $\boldsymbol{r}$. Una analoga convenzione varrà per tutti i segni di disuguaglianza che coinvolgono *array* nel corso del lavoro.

In Theil-Goldberger (1961) e Theil (1963) viene proposto l'uso delle (1.2) come informazioni a priori e tracciato lo schema della procedura mista di stima del modello di regressione lineare. Il metodo prevede l'inclusione delle relazioni già note direttamente nella procedura di stima, come parte integrante dell'insieme di ipotesi su cui il modello si regge. Ai dati di tipo statistico (pe esempio risultati di una stima precedente) vengono però uniformati quelli matematici (per esempio condizioni teoriche sui segni dei parametri) che di per se sono privi di una dimensione casuale. Un approccio alternativo è quello di considerare le (1.2) come una limitazione dello spazio parametrico e di procedere quindi con l'ottimizzazione solo nel sottoinsieme di E^m da esse circoscritto. L'obiettivo di questo lavoro, seguendo una impostazione analoga di Efron-Morris (1972) è di ricomporre la procedura di stima mista e quella vincolata in un formulazione bayesiana comune e di attuare delle simulazioni per cercare soprattutto una risposta alle seguenti domande:

1) Quale delle due procedure è maggiormente idonea al trattamento dei vincoli?
2) Si ottiene un significativo miglioramento nell'errore quadratico medio rispetto alla tecnica ordinaria dei minimi quadrati?

Il termine di paragone per le due procedure sarà l'Errore Quadratico Medio (EQM)

$$EQM(\boldsymbol{b}) = E\left\{\left(\boldsymbol{b} - \beta^0\right)^t \left(\boldsymbol{b} - \beta^0\right)\right\} \tag{1.4}$$

dove $\beta°$ è il valore vero dei parametri, $\boldsymbol{b}$ è lo stimatore di volta in volta considerato. Nel paragrafo 2 viene esposta la logica che produce lo stimatore misto, nel paragrafo 3 quella che porta allo stimatore vincolato. Nel paragrafo 4 viene presentato il modello di prova ed il piano di simulazione; infine, nel paragrafo 5 si riporteranno risultati e commenti.

2. La procedura di stima mista

Secondo questo approccio i vincoli (1.2) vengono rappresentati nella forma

$$\boldsymbol{r} = \boldsymbol{R}\beta + \boldsymbol{v} \tag{2.1}$$

con l'assunzione

$$\boldsymbol{v} \sim \boldsymbol{N}\left(\boldsymbol{0}, \lambda^2 \boldsymbol{I}_p\right) \tag{2.2}$$

Le informazioni a questo punto disponibili su β e su λ possono essere condensate nella relazione

$$P\left(\beta, \lambda | \boldsymbol{r}, \boldsymbol{R}\right) \propto \left(\frac{1}{\lambda^p}\right) e^{-\frac{1}{2\lambda^2}\left[v_1 c_1^2 + (\beta - \dot{\beta})^t \left(\boldsymbol{R}^t \boldsymbol{R}\right)(\beta - \dot{\beta})\right]}, \; c_1^2 = \left(\boldsymbol{r} - \boldsymbol{R}\dot{\beta}\right)^t \left(\boldsymbol{r} - \boldsymbol{R}\dot{\beta}\right), \; v_1 = p - m \tag{2.3}$$

dove

$$\dot{\beta} = \left(\boldsymbol{R}^t \boldsymbol{R}\right)^{-1} \boldsymbol{R}\boldsymbol{r} \tag{2.4}$$

Integrando la (2.3) rispetto a λ la distribuzione a priori dei β diventa

$$P\left(\beta | \boldsymbol{r}, \boldsymbol{R}\right) \propto \left[v_1 c_1^2 + \left(\beta - \dot{\beta}\right)^t \left(\boldsymbol{R}^t \boldsymbol{R}\right)\left(\beta - \dot{\beta}\right)\right]^{-\frac{p}{2}} \tag{2.5}$$

cioè una *t* di Student m-variata con media $\dot{\beta}$. La distribuzione (2.3) viene ora combinata con la funzione di verosimiglianza campionaria cioè

$$P\left(y | \beta, \sigma, \boldsymbol{X}\right) \propto e^{-\frac{1}{2\sigma^2}\left[v_2 c_2^2 + (\beta - \hat{\beta})^t \left(\boldsymbol{X}^t \boldsymbol{X}\right)(\beta - \hat{\beta})\right]}, \; c_2^2 = \left(\boldsymbol{y} - \boldsymbol{X}\hat{\beta}\right)^t \left(\boldsymbol{y} - \boldsymbol{X}\hat{\beta}\right), \; v_1 = n - m \tag{2.6}$$

e poi

$$\hat{\beta} = \left(\boldsymbol{X}^t \boldsymbol{X}\right)^{-1} \boldsymbol{X}^t \boldsymbol{y} \tag{2.7}$$

con $\hat{\beta}$ stimatore ai minimi quadrati ordinari. La distribuzione a posteriori di β e λ sarà

$$P\left(\beta, \lambda | y\right) \propto \left(\frac{1}{\lambda^{p+1} \sigma^n}\right) e^{-\frac{1}{2}\left\{\frac{1}{\lambda^2}\left[v_1 c_1^2 + (\beta - \dot{\beta})^t \left(\boldsymbol{R}^t \boldsymbol{R}\right)(\beta - \dot{\beta})\right] + \frac{1}{2\sigma^2}\left[v_2 c_2^2 + (\beta - \hat{\beta})^t \left(\boldsymbol{X}^t \boldsymbol{X}\right)(\beta - \hat{\beta})\right]\right\}} \tag{2.8}$$

La procedura di stima mista si concreta nel fissare ad un livello determinato il rapporto

$$k = \left(\frac{\sigma}{\lambda} \right)^2 \tag{2.9}$$

La (2.8) viene riscritta come

$$P(\beta, \lambda | y) \propto \left(\frac{1}{\sigma^{n+p+1}} \right) e^{-\frac{1}{2}\left\{ \frac{1}{\sigma^2}\left[c^2 + (\beta - \bar{\beta})^t \left[\left(\boldsymbol{X}^t \boldsymbol{X} \right) + k\left(\boldsymbol{R}^t \boldsymbol{R} \right) \right] (\beta - \bar{\beta}) \right] \right\}} \tag{2.10}$$

dove

$$c^2 = k\nu_1 c_1^2 + k\dot{\beta}^t \left(\boldsymbol{R}^t \boldsymbol{R} \right)\dot{\beta} - \bar{\beta}^t \left[\left(\boldsymbol{X}^t \boldsymbol{X} \right) + k\left(\boldsymbol{R}^t \boldsymbol{R} \right) \right] \bar{\beta}$$

ed inoltre

$$\bar{\beta} = \left[\left(\boldsymbol{X}^t \boldsymbol{X} \right) + k\left(\boldsymbol{R}^t \boldsymbol{R} \right) \right]^{-1} \left[\left(\boldsymbol{X}^t \boldsymbol{X} \right)\hat{\beta} + k\left(\boldsymbol{R}^t \boldsymbol{R} \right)\dot{\beta} \right] \tag{2.11}$$

Se si integra la (2.10) rispetto a σ si ottiene la distribuzione a posteriori di β

$$P(\beta | y) \propto \left\{ c^2 + (\beta - \bar{\beta})^t \left[\left(\boldsymbol{X}^t \boldsymbol{X} \right) + k\left(\boldsymbol{R}^t \boldsymbol{R} \right) \right] (\beta - \bar{\beta}) \right\}^{-\frac{n+p}{2}} \tag{2.12}$$

che è ancora in forma di una *t* di Student m-variata con media $\bar{\beta}$.

La moda della distribuzione a posteriori (2.12) è la stima puntuale che sostituisce la stima usuale ottenuta con il metodo dei minimi quadrati. Ad un risultato molto vicino a quello esposto arrivano Saxena (1984) e Lindley- Smith (1972) muovendo dall'idea definettiana di interscambiabilità delle funzioni di distribuzione. In questo senso, vedasi anche Zellner (1971).

Lo stimatore (2.11) ha avuto considerevole fortuna nella letteratura statistica (cfr. Goldstein e Smith, 1974; Mayer,1973; Lowerre,1974;Swindel ,1976; Farebro ther,1977. Lo stimatore $\bar{\beta}$ condivide la stessa formula generale con gli stimatori ridge e gli stimatori di Stein. Per meglio evidenziarne le caratteristiche lo si può riesporre nelle due espressioni equivalenti

$$\begin{aligned} \bar{\beta} &= \dot{\beta} + \left[\boldsymbol{I} + k\left(\boldsymbol{X}^t \boldsymbol{X} \right)^{-1} \left(\boldsymbol{R}^t \boldsymbol{R} \right) \right]^{-1} \left(\hat{\beta} - \dot{\beta} \right) \\ \bar{\beta} &= \hat{\beta} + \left[\boldsymbol{I} + \frac{1}{k}\left(\boldsymbol{R}^t \boldsymbol{R} \right)^{-1} \left(\boldsymbol{X}^t \boldsymbol{X} \right) \right]^{-1} \left(\dot{\beta} - \hat{\beta} \right) \end{aligned} \tag{2.13}$$

usando la relazione $(I + G)^{-1} = I - (I + G^{-1})^{-1}$ quando esista la G^{-}l. Quando il rapporto *k* tende a zero, cioè quando le informazioni a priori risultano vaghe rispetto a quelle campionarie, dalla prima delle (2.13) si vede che lo stimatore misto tende a quello ordinario (2.7). Dalla seconda delle (2.13) si vede peraltro come all'aumentare di *k*, indicando maggiore dispersione nel campione rispetto alle convinzioni ex-ante, lo stimatore $\bar{\beta}$ tende a concentrarsi sulla aspettativa a priori $\dot{\beta}$. La definizione del parametro *k* è quindi cruciale nello sviluppo della procedura mista e deve essere orientata al fine di garantire una equilibrata fusione di informazioni a priori ed informazioni campionarie. Una possibile tecnica di definizione di *k* potrebbe essere quella di stimare il numeratore a mezzo della varianza campionaria

$$\hat{\sigma}^2 = \frac{\|\boldsymbol{y} - \boldsymbol{X}\beta\|}{v_2} \tag{2.13}$$

Meno limpida la stima del denominatore che, spingendo al limite la interpretazione stocastica dei vincoli viene stimato, in analogia alla (2.14) con

$$\hat{\lambda}^2 = \frac{\|\boldsymbol{r} - \boldsymbol{R}\beta\|}{v_1} \tag{2.15}$$

Le proprietà statistiche, dal punto di vista campionario, dello stimatore (2.11) sono state approfondite in Nagar e Kakwani (1964).

3. La procedura di stima vincolata

La prospettiva di questa procedura coglie i vincoli (1.2) come condizioni analiti che del tutto prive di peculiarità stocastiche: essi definiscono un sottoinsieme di E^m

$$B = \left\{\beta \in E^m \middle| \boldsymbol{R}\beta \geq \boldsymbol{r}\right\} \tag{3.1}$$

che costituisce l'ambito di ammissibilità per i valori di β. Sui parametri e sulla varianza dei residui non si hanno informazioni, se non che, a priori, β e σ sono indipendenti e distribuiti in modo uniforme

$$P(\beta,\sigma) \propto \frac{1}{\sigma}; \quad \sigma > 0, \ -\infty < \beta < \infty \tag{3.2}$$

la formula bayesiana

$$P(\beta|y) \propto P(\beta)P(\boldsymbol{y}|\beta) \tag{3.3}$$

è invariante rispetto a restrizioni dello spazio parametrico (Hartigan, 1964): la distribuzione a posteriori basata su $P(y|\beta)$ definita per $\beta \in B$ sarà proporzionale alla $P(\beta\,|\,\boldsymbol{y})$ ristretta su *B*. In pratica, l'invarianza implica che, se si parte da una distribuzione a priori come la 3.2, ad informazione minima, si otterrà la stessa distribuzione a posteriori sia che si operi sulla funzione di verosimiglianza costruita solo su B, sia che si operi sulla funzione di verosimiglianza costruita su E^m forzando poi quest'ultima a zero nei punti esterni a *B*. (Waterman, 1974). Una impostazione basata su queste considerazioni è stata adoperata in O'Hagan (1973) per l'adattamento di una quadratica vincolata alla convessità.

Combinando la (3.2) con la funzione di verosimiglianza (2.6) si otterrà

$$P(\beta,\sigma|\boldsymbol{y},\boldsymbol{X}) \propto \left(\frac{1}{\sigma^{n+1}}\right) e^{-\frac{1}{2\sigma^2}\left\{\boldsymbol{y}^t\boldsymbol{M}\boldsymbol{y}+(\beta-\hat{\beta})^t(\boldsymbol{X}^t\boldsymbol{X})(\beta-\hat{\beta})\right\}}; \quad M = I - \boldsymbol{X}\left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t \qquad (3.4)$$

Integrando la (3.4) rispetto a σ si definisce la distribuzione a posteriori dei β

$$P(\beta|\boldsymbol{y},\boldsymbol{X}) \propto \left[\boldsymbol{y}^t\boldsymbol{M}\boldsymbol{y} + (\beta-\hat{\beta})^t(\boldsymbol{X}^t\boldsymbol{X})(\beta-\hat{\beta})\right]^{-\frac{n}{2}} \qquad (3.5)$$

che è ancora in forma di una t di Student m-variata con variabili troncate. Per calcolare il valore della stima vincolata bisogna cercare quel punto $\beta \in B$ che rende massima la (3.5) e quindi risolvere il problema di programmazione quadratica:

$$\underset{\beta\in B}{Min}\left\{(\beta-\hat{\beta})^t(\boldsymbol{X}^t\boldsymbol{X})(\beta-\hat{\beta})\right\} \text{ soggetto a } \quad \boldsymbol{R}\beta \geq \boldsymbol{r} \qquad (3.6)$$

Se $\hat{\beta} \in B$ allora la stima vincolata $\tilde{\beta}$ coincide con la stima ordinaria, se invece quaIcuno dei vincoli risulta effettivo, $\tilde{\beta}$ sarà quel valore $\beta \in B$ che ha distanza minima da $\hat{\beta}$ secondo la metrica $(\boldsymbol{X}^t\,\boldsymbol{X})$.

Supponiamo che $\hat{\beta}$ soddisfi p_1, con $p_1 < p$, vincoli e violi gli altri $p_2 = p - p_1$. Avremo $\boldsymbol{R_1}\hat{\beta} \geq \boldsymbol{r_1}; \quad \boldsymbol{R_2}\hat{\beta} < \boldsymbol{r_2}$ (3.7)

dove $(\boldsymbol{R_1}, \boldsymbol{R_2})$ ***e*** $(\boldsymbol{r_1}, \boldsymbol{r_2})$ sono adeguate partizioni di $\boldsymbol{R}$ ed $\boldsymbol{r}$ rispettivamente. L'esame delle condizioni di Kuhn-Tucker condotta in Sposito (1975), per l'ottimo del problema (3.6) porta a definire $\tilde{\beta}$ come

$$\tilde{\beta} = \begin{cases} \hat{\beta} \quad \text{se } \boldsymbol{R}\beta \geq \boldsymbol{r} \\ \hat{\beta} + (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{R_2^t}\left[\boldsymbol{R_2}(\boldsymbol{X}^t\boldsymbol{X})\boldsymbol{R_2^t}\right]^{-1}(\boldsymbol{r_2} - \boldsymbol{R_2}\hat{\beta}) \end{cases} \qquad (3.8)$$

se valgono le (3.7) e purché si abbia pure

$$\begin{cases} \boldsymbol{R_1}(\boldsymbol{X}^t\boldsymbol{X})\boldsymbol{R_1^t} \geq 0 \\ \boldsymbol{R_2^t}\left[\boldsymbol{R_2}(\boldsymbol{X}^t\boldsymbol{X})\boldsymbol{R_2^t}\right]^{-1}(\boldsymbol{r_2} - \boldsymbol{R_2}\hat{\beta}) \geq 0 \end{cases} \qquad (3.9)$$

La (3.8) e la (3.9) possono essere riassunte nel dire che se $\hat{\beta}$ viola qualche vincolo e si verificano le (3.9) i vincoli violati possono essere considerati uguaglianze e la soluzione esplicita è in tal caso data dalla seconda delle (3.8). Le condizioni (3.9) sono sufficienti ad assicurare l'ammissibilità della (3.8) e quin- secondo le condizioni di Kuhn e Tucker, la sua ottimalità. In realtà a fronte della (3.7) è certo solo che almeno uno dei vincoli violati sarà soddisfatto in forma di uguaglianza nella soluzione ottima. Questo non necessariamente implica (come sembrano indicare le (3.8) che tutti i vincoli saranno completamente esauriti da $\tilde{\beta}$, come non esclude che i vincoli attualmente verificati con residuo da $\hat{\beta}$ non si ritrovino poi in eguaglianze nel punto di ottimo. (Boot ,1964).

Se le (3.9) reggono non è necessario ricorrere ad algoritmi specifici per il calcolo dei $\tilde{\beta}$. Se tali condizioni non ci sono, l'espressione dei $\tilde{\beta}$ non è più così nitida ed anche se non si discosta molto dalla (3.8) (Liew , 1976) resta legata al problema particolare tanto da non essere più utilizzabile come termine di paragone esplicito con quella di altri stimatori.

Lo studio campionario dello stimatore b viene completato in Judge e Takayama (1966), e, in un contesto più generale, in Malinvaud (1966).

4. Configurazione del modello di prova

Le relazioni sulle quali si è incentrato il confronto sono

$$y_i = 3.0X_{i1} + 2.3X_{i2} + 1.6X_{i3} + 0.8X_{i4} + 0.1X_{i5}; \quad i = 1,2,\ldots,n \tag{4.1}$$

$$\beta_j \geq 0 \quad (j = 1,2,\ldots,5) \tag{4.2}$$

$$\beta_j - \beta_{j+1} \geq 0.5 \quad (j = 1,2,\ldots,4) \tag{4.3}$$

I vincoli, per questo problema, possono essere riassunti nella forma matriciale

$$\begin{bmatrix} \boldsymbol{0}_5 \\ 0.5\boldsymbol{U}_4 \end{bmatrix} \leq \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 & \boldsymbol{e}_4 & \boldsymbol{e}_5 \\ \boldsymbol{e}_1 & \boldsymbol{e}_1 - \boldsymbol{e}_2 & \boldsymbol{e}_2 - \boldsymbol{e}_3 & \boldsymbol{e}_3 - \boldsymbol{e}_4 & \boldsymbol{e}_4 - \boldsymbol{e}_2 \end{bmatrix} \beta \tag{4.4}$$

dove $\boldsymbol{e}_i$ è un vettore identità con un uno in posizione i-esima e per il resto zero; $\boldsymbol{0}_5$ è un vettore nullo di dimensione 5, $\boldsymbol{U}_4$ è un vettore somma composto da soli uno e di dimensione 4. I valori veri nella (4.1) discendono dalla considerazione che l'effetto delle restrizioni potrebbe essere diverso secondo la distanza dei β dal loro rispettivo contorno dell'insieme di ammissibilità. La (4.3) oltre ad imporre una relazione d'ordine fra i parametri, permette di eludere problemi di multicollinearità che potrebbero essere presenti nella matrice dei regressori. Le $\boldsymbol{X}_j$ infatti vengono simulate da distribuzioni uniformi indipendenti il cui contributo alla variabilità della $\boldsymbol{y}$ è misurato da

$$\beta_j^2 var\left(X_j\right) = 1 \tag{4.5}$$

tutte le $\boldsymbol{X}_j$ avranno dunque la stessa importanza nella composizione della variabile dipendente, pur muovendo da distribuzioni diverse. In particolare

$$X_1 \sim U(0,1.1547);\ X_2 \sim U(0,1.5061);\ X_3 \sim U(0,2.1551)$$
$$X_4 \sim U(0,4.3301);\ X_5 \sim U(0,34.641) \tag{4.6}$$

Fissata la matrice dei regressori, i fattori guida della sperimentazione sono l'ampiezza campionaria n e la varianza dei residui σ^2. Per la prima si è prevista una variazione nell'insieme

$$n \in N = \{30,\ 60,\ 200\} \tag{4.7}$$

in cui, i primi due valori, sono tipiche ampiezze campionarie incontrate di frequente nei lavori econometrici e negli studi di simulazione. Il terzo elemento di N darà una certa tonalità asintotica alla sperimentazione.

La varianza dei residui è stata definita come una percentuale a della variabilità della y, per cui

$$\sigma^2 = \frac{m\alpha}{1-\alpha} \tag{4.8}$$

e per a è stata prevista una rotazione nell'insieme

$$\alpha \in A = \{0.001,\ 0.05,\ 0.20,\ 0.50\} \tag{4.9}$$

i cui elementi sono motivati dalla esigenza di saggiare le procedure di stima secondo una certa gamma di coerenza fra il modello a cui si applicano ed i dati analizzati.

Ogni elemento del prodotto cartesiano $N \otimes A$ propone una situazione sperimentale e per ognuna di queste 12 occasioni saranno effettuate $L=1000$ repliche al fine di stabilizzare i risultati. Il confronto degli stimatori verrà realizzato attraverso le misure suggerite in Spitzer (1978) e cioè

$$\text{Valore Asintotico Medio (VAM): } (l/L) \sum \boldsymbol{b} \tag{410}$$

$$\text{Errore Quadratico Medio (EQM): } (l/L) \sum || \boldsymbol{b} - \beta° ||^2 \tag{411}$$

dove L è il numero di repliche (1000) condotte, $\beta°$ è il valore vero dei parametri e $\boldsymbol{b}$ è il valore corrente prodotto da una procedura di stima. Lo schema delineato è quello più usuale adoperato nei confronti di stimatori (cfr. Dempster *et al*, 1977; Gunst e Mason, 1977; Lawless e Wang, 1976; McDonald e Galarneau, 1973). Dettagli maggiori sulla simulazione e sul calcolo degli stimatori vengono dati nell'appendice al lavoro.

5 Risultati e commenti

L'analisi dei risultati indica comportamenti perlopiù attesi, concordi con lo sviluppo teorico condotto nei parte 2 e 3. In particolare viene messa in risalto la scarsa incisività delle informazioni a priori quando il modello è coerente ai dati; si evidenzia inoltre la diversità del fine raggiunto quando i dati ex-ante vengono inseriti nel modello secondo l'una o l'altra interpretazione.

Le tabelle 1 e 2 riportano gli esiti delle 12 configurazioni di stima per la procedura ordinaria (MQO), per quella mista (MQV) e per quella vincolata (MQV). I risultati sono stati affiancati in modo da far emergere il diverso contegno delle procedure nei due loro momenti più significativi: Valore Asintotico Medio (VAM) ed Errore Quadratico Modio (EQM).

5.1 Valore asintotico medio

Per i valori più piccoli di *a* il VAM si avvicina al valore vero all'aumentare dell'ampiezza campionaria per tutte le tre procedure. L'adesione è più celere per gli stimatori misti rispetto a quelli ordinari e di questi rispetto a quelli vincolati. Per *n = 200* le differenze fra i tre approcci sono poco significative. Evidentemente, laddove la componente stocastica sia debole e si disponga di un numero appena sufficiente di osservazioni, tutte le procedure anche solo plausibili danno risultati soddisfacenti. Per i valori più grandi di α si realizza la stessa tendenza di cui al punto precedente, tranne che lo stimatore ordinario mantiene la distorsione entro limiti ridotti anche per i valori piccoli di *n*. Lo stimatore misto viene prima contratto sul suo valore a priori per espandersi poi sotto l'effetto dell'aumento di *n* Tale effetto è pero diversificato secondo la magnitudo dei parametri: più tenue per i valori piccoli, più intenso per quelli grandi.

Lo stimatore vincolato risulta poco affidabile quando *n* è piccolo dando luogo a valori medi di stima lontani da quelli attesi. La qualità migliora quando *n* aumenta. Anche questo miglioramento risulta selettivo secondo l'ordine di grandezza dei parametri analogamente a quando succede per lo stimatore misto. Per *n = 30* i V AM si allontanano dal valore atteso quando α aumenta. Gli scostamenti sono poco marcati per gli MQO; gli MQM collassano sulla media a priori; gli MQV oscillano senza alcuna sistematicità. Per *n = 200* la tendenza di fondo non cambia, si hanno pero deviazioni più lievi per gli MQO, collassamento più lento per gli MQM od oscillazioni maggiormente ristrette per gli MQV.

5.2. Errore Quadratico Medio

Quando *a* è piccolo, indipendentemente da *n*, l'EQM degli MQO e degli MQM è quasi sempre inferiore a quello degli MQV. La differenza fra MQO ed MQM è poca cosa e tale rimane anche per *a= 0,05*. Per i paramentri b_2 e b_3 gli MQM risultano leggermente più precisi degli ordinari che, a loro volta, hanno aggio sugli altri parametri. Gli stimatori vincolati, a questi livelli di *a* presentano EQM sempre più alti degli MQO, raramente più bassi degli MQM. Per i valori più grandi di *a*, gli stimatori vincolati ottengono EQM inferiori a quelli misti, segnatamente per i parametri b_2 e b_3 con esclusione del parametro b_5 . Quando a= 0.5 anche gli stimatori ordinari hanno EQM tendenzialmente superiori a quelli vincolati ad eccezione del b_5 (troppo spesso, forse, posto pari a zero dell'algoritmo di programmazione quadratica). All'aumentare dell'ampiezza campionaria l'EQM diminuisce per gli MQO e per gli MQM, ma più rapidamente per i secondi che per i primi. Gli MQV non denotano tale tendenza rivelando forse una distorsione non eliminabile dal ripetersi delle prove. Il momento migliore degli stimatori vincolati è nei valori alti di *a* e piccoli di *n*. Per *a=0.50* ed *n=30* ottengono l'EQM più basso per quasi tutti i parametri.

Tab.1 -*Valore asintotico medio*

n	α		β_1	β_2	β_3	β_4	β_5
30	0.001	MQO	3.0023	2.2992	1.5997	0.8001	0.1000
		MQM	2.9861	2.2971	1.6020	0.8015	0.1002
		MQV	2.8700	2.2705	1.5750	0.8090	0.1049
	0.050	MQO	2.9990	2.3060	1.6013	0.8003	0.1000
		MQM	2.5740	2.1470	1.6026	0.8640	0.1094
		MQV	2.5268	2.1388	1.7269	0.9010	0.1105
	0.200	MQO	2.9918	2.1325	1.5957	0.8054	0.0995
		MQM	1.9366	1.7444	1.4917	0.9590	0.1354
		MQV	2.4711	2.0156	1.3742	0.6734	0.1063
	0.500	MQO	3.0238	2.2883	1.5829	0.7882	0.1031
		MQM	1.2874	1.1818	1.1454	0.8420	0.1063
		MQV	2.6845	2.4345	1.3112	1.0312	0.0931
60	0.001	MQO	3.0032	2.3029	1.6013	0.8008	0.1001
		MQM	2.9980	2.2977	1.6002	0.8008	0.1001
		MQV	2.9410	2.2834	1.5952	0.8014	0.1021
	0.050	MQO	3.0003	2.3001	1.6025	0.8009	0.1002
		MQM	2.7716	2.2343	1.6131	0.8278	0.1047
		MQV	2.8717	2.2996	1.6190	0.8301	0.0989
	0.200	MQO	3.0115	2.3228	1.5942	0.8029	0.0995
		MQM	2.2793	1.9781	1.5774	0.9055	0.1202
		MQV	2.7702	2.4893	1.5070	0.8324	0.1082
	0.500	MQO	2.9842	2.2914	1.5876	0.7965	0.1022
		MQM	1.6249	1.4911	1.3676	0.9741	0.1552
		MQV	2.3840	2.1340	1.8840	0.9463	0.0949
200	0.001	MQO	3.0025	2.3025	1.6016	0.8008	0.1001
		MQM	2.9980	2.3000	1.5996	0.8003	0.1000
		MQV	2.9695	2.2979	1.6022	0.8057	0.1003
		MQO	3.0061	2.3006	1.5974	0.8020	0.1001
		MQM	2.9243	2.2817	1.6047	0.8084	0.1015
		MQV	3.0524	2.2438	1.6149	0.8183	0.0992
	0.200	MQO	3.0143	2.2995	1.5947	0.8024	0.1002
		MQM	2.7075	2.2063	1.6045	0.8420	0.1063
		MQV	2.8866	2.2387	1.6245	0.7889	0.1063
	0.500	MQO	3.0147	2.2908	1.5935	0.8082	0.0987
		MQM	2.2144	1.9372	1.5686	0.9184	0.1214
		MQV	2.8157	2.3272	1.6108	0.8697	0.0916

Tab.2 - *Errore Quadratico medio*

n	α		β_1	β_2	β_3	β_4	β_5
30	0.001	MQO	0.0013	0.0008	0.0005	0.0001	0.0000
		MQM	0.0018	0.0008	0.0004	0.0001	0.0000
		MQV	0.3261	0.2213	0.0314	0.0847	0.6323
	0.05	MQO	0.0732	0.0392	0.0258	0.0039	0.0002
		MQM	0.2311	0.0465	0.0154	0.0093	0.0002
		MQV	0.0795	0.1157	0.1244	0.0423	0.6235
	0.2	MQO	0.3603	0.1848	0.0999	0.0171	0.0007
		MQM	1.2098	0.3575	0.0444	0.0423	0.0017
		MQV	1.3868	0.1021	0.0248	0.2644	0.6277
	0.5	MQO	1.4254	0.6972	0.4394	0.0705	0.0027
		MQM	1.9905	1.3014	0.2499	0.0535	0.0088
		MQV	0.2111	0.4659	0.0572	0.0683	0.6258
60	0.001	MQO	0.0007	0.0004	0.0002	0.0000	0.0000
		MQM	0.0007	0.0004	0.0002	0.0000	0.0000
		MQV	0.4126	0.2337	0.0383	0.0892	0.6367
	0.05	MQO	0.0393	0.0211	0.0126	0.0022	0.0000
		MQM	0.0780	0.0192	0.0084	0.0034	0.0001
		MQV	0.3268	0.2499	0.0594	0.0747	0.6418
	0.2	MQO	0.1735	0.0978	0.0567	0.0103	0.0002
		MQM	0.5713	0.1358	0.0231	0.0212	0.0006
		MQV	0.2758	0.5452	0.0169	0.0936	0.6270
	0.5	MQO	0.7852	0.3429	0.2045	0.0452	0.0007
		MQM	1.9415	0.6950	0.0838	0.0508	0.0037
		MQV	0.1003	0.2058	0.3275	0.0773	0.6487
200	0.001	MQO	0.0002	0.0001	0.0001	0.0000	0.0000
		MQM	0.0002	0.0001	0.0001	0.0000	0.0000
		MQV	0.4483	0.2480	0.0409	0.0866	0.6396
	0.05	MQO	0.0098	0.0054	0.0034	0.0006	0.0000
		MQM	0.0143	0.0052	0.0028	0.0008	0.0000
		MQV	0.5654	0.2039	0.0467	0.0797	0.6413
	0.2	MQO	0.0463	0.0257	0.0144	0.0031	0.0000
		MQM	0.1138	0.0255	0.0103	0.0049	0.0001
		MQV	0.3464	0.3147	0.0506	0.0974	0.6302
	0.5	MQO	0.1834	0.1131	0.0616	0.0119	0.0002
		MQM	0.6594	0.1604	0.0231	0.0232	0.0006
		MQV	0.2835	0.2892	0.0466	0.0593	0.6538

Appendice

I calcoli riportati in questo articolo sono stati effettuati sul VAX 11/780 dell'Uni versità di Calabria. Si è sempre utilizzata la doppia precisione aritmetica ed i risultati sono accurati nelle cifre riportate. L'autore ha adattato le routine riportale in Land e Powell (1975) per la soluzione del problema (3.6) e la messa in opera del programma di Douglas Stirling (1981) per il calcolo degli stimaltri misti Ha inoltre curato il raccordo dei programmi e la redazione di quelli non dispombili in letteratura.

Le variabili indipendenti $(X_j, j = 1,2,,5)$ sono state simulate usando l'algoritmo proposto in Wichmann e Hill(1982) con valori di partenza *IX = 13367, IY = 23311, IZ = 26317* per acquisire numeri pseudocasuali aull'intervallo unitario, riportati poi, con una trasformazione lineare, sull'intervallo fissato. La matrice così ottenuta è rimasta poi inalterata nel corso di tutte le sperimentazioni

Il generatore di numeri casuali già citato è stato pure richiamato per la simulazione dei residui del modello (4.1)per i quali si è adottata la trasformazione "polare" descritta in Alkinson e Pierce (1976) per ottenere valori da una distribuzione nonnale

Riferimenti bibliografici

Atkinson A.C., Pierce M.C. (1976). The Computer Generation of Beta, Gamma and Normal Random Variable, *Journal of the Rojal Statistical Society, A,* 139, 431-461

Boot J. (1964). Ouadratic Prog,amming. Rand-MeNall, Chicago

Dempster A.P. - Shatzcoff M. - Wermuth N(I977). A Simulation Study of Alternatives to Ordinary Least Squaes, *Journal of the American Statistical Association,* 72, 77-90

Douglas Sterling W. (1981). Algorithm AS 164: Least Squares Subject to Linear Constraints, *Applied Statistics*, 30, 204-210

Douglas Sterling W. (1981). Algorithm AS 164: a correction. *Applied Statistics,* 30, 357

Efron B. - Morris C. (1972). Empirical Bayes on Vector Observations: An Extension of *Association*", 65, 913-925.

Stein's Method. *Biometrika,* 335-347

Farebrother R. W. (1977). A Class of Shrinkage Estimators. *Journal of the Royal Statistical Society, B*, 40, 47-49

Goldstein M. Smith A.F.M. (1974). Ridge-Type Estimators for Regression Analysis. *Journal of the Royal Statislical Society,* B, 36,284-291

Gunst R.F. - Mason R.L. (1977) Biased Estimation in Regression: An Evaluation Using Mean Square Error. *Journal of the American Statistical Association*,72, 616-628

Hartigan J. (1964). lnvariant Prior Distributions. *Annals of Mathematical Statistics,* 35, 836-845

Judge G.G. - Takajama T. (l966). lnrquality Restrictions in Regression Analysis. *Journal of the Amencan Statistical Association*, 61,166-181

Land A. Powell S. (1975). Fortran Codes for Mathematical Programming. John Wiley and Sons, New York

Lawless I.F. - Wang P. (1976). A Simulation Study of Ridge and Other Regression Estimators. *Communications in Statistics, Theory and Methods, A*,10, 307-323

Liew C.K. (1976). Inequality Constrained Least Squares Estimation. *Journal of the American Statistical Association*,71,746-751

Lindley D.V. - Smith A.F.M. (1972). Bayes Estimates for the Linear ModeI. *Journal of the Royal Statistical Society, B*, 34,1-18

Lovell M.C. - Prescott B. (1970). Multiple Regression with Inequality Constraints: Pretesting Bias, Hypothesis Testing and Efficiency. *Journal of the American Statistical*

Lowerre J.M. (1974). On the Mean Square Error of Parameter Estimates for Some Biased Estimators, *Technometrics*, v. 16, pp. 461-467.

Malinvaud E. (1966). Statistical Methods of Econometrics. Rand-McNally, Chicago.

Mayer L.S,Wilke T.A.(1973). On Biased Estimation in Linear Models. *Technometrics*, 15,497-508.

McDonald G.C. - Galarnau D.I. (1973). A Monte Carlo Evaluation of Some Ridge-Type Estimators. *Journal of the American Statistical Association*,70, 407-416

Nagar A.L. - Kakwani N.C. (1964). The Bias and the Moment Matrix of a Mixed Regression Estimators. *Econometrica*, 32, 174-182.

O'Hagan A. (1973). Biased Estimation of a Convex Quadratic. *Biometrika*, 60, 565-571.

Rothenberg J. (1973). Efficient Estimation with a Piori Information Yale University Press, New Haven-London.

Saxena A.K. (1984). Bayesian Estimation of T-G Ridge Model. *Statistica Neerlandica*, 38,257-260.

Spitzere J.J. (1978). A Montecarlo Investigation of the Box-Cox Transformation in Small Samples. *Journal of the American Statistical Association*,73, 488-495.

Sposito V.D. (1975). Linear and Nonlinear Programming. Iowa State University Press. Ames.

Swindel B.F. (1976). Good Ridge Estimators Based on Priori Information. *Communications in Statistics-Theory and Methods, A*, 11,1065-1075.

Theil H. (1963). On the Use of Incomplete Prior Information in Regression Analysis, *Journal of the American Statistical Association*, 58, 401-414.

Goldberg A.S. (1961). On Pure and Mixed Statistical Estimation in Economics. *International Economic Review*, 2, 65-78.

Watermann M.S. (1974). A Restricted Least Squares Problem. *Technometrics*,16,135-136.

Wichmann B.A. - Hill I.D. (1982). An Efficient and Portable Pseudo-Random Number Generator, *Applied Statistics*, 31,188-190.

Zellner A. (1971). An Introduction to Bayesian Inference in Econometrics. John Wiley & Sons, New York