

# 9. Alcuni problemi di specificazione e stima di modelli a parametri variabili multiequazionali

di Carlo Carraro e Domenico Sartore

## 1. Introduzione

L'analisi dei problemi legati all'applicazione econometrica di modelli teorici la cui struttura parametrica non è costante nel tempo si è notevolmente sviluppata nel corso degli ultimi anni. Nel dibattito teorico che ne è conseguito, non sembra giustificata l'obiezione che una struttura dei parametri variabile debba essere considerata come un errore di specificazione del modello. Infatti, quanto più ampio è il periodo campionario di riferimento, tanto più elevata risulta la probabilità che alcuni parametri del modello non possano essere tenuti costanti perché su di essi si riflettono i cambiamenti strutturali dei fenomeni economici. Questo avviene tipicamente per le equazioni di un modello che esprimono delle funzioni di reazione a decisioni di politica economica. Come è evidente, cambiamenti di regime di politica economica non possono essere descritti mantenendo costanti i parametri in queste equazioni.

È ben vero che scopo rilevante dell'indagine econometrica rimane la ricerca dei parametri invarianti rispetto a tali cambiamenti, come suggerisce l'applicazione del concetto di *super-esogeneità* [Engle, Hendry e Richard 1983], tuttavia anche la conoscenza di come variano i parametri non costanti risulta utile per l'analisi strutturale del modello, a fini previsivi e di controllo. A ciò si aggiunga anche l'ampia quantità di modelli teorici con parametri non costanti proposti negli sviluppi più recenti della letteratura economica (ad esempio, in economia industriale, nella teoria del ciclo reale, nella teoria del consumo). L'econometria dei modelli a parametri variabili va quindi considerata uno strumento importante di verifica empirica, valutazione quantitativa e simulazione di numerosi modelli teorici.

Gli strumenti analitici più frequentemente utilizzati per la stima di modelli a struttura variabile sono il Filtro di Kalman e le sue estensioni (*information filter* e *smoothing*). L'utilizzo di tali strumenti pone tuttavia numerosi problemi al ricercatore, data la difficoltà di definire in modo appropriato le condizioni iniziali di stima, la struttu-

ra dinamica del modello e l'algoritmo di stima degli iperparametri, cioè della matrice di transizione, e delle matrici di covarianza. Ciò rende tale strumentazione di difficile adozione per l'economista applicato e ne limita l'utilizzo.

L'obiettivo del programma di ricerca sviluppato in questi anni presso l'Università di Venezia è stato proprio quello di facilitare l'applicazione del Filtro di Kalman ai problemi economici agendo in due direzioni: da un lato combinando alcuni algoritmi di stima in modo da rendere ottimale ed automatico il processo di identificazione dei parametri; dall'altro, rendendo più affidabile, dal punto di vista numerico, la procedura proposta da Kalman. I risultati teorici e numerici ottenuti sono stati presentati in precedenti lavori [si veda Carraro-Sartore 1984 e 1987; Carraro 1988; Sartore 1990]. La messa in opera di quei risultati ha condotto alla costruzione di un algoritmo, denominato *Square Root Iterative Filter, Information and Smoothing* (SRIFIS)<sup>1</sup>, la cui struttura viene discussa nel terzo paragrafo di questo lavoro, e alla realizzazione di un pacchetto *software*<sup>2</sup> che permette la stima del vettore di stato e degli iperparametri di modelli rappresentati nello spazio degli stati in modo efficiente e numericamente affidabile.

La trasformazione nello spazio degli stati di un modello multiequazionale con una struttura parametrica variabile nel tempo viene discussa nel paragrafo 2. Come verrà evidenziato, una delle caratteristiche di SRIFIS consiste nello stimare gli iperparametri in termini non condizionali rispetto alle stime del vettore di stato. Esperimenti di Monte-Carlo hanno mostrato la maggiore efficienza delle stime non condizionali, rispetto a quelle condizionali.

Obiettivo di questo articolo è il compimento di un altro passo nella direzione del programma di ricerca prima illustrato. Risolti i problemi di definizione delle condizioni iniziali sul vettore di stato, della stima degli iperparametri (si veda il paragrafo 4), della convergenza dell'algoritmo e della sua ottimizzazione numerica [Carraro-Sartore 1987], vanno evidenziate alcune difficoltà per la stima di modelli a parametri variabili, discusse nel paragrafo 5 [si veda anche Sartore 1990].

Nella stima di modelli a parametri variabili possono infatti sorgere due diversi problemi. Il primo proviene dal mancato rispetto di una *condizione di minima flessibilità strutturale*. Qualora il modello a parametri variabili non goda di questa condizione nella specificazione del modello, la procedura iterativa per la stima degli iperparametri

<sup>1</sup> Versioni precedenti dello stesso algoritmo erano denominate semplicemente SRIF.

<sup>2</sup> Il pacchetto è realizzato in Gauss con la pregevole collaborazione di Paolo Vidulich.

conduce ad un *modello singolare*, cioè ad un modello in cui diventa singolare la matrice di covarianza dei residui del modello. Si ha quindi un problema di *eccesso* di flessibilità strutturale.

Nel caso di un modello con una sola variabile dipendente è la varianza dei residui che si annulla come effetto della completa capacità di adattamento del modello ai dati osservati, si ha cioè *perfetta* flessibilità strutturale.

Come si vedrà nel paragrafo 5, per questi casi, l'algoritmo SREMS propone indicatori che suggeriscono il punto di arresto nella procedura iterativa, fondati essenzialmente sulla definizione di misure di distanza tra stime successive del vettore di stato.

Il secondo problema riguarda le strategie di assegnazione dei valori iniziali a priori degli iperparametri per rendere efficaci le procedure di stima. Valori iniziali devono essere fissati sia per il vettore di stato, sia per gli iperparametri. Il Filtro di Kalman fornisce stimatori robusti rispetto ai diversi valori iniziali del vettore di stato mentre risulta più sensibile ai valori iniziali assegnati agli iperparametri. Questo *eccesso di sensibilità* ai valori iniziali assegnati agli iperparametri è ancor più marcato per la stima della matrice di covarianza dello stato, rilevante per la definizione delle statistiche utilizzate nella verifica delle ipotesi sui parametri del modello. In questo lavoro si delinea una procedura di ricerca dei valori iniziali da assegnare agli iperparametri che sembra in grado di condurre a stime soddisfacenti della distribuzione statistica dei parametri del modello.

Va comunque notato che il problema della eccessiva flessibilità strutturale e quello della eccessiva sensibilità ai valori iniziali degli iperparametri sono tra loro collegati, e l'assenza del primo agevola la soluzione del secondo.

Nell'articolo non si è voluto proporre un paragrafo conclusivo. I risultati presentati hanno in alcuni casi natura preliminare in quanto evidenziano problemi di fondo legati alla natura dei modelli a parametri variabili; problemi in cui si intersecano difficoltà di identificazione, specificità degli algoritmi, molteplicità dei processi decisionali. L'evidenza simulativa sembra suggerire l'importanza dei temi in discussione e la validità delle soluzioni proposte: ulteriori analisi sono tuttavia in corso di elaborazione.

## 2. Il modello a parametri variabili multiequazionali: sua rappresentazione nello spazio degli stati

Il modello multivariato è rappresentato congiuntamente dalle equazioni delle variabili endogene,  $\mathbf{y}_t = [y_{1t}, \dots, y_{rt}]'$ , e dalle equazioni dei parametri variabili  $\boldsymbol{\beta}_t = [\beta_{1t}, \dots, \beta_{rt}]'$ , come segue:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{x}_t' \boldsymbol{\beta}_t + \mathbf{w}_t' \boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t \\ \boldsymbol{\beta}_{it} &= \mathbf{M}_i \boldsymbol{\beta}_{i,t-1} + \mathbf{v}_{it} \boldsymbol{\delta}_i + \mathbf{u}_{it}, \quad i = 1, \dots, r \end{aligned}$$

dove le quantità che entrano nell'equazione delle variabili endogene sono:

$\mathbf{x}_t = \text{diag} \{ \mathbf{x}_{1t}, \dots, \mathbf{x}_{it}, \dots, \mathbf{x}_{rt} \}$ , con  $\mathbf{x}_{it}$  vettori di ordine  $(k_i \times 1)$  di variabili esogene relative ai parametri variabili del modello;

$\mathbf{w}_t = \text{diag} \{ \mathbf{w}_{1t}, \dots, \mathbf{w}_{it}, \dots, \mathbf{w}_{rt} \}$ , con  $\mathbf{w}_{it}$  vettori di ordine  $(l_i \times 1)$  di variabili esogene relative ai parametri costanti  $\boldsymbol{\gamma}$ ;

i disturbi  $\boldsymbol{\varepsilon}_t \sim \text{NID}(\mathbf{0}, \mathbf{R})$ , con  $\mathbf{R}$  matrice definita positiva di ordine  $r$ ;

e le quantità che entrano nelle equazioni dei parametri variabili sono:

$\mathbf{M}_i$ ,  $i = 1, \dots, r$ , matrici di transizione dei parametri variabili di ordine  $(k_i \times k_i)$ ;

$\mathbf{v}_{it}$ ,  $i = 1, \dots, r$ , matrici di variabili esogene di ordine  $(k_i \times p_i)$ ;

i disturbi  $\mathbf{u}_{it} \sim \text{NID}(\mathbf{0}, \mathbf{Q}_i)$ , con  $\mathbf{Q}_i$ ,  $i = 1, \dots, r$ , matrici definite positive di ordine  $k_i$ .

Il modello a parametri variabili multivariato deve essere posto in forma di modello nello spazio degli stati per l'applicazione dell'algoritmo SRIFIS. La rappresentazione è la seguente:

$$\begin{aligned} \mathbf{y}_t &= \mathbf{H}_t \mathbf{z}_t + \boldsymbol{\varepsilon}_t \\ \mathbf{z}_t &= \mathbf{F}_t \mathbf{z}_{t-1} + \mathbf{G}_t \mathbf{u}_t \end{aligned}$$

dove  $\boldsymbol{\varepsilon}_t \sim \text{NID}(\mathbf{0}, \mathbf{R})$  sono i disturbi dell'equazione di output e  $\mathbf{u}_t \sim \text{NID}(\mathbf{0}, \mathbf{Q})$  rappresentano i disturbi dell'equazione di stato e si assume  $\mathbf{Q} = \text{diag}(\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_r)$ .

La trasposizione si ha ponendo:

$$[1] \quad \mathbf{z}_t = [\boldsymbol{\beta}_{1t} \ \boldsymbol{\delta}_{1t} \ \boldsymbol{\gamma}_{1t} \mid \boldsymbol{\beta}_{2t} \ \boldsymbol{\delta}_{2t} \ \boldsymbol{\gamma}_{2t} \mid \dots \mid \boldsymbol{\beta}_{rt} \ \boldsymbol{\delta}_{rt} \ \boldsymbol{\gamma}_{rt}]'$$

$$[2] \quad \mathbf{F}_t = \text{diag}(\mathbf{F}_{11t}, \mathbf{F}_{22t}, \dots, \mathbf{F}_{rrt}) \text{ con:}$$

$$\mathbf{F}_{iit} = \begin{bmatrix} \mathbf{M}_i & \mathbf{v}_{it}' & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{p_i} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{k_i} \end{bmatrix}, \quad i = 1, 2, \dots, r$$

$$[3] \quad \mathbf{G}_t = \begin{bmatrix} \mathbf{I}_{k_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I}_{k_2} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \hline \dots & \dots & \dots & \dots \\ \hline \mathbf{0} & \dots & \mathbf{0} & \mathbf{I}_{k_r} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$[4] \quad \mathbf{H}_t = \text{diag} (\mathbf{H}_{11t}, \mathbf{H}_{22t}, \dots, \mathbf{H}_{rrt}),$$

con  $\mathbf{H}_{iit} = [\mathbf{x}'_{it} \quad \mathbf{0}' \quad \mathbf{w}'_{it}]$ ,  $i = 1, 2, \dots, r$

La dimensione del vettore di stato  $\mathbf{z}_t$  è data da  $n = \sum_{i=1}^r (k_i + p_i + l_i)$ . Ponendo  $m = \sum_{i=1}^r k_i$ , la matrice  $\mathbf{H}_t$  risulta di ordine  $(r \times n)$ , la matrice  $\mathbf{F}_t$  di ordine  $(n \times n)$  e  $\mathbf{G}_t$  di ordine  $(n \times m)$ .

Si noti che nel vettore di stato sono presenti contemporaneamente sia parametri variabili che costanti del modello. Questa rappresentazione è molto generale, permette infatti di avere più parametri variabili e più parametri costanti per ogni equazione. Inoltre, attraverso semplici modifiche nella composizione delle matrici  $\mathbf{F}_t$  ed  $\mathbf{H}_t$  si possono ottenere:

a) *modelli a soli parametri costanti*, quando:

$$[a.1] \quad \mathbf{F}_t = \mathbf{I}_n$$

$$[a.2] \quad \mathbf{H}_t = \mathbf{w}'_t$$

$$[a.3] \quad \mathbf{G}_t = \mathbf{0};$$

b) *modelli a soli parametri variabili*<sup>3</sup>, se:

$$[b.1] \quad \mathbf{H}_{iit} = \mathbf{x}'_{it}$$

$$[b.2] \quad \mathbf{F}_{iit} = \mathbf{M}_i \quad i = 1, 2, \dots, r;$$

<sup>3</sup> Questi tipi di modelli sono discussi, ad esempio, in Chow [1983].

c) modelli a parametri variabili con equazioni a soli parametri costanti.

Supponendo ad esempio che la  $j$ -esima equazione del sistema contenga solo parametri costanti, allora all'interno delle matrici  $F$  e  $H$  definite in [2] e [4] si ha:

$$[c.1] \quad F_{jft} = I_{n_j}, \quad n_j = k_j + p_j + l_j$$

$$[c.2] \quad H_{jft} = w'_{jt}$$

$$[c.3] \quad G_{jft} = 0$$

### 3. L'algoritmo: «Square Root Iterative Filter, Information and Smoothing (SRIFIS)»

L'algoritmo, proposto in Carraro e Sartore [1987], considera due diversi metodi per stimare le matrici  $M$ ,  $R$  e  $Q$ . Il primo utilizza, per ogni iterazione, i residui dello *smoothing*; il secondo è basato sul processo di innovazione del filtro. In questo lavoro si farà riferimento al primo metodo per la maggiore semplicità computazionale. Questo requisito diventa essenziale qualora si voglia tradurre l'algoritmo su *personal computer*. Un'ulteriore motivazione è data dal fatto che i residui dello *smoothing*, a differenza delle innovazioni del filtro, sono determinati condizionatamente a tutta l'informazione campionaria. Nel caso di fenomeni economici, le osservazioni campionarie vengono analizzate quasi sempre in modo congiunto, raramente infatti si producono stime del vettore di stato  $z_{t/t-1}$  oppure  $z_{t/t}$ , sequenzialmente ad ogni osservazione pervenuta, per poi utilizzarle (ad esempio attraverso il controllo ottimale) nel condizionare lo stato reale dell'economia. A fini di analisi strutturale diventa invece rilevante l'uso dello *smoothing*, rispetto al filtro di Kalman, poiché ciò significa non rinunciare a utilizzare tutta l'informazione campionaria disponibile e quindi a produrre delle stime più accurate dei parametri del modello.

Si riportano qui di seguito i passi principali della procedura algoritmica dello SRIFIS:

*Passo 1:* Si scelgono i valori iniziali di  $M_0$ ,  $R_0$ ,  $Q_0$ , per le matrici di covarianza  $R$ ,  $Q$ , e per la matrice  $M$  che contiene gli elementi incogniti della matrice di transizione  $F$  e si inizializza il filtro di informazione con una matrice di covarianza dello stato a priori diffusa, cioè  $P_0^{-1} = 0$ .

*Passo 2:* Si applica il filtro informazione per le prime  $c$  ( $c \leq n$ ) osservazioni<sup>4</sup>, dove  $n$  rappresenta la dimensione dello stato, e si inverte la matrice  $P_{c/c}^{-1}$ .

<sup>4</sup> Si ha  $c = n$  quando il sistema ha un output scalare e variabilità deterministica del vettore di stato [si veda Carraro 1988].

*Passo 3:* Si applica il filtro di Kalman dalla  $c + 1$ -esima osservazione fino alla  $T$ -esima. Il filtro viene inizializzato con i valori dello stato  $\mathbf{z}_{c/c}$  e la sua matrice di covarianza  $\mathbf{P}_{c/c}$  determinati al passo precedente.

*Passo 4:* Si applica lo smoothing dalla  $T$ -esima, a ritroso, fino alla  $c$ -esima osservazione<sup>5</sup>. Si ottengono così  $\{x_{t/T}; t = c + 1, \dots, T\}$ , le sequenze dei residui  $\{\hat{\mathbf{e}}_{t/T}$  e  $\hat{\mathbf{u}}_{t/T}; t = c + 1, \dots, T\}$  oltre che le condizioni iniziali  $\mathbf{z}_{c/T}$  e  $\mathbf{P}_{c/T}$  utilizzabili nel filtro di Kalman in successive iterazioni.

*Passo 5:* Si regredisce  $\mathbf{z}_{t/T}$  su  $\mathbf{z}_{t-1/T}$  per la stima di  $\mathbf{M}$  e si usano opportuni stimatori (discussi in seguito) per  $\mathbf{R}$  e  $\mathbf{Q}$ .

*Passo 6:* Si re-inizializza il filtro di Kalman usando le condizioni iniziali determinate al Passo 4 e le stime calcolate al Passo 5. Si ripete la procedura dal Passo 3 al Passo 6 fino alla convergenza.

In realtà per quanto riguarda la matrice di covarianza dello stato si dimostra che la procedura iterativa richiede al Passo 6 che la re-inizializzazione utilizzi come condizione iniziale la matrice  $2 \times \mathbf{P}_{c/T}$ . Ciò permette di eliminare l'effetto *consistenza* artificialmente indotto dalla replicazione dell'algoritmo sullo stesso insieme campionario. Secondo questo effetto, la matrice  $\mathbf{P}_{c/T}$  tende ad una matrice singolare, qualora i suoi elementi non vengano raddoppiati all'inizio di ogni iterazione. La correzione si dimostra appropriata nel caso di stime di una matrice di covarianza di uno stato costante nel tempo, poiché, pur essendo condizionale al valore iniziale di  $\mathbf{R}_0$ , converge alla matrice di covarianza campionaria non condizionale. La stessa correzione è inoperante qualora sia riferita alla matrice di covarianza di uno stato variabile [si veda Sartore 1990].

Negli esempi riportati nell'ultimo paragrafo si mostra come la procedura iterativa applicata dall'algoritmo SRIFIS a modelli con stato variabile porti la matrice di covarianza stimata  $\hat{\mathbf{R}}$ , o alternativamente, la matrice  $\hat{\mathbf{Q}}$  a divenire singolare. La singolarità non è imputabile alla statistica utilizzata per la stima, ma va interpretata come risultato della estrema flessibilità del modello, indotta dalla presenza di parametri variabili stocastici. In questa classe di modelli, la procedura iterativa porta la curva dei parametri variabili ad adattarsi perfettamente ai dati, cioè a coincidere con la curva dell'output (caso della stima singolare di  $\hat{\mathbf{R}}$ ), oppure ad adattarsi perfettamente al suo valore atteso condizionale:  $E(\beta_t/\beta_{t-1}, \mathbf{v}_t)$  (caso della stima singolare di  $\hat{\mathbf{Q}}$ ).

<sup>5</sup> Qualora non si applichi la procedura di Information Filter o si rinunci all'utilizzazione di una distribuzione a priori diffusa sulle condizioni iniziali, risulta  $c = 0$ .

#### 4. Stimatori della matrice di transizione $M$ e delle matrici di covarianza $R$ e $Q$

La verifica della affidabilità ed efficienza dell'algoritmo SRIFIS si è ottenuta con il metodo Monte-Carlo. In un precedente lavoro [Carraro e Sartore 1984] sono stati discussi diversi metodi di stima degli iperparametri riferiti all'applicazione del filtro radice quadrata di Kalman iterativo (denominato semplicemente *Iterative Kalman Filter* [IKF]). In modo analogo si sono verificate le «performances» di stimatori, alcuni dei quali proposti in letteratura, applicandoli alle sequenze dei residui ottenuti al Passo 4 dello SRIFIS. Si possono elencare tali stimatori come segue:

a) *Stimatori di  $M$  e  $Q$* . Possono essere suddivisi in stimatori *condizionali* e *non condizionali* rispetto ai valori stimati dello stato:

a.1) *stimatori condizionali*. Appartengono al primo tipo gli stimatori di  $M$  e di  $Q$  costruiti nel seguente modo. Si supponga per semplicità di far riferimento alla generica equazione  $i$ -esima del modello in cui compare un solo parametro variabile. Nel vettore di stato  $z_{t/T}$  saranno comprese anche le stime dei parametri incogniti dell'equazione del parametro variabile di riferimento, cioè  $\beta_{i,t/T}$ , e le stime dei parametri costanti  $\delta_{i,t/T}$ . Si può quindi definire una nuova variabile  $\beta_{i,t/T}^* = \beta_{i,t/T} - v_{i,t/T} \delta_{i,t/T}$  e stimare  $M$  con il metodo dei minimi quadrati ordinari applicato all'equazione:

$$\beta_{i,t/T}^* = M_i \beta_{i,t-1/T} + u_{i,t}, \quad t = c + 1, \dots, T$$

Lo stimatore della varianza  $Q_i$  è dato da:

$$\hat{Q}_i = \frac{1}{T-c} \sum_{t=c+1}^T \hat{u}_{i,t}^2$$

dove:

$$\hat{u}_{i,t} = \beta_{i,t/T}^* - \hat{M}_i \beta_{i,t-1/T}$$

Nel caso di  $k_i$  parametri variabili,  $M$  non è più uno scalare ma una matrice propria. Il metodo di stima rimane quello dei minimi quadrati ordinari sia che  $Q$  sia supposta diagonale o piena, poiché le variabili esplicative sono identiche in tutte le equazioni. La formulazione compatta dello stimatore di  $M$ , riferito all'equazione  $i$ -esima, diventa:

$$\hat{M}_i = Y_i X_i' (X_i X_i')^{-1}$$

dove:

$$\begin{aligned}
 \mathbf{Y}_i &= \begin{bmatrix} \beta_{1,c+1/T} - \mathbf{v}_{1,c+1} \delta_{1,c+1/T} & \dots & \beta_{1,T/T} - \mathbf{v}_{1,T} \delta_{1,T/T} \\ \dots & \dots & \dots \\ \beta_{k_p,c+1/T} - \mathbf{v}_{k_p,c+1} \delta_{k_p,c+1/T} & \dots & \beta_{k_p,T/T} - \mathbf{v}_{k_p,T} \delta_{k_p,T/T} \end{bmatrix} \\
 \mathbf{X}_i &= \begin{bmatrix} \beta_{1,c/T} & \dots & \beta_{1,T/T} \\ \dots & \dots & \dots \\ \beta_{k_p,c/T} & \dots & \beta_{k_p,T/T} \end{bmatrix}
 \end{aligned}$$

Lo stimatore della matrice di covarianza  $\mathbf{Q}$  sarà dato semplicemente da  $\hat{\mathbf{Q}} = \hat{\mathbf{U}}\hat{\mathbf{U}}'/(T-c)$ .

a.2) *Stimatori non condizionali.* Il secondo tipo di stimatori sfrutta la proprietà di consistenza delle stime dello stato ed utilizza i valori dei parametri variabili stimati alla stregua di *osservabili*, per determinare non solo le stime degli elementi della matrice  $\mathbf{M}$ , ma anche le stime dei parametri  $\delta$ .

Se l'equazione di riferimento  $i$ -esima possiede un solo parametro variabile, allora  $\mathbf{M}$  viene stimato facendo riferimento all'espressione:

$$\beta_{i,t/T} = M_i \beta_{i,t-1/T} + \mathbf{v}_{i,t} \delta_{i,t/T} + \mathbf{u}_{i,t}^*, \quad t = c + 1, \dots, T$$

attraverso l'applicazione del metodo dei minimi quadrati ordinari.

L'estensione al caso di  $k_i$  parametri variabili si ha mediante una opportuna generalizzazione dell'espressione precedente, nel seguente modo:

$$\beta_{i,t/T} = (\beta'_{i,t-1/T} \otimes \mathbf{I}_{k_i}) \text{vec}(\mathbf{M}_i) + \mathbf{v}_{i,t} \delta_{i,t/T} + \mathbf{u}_{i,t}^*, \quad t = c + 1, \dots, T$$

dove i parametri  $\beta_{i,t/T}$  rappresentano vettori di ordine  $k_i$ . Ora, ponendo in forma matriciale l'intero sistema, si ottiene:

$$\begin{bmatrix} \beta_{i,c+1/T} \\ \vdots \\ \beta_{i,T/T} \end{bmatrix} = \begin{bmatrix} \beta_{i,c/T} \otimes \mathbf{I}_{k_i} \\ \vdots \\ \beta_{i,T/T} \otimes \mathbf{I}_{k_i} \end{bmatrix} \boldsymbol{\theta}_i + \begin{bmatrix} \mathbf{v}_{i,c} \\ \vdots \\ \mathbf{v}_{i,T} \end{bmatrix} \boldsymbol{\delta}_i + \begin{bmatrix} \mathbf{u}_{i,c}^* \\ \vdots \\ \mathbf{u}_{i,T}^* \end{bmatrix}$$

dove  $\boldsymbol{\theta}_i = \text{vec}(\mathbf{M}_i)$ . Infine, il sistema viene posto nella nota forma:

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\alpha}_i + \mathbf{U}_i^*, \text{ mediante le relazioni:}$$

$$Y_i = \begin{bmatrix} \beta_{i,c+1/T} \\ \vdots \\ \beta_{i,T/T} \end{bmatrix} \quad X_i = \begin{bmatrix} \beta_{i,c/T} \otimes I_{k_i} & \vdots & v_{i,c} \\ \vdots & \vdots & \vdots \\ \beta_{i,T/T} \otimes I_{k_i} & \vdots & v_{i,T} \end{bmatrix} \quad U_i^* = \begin{bmatrix} u_{i,c}^* \\ \vdots \\ u_{i,T}^* \end{bmatrix}$$

e

$$\alpha_i = \begin{bmatrix} \theta_i \\ \delta_i \end{bmatrix}.$$

I parametri  $\alpha_i$  vengono stimati con il metodo dei minimi quadrati, cioè:  $\hat{\alpha}_i = (X_i' X_i)^{-1} X_i' Y_i$ .

I risultati degli esperimenti di Monte-Carlo, qui tralasciati per brevità<sup>6</sup>, mostrano la superiorità del metodo non condizionale rispetto a quello condizionale. La ragione va probabilmente cercata nel fatto che il metodo di stima condizionale può essere interpretato come metodo di stima vincolato, dove il vincolo è rappresentato dal fatto che si fissano i valori stimati dei parametri costanti relativi alle equazioni dei parametri variabili. Queste stime possono risultare sensibili ai valori iniziali delle matrici  $R$  e  $Q$ , qualora essi siano abbastanza lontani da quelli che hanno generato i dati.

b) *Stimatori di R*. Utilizzando il metodo Monte-Carlo sono stati sottoposti a verifica diversi estimatori della matrice di covarianza  $R$ .

Lo stimatore di partenza utilizza i residui dello smoothing alla stregua dei residui di un qualsiasi modello di regressione multiequazionale stimato in modo efficiente. Si ha quindi:

$$\hat{\epsilon} = y_t - H_t z_{t/T}$$

$$\hat{R} = \frac{1}{T - c} \sum_{t=c+1}^T \hat{\epsilon}_t \hat{\epsilon}_t'$$

Gli esperimenti di simulazione condotti con l'impiego delle stime ricavate dalla precedente formulazione mostrano che, qualsiasi sia la generazione campionaria, la procedura iterativa dello SRIFIS determina una sequenza di  $\hat{R}$  che converge a zero nel caso in cui  $R$  sia uno scalare, e più generalmente ad una matrice singolare.

Questo comportamento ha suggerito l'idea di apportare delle correzioni allo stimatore iniziale della matrice di covarianza, nel tentativo di ottenere la convergenza ai valori teorici utilizzati nella procedura di generazione campionaria.

Si riportano qui di seguito gli algoritmi sperimentati:

<sup>6</sup> Alcuni esperimenti sono riportati in Sartore [1990].

a) Il primo algoritmo ha come punto di partenza la seguente relazione:

$$\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \mathbf{H}_t \mathbf{z}_{t/T} - \mathbf{H}_t (\mathbf{z}_t - \mathbf{z}_{t/T})$$

che si può scrivere:

$$\mathbf{y}_t - \mathbf{H}_t \mathbf{z}_{t/T} = \boldsymbol{\varepsilon}_t + \mathbf{H}_t (\mathbf{z}_t - \mathbf{z}_{t/T})$$

Sommando i termini di destra e di sinistra dell'equazione e dividendo per la numerosità campionaria, si ottiene:

$$\begin{aligned} \frac{1}{T-c} \sum_{t=c+1}^T (\mathbf{y}_t - \mathbf{H}_t \mathbf{z}_{t/T}) (\mathbf{y}_t - \mathbf{H}_t \mathbf{z}_{t/T})' &= \frac{1}{T-c} \sum_{t=c+1}^T \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' + \\ &+ \frac{1}{T-c} \sum_{t=c+1}^T \mathbf{H}_t (\mathbf{z}_t - \mathbf{z}_{t/T}) (\mathbf{z}_t - \mathbf{z}_{t/T})' \mathbf{H}_t' \end{aligned}$$

Da cui si determina facilmente il termine di correzione di  $\hat{\mathbf{R}}$  supponendo che il prodotto  $(\mathbf{z}_t - \mathbf{z}_{t/T}) (\mathbf{z}_t - \mathbf{z}_{t/T})'$  sia ben approssimato dal suo valore atteso, cioè  $\mathbf{P}_{t/T}$ . Si ha quindi:

$$\tilde{\mathbf{R}} = \hat{\mathbf{R}} - \frac{1}{T-c} \sum_{t=c+1}^T \mathbf{H}_t \mathbf{P}_{t/T} \mathbf{H}_t'$$

b) Un secondo algoritmo tenta di determinare la correzione di  $\hat{\mathbf{R}}$  attraverso l'utilizzazione di variabili ausiliarie. Si consideri la relazione:

$$\begin{aligned} \hat{\boldsymbol{\varepsilon}}_t &= \mathbf{y}_t - \mathbf{H}_t \mathbf{z}_{t/T} \\ &= \boldsymbol{\varepsilon}_t + \mathbf{H}_t (\mathbf{z}_t - \mathbf{z}_{t/T}) \\ &= \boldsymbol{\varepsilon}_t + \mathbf{H}_t \mathbf{z}_t - \mathbf{H}_t \mathbf{z}_{t/T} \end{aligned}$$

Ponendo  $\tilde{\boldsymbol{\varepsilon}}_t = \boldsymbol{\varepsilon}_t + \mathbf{H}_t \mathbf{z}_t$  si può scrivere la relazione:

$$\hat{\boldsymbol{\varepsilon}}_t = \tilde{\boldsymbol{\varepsilon}}_t - \tilde{\mathbf{H}}_t \mathbf{z}_{t/T}$$

Quest'ultimo risultato suggerisce l'uso della variabile ausiliaria  $\mathbf{H}_t \mathbf{z}_{t/T}$ . Si definisca quindi:  $\mathbf{W}_t \equiv \mathbf{H}_t \mathbf{z}_{t/T}$ . Ciò permette di costruire la regressione:

$$\mathbf{y}_t = \boldsymbol{\theta} \mathbf{W}_t + \boldsymbol{\eta}_t \quad t = c + 1, \dots, T$$

L'intero sistema in forma matriciale può essere scritto  $\mathbf{Y} = \boldsymbol{\theta} \mathbf{W} + \boldsymbol{\eta}$ , dove:

$$Y = [y_{c+1} \dots y_T] \quad W = [W_{c+1} \dots W_T] \quad \eta = [\eta_{c+1} \dots \eta_T]$$

La stima del parametro  $\theta$  si ottiene mediante l'applicazione del metodo dei minimi quadrati, e risulta:  $\hat{\theta} = YW'(WW')^{-1}$ .

La stima di  $R$  è fondata sui valori residui della regressione, cioè:

$$\eta = Y - \hat{\theta}W$$

$$\hat{R} = \frac{1}{T-c} \eta\eta'$$

c) Il terzo algoritmo è ancora fondato sulla relazione:

$$\hat{\epsilon}_t = \epsilon_t + H_t(z_t - z_{t/T})$$

ma si considera ora la sua forma quadratica:

$$\hat{\epsilon}_t \hat{\epsilon}_t' = \epsilon_t \epsilon_t' + H_t(z_t - z_{t/T})(z_t - z_{t/T})' H_t' + \epsilon_t(z_t - z_{t/T})' H_t' + H_t(z_t - z_{t/T}) \epsilon_t'$$

La procedura di stima di  $R$  risulta dalle seguenti formulazioni:

$$\begin{bmatrix} \hat{\epsilon}_{c+1} \hat{\epsilon}_{c+1}' \\ \vdots \\ \hat{\epsilon}_T \hat{\epsilon}_T' \end{bmatrix} = \begin{bmatrix} R \\ \vdots \\ R \end{bmatrix} + \begin{bmatrix} H_{c+1} z_{c+1/T} z_{c+1/T}' H_{c+1}' \\ \vdots \\ H_T z_{T/T} z_{T/T}' H_T' \end{bmatrix} + \begin{bmatrix} H_{c+1} z_{c+1/T} \\ \vdots \\ H_T z_{T/T} \end{bmatrix} \epsilon' + \eta$$

$$Y = \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix} R + \begin{bmatrix} H_{c+1} z_{c+1/T} z_{c+1/T}' H_{c+1}' \\ \vdots \\ H_T z_{T/T} z_{T/T}' H_T' \end{bmatrix} \theta_1 + \begin{bmatrix} H_{c+1} z_{c+1/T} \\ \vdots \\ H_T z_{T/T} \end{bmatrix} \theta_2 + \eta$$

$$Y = \left[ \begin{array}{c|c|c} I & H_{c+1} z_{c+1/T} z_{c+1/T}' H_{c+1}' & H_{c+1} z_{c+1/T} \\ \vdots & \vdots & \vdots \\ I & H_T z_{T/T} z_{T/T}' H_T' & H_T z_{T/T} \end{array} \right] \begin{bmatrix} R \\ \theta_1 \\ \theta_2 \end{bmatrix} + \eta$$

$$Y = X\theta + \eta$$

$$\hat{\theta} = (X'X)^{-1} X'Y$$

I risultati ottenuti con i tre algoritmi riportati non sono stati confortanti poiché si sono dimostrati estremamente sensibili ai valori ini-

ziali assegnati alle matrici  $R$  e  $Q$ . Nella totalità dei casi si sono ottenuti, infatti, valori che si scostavano da quelli iniziali in modo poco significativo.

## 5. Alcuni risultati con simulazioni di Monte-Carlo

In questo paragrafo sono riportati due esempi di esperimenti con il metodo Monte-Carlo. Ad essi non si vuole assegnare lo stesso significato dimostrativo proprio degli esperimenti ampiamente replicati. Se, per questi ultimi, i risultati sono condizionali al generatore dei disturbi casuali, i risultati di una sola replicazione sono ovviamente condizionali ad una particolare generazione di disturbi casuali. Lo scopo quindi delle esemplificazioni sotto riportate è semplicemente illustrativo. Va sottolineato, tuttavia, che tutte le esperienze di simulazione effettuate dagli Autori hanno presentato connotazioni del tutto simili.

ESEMPIO 1. Il primo modello riportato presenta solo parametri costanti, ed è definito come segue:

$$[M1] \quad y_t = \gamma_0 + \sum_{j=1}^2 \gamma_j w_{jt} + \varepsilon_t$$

$$\text{dove: } \gamma_0 = 10; \quad \gamma_1 = -1,5; \quad \gamma_2 = 0,7$$

$$\text{Var}(\varepsilon_t) = 100, \quad \forall t$$

$$\hat{\sigma}_y = 9448,3061$$

La dimensione campionaria prescelta è di 200 osservazioni. I disturbi dell'equazione della variabile dipendente del modello sono generati da una distribuzione normale in modo indipendente dai disturbi generati per la costruzione delle serie delle variabili esplicative. Queste ultime seguono dei semplici processi autoregressivi del primo ordine. Poiché tutta l'analisi è condotta in modo condizionale a queste variabili, non è rilevante illustrare il processo generatore di tali dati.

Essendo nota la specificazione del modello si è proceduto alla stima con il metodo dei minimi quadrati ordinari utilizzando il pacchetto Micro-TSP, al fine di avere termini di confronto con i risultati ottenuti dalla procedura iterativa dello SRIFIS. Nella seguente tabella sono riportati i risultati ottenuti<sup>7</sup>:

<sup>7</sup> Per brevità non si riportano i valori degli errori standard. Essi dipendono in ogni caso interamente dalla stima del parametro  $\sigma_\varepsilon$ .

TAB. 1. *Stime dei parametri del modello (M1) con il metodo dei minimi quadrati ordinari non ricorsivo*

$$\hat{\gamma}_0 = 9,8968, \hat{\gamma}_1 = -1,5099, \hat{\gamma}_2 = 0,6914$$

$$\hat{\sigma}_e = 107,6354$$

L'applicazione dello SRIFIS ha fornito, invece, i seguenti valori:

TAB. 2. *Stime dei parametri del modello (M1) ottenute con l'algoritmo SRIFIS partendo dai valori iniziali  $R_0$  riportati. Applicazione dell'«information filter»*

It. n. 1	$\hat{\gamma}_0 = 9,8968, \hat{\gamma}_1 = -1,5099, \hat{\gamma}_2 = 0,6914$ $R_0 = 9448,3061, \hat{R} = 115,8868$
It. n. 2	$\hat{\gamma}_0 = 9,9478, \hat{\gamma}_1 = -1,5082, \hat{\gamma}_2 = 0,6919$ $R_0 = 115,8868, \hat{R} = 114,5427$
It. n. 3	$\hat{\gamma}_0 = 9,9482, \hat{\gamma}_1 = -1,5082, \hat{\gamma}_2 = 0,6919$ $R_0 = 114,5427, \hat{R} = 109,4931$
It. n. 16	$\hat{\gamma}_0 = 9,9485, \hat{\gamma}_1 = -1,5082, \hat{\gamma}_2 = 0,6919$ $R_0 = 108,5094, \hat{R} = 108,5009$

Per il confronto con i valori della tabella precedente si tenga presente che  $\hat{R}$  rappresenta la stima della varianza  $\sigma_e$ . Si osservi che ad ogni iterazione si è utilizzato come valore iniziale della varianza, il valore stimato calcolato nell'iterazione precedente. Alla 16<sup>a</sup> iterazione le stime dei coefficienti delle variabili esplicative rimangono costanti nella quarta cifra decimale e la stima della varianza è molto prossima a quella della tabella 1, dopo 30 iterazioni la stima scende al valore 108,4474, quindi non vi è nessun guadagno sensibile.

Il confronto tra varianze deve tener conto del fatto che l'applicazione dell'*information filter* riduce, in questo caso, l'ordine del vettore dei residui di una quantità pari alla dimensione dello stato. Quindi il calcolo della stima della varianza è basato solo su 197 residui. Qualora si rinunci all'applicazione dell'*information filter*, non si ha riduzione della dimensione campionaria. I risultati ottenuti in questo modo sono riportati nella tabella 3.

TAB. 3. Stime dei parametri del modello (M1) ottenute con l'algoritmo SRIFIS partendo dai valori iniziali  $R_0$  riportati. Senza applicazione dell'«information filter»

It. n. 1	$\hat{\gamma}_0 = 9,8963, \hat{\gamma}_1 = -1,5099, \hat{\gamma}_2 = 0,6914$ $R_0 = 9448,3061, \hat{R} = 285,4995$
It. n. 2	$\hat{\gamma}_0 = 9,8968, \hat{\gamma}_1 = -1,5099, \hat{\gamma}_2 = 0,6914$ $R_0 = 285,4995, \hat{R} = 112,5844$
It. n. 3	$\hat{\gamma}_0 = 9,8968, \hat{\gamma}_1 = -1,5099, \hat{\gamma}_2 = 0,6914$ $R_0 = 112,5844, \hat{R} = 109,5536$

In questo caso le stime dei coefficienti delle variabili esplicative rimangono costanti nella quarta cifra decimale già dalla seconda iterazione, mentre la convergenza della stima della varianza risulta più lenta. Alla 30<sup>a</sup> iterazione il valore è 107,6969, questa volta comparabile con il valore riportato dalla tabella 1.

Si osservi come la convergenza alla varianza teorica sia in questo caso leggermente più lenta. Questo risultato è coerente con l'applicazione dell'*information filter*, che permette un più rapido apprendimento dai dati poiché la distribuzione a priori implicitamente utilizzata è *perfettamente* diffusa, cioè possiede varianza infinita.

ESEMPIO 2. Ulteriori esperimenti sono stati condotti sul seguente modello con un solo parametro variabile e sei parametri costanti:

$$\begin{aligned} y_t &= x_t \beta_t + \gamma_0 + \sum_{j=1}^3 \gamma_j w_{jt} + \varepsilon_t \\ \beta_t &= M \beta_{t-1} + \delta_0 + \delta_1 v_t \end{aligned}$$

dove:  $M = 0,7; \gamma_0 = 5; \gamma_1 = 1; \gamma_2 = 0,5; \gamma_3 = -2; \delta_0 = 10; \delta_1 = -1,5$   
 $Var(\varepsilon_t) = 625, \forall t; Var(u_t) = 9, \forall t.$

Anche per questo modello i disturbi dell'equazione della variabile endogena sono stati generati da una distribuzione normale in modo indipendente dai disturbi generati per la costruzione delle serie delle variabili esplicative, che seguono dei processi autoregressivi del primo ordine.

Conoscendo i veri valori degli iperparametri si è applicato l'algoritmo SRIFIS utilizzando questo punto di partenza, utile per i confronti con risultati successivi. Si riporta quanto ottenuto nella seguente tabella:

TAB. 4. Valori degli iperparametri ottenuti con l'algoritmo SRIFIS partendo dai valori iniziali scritti nel primo riquadro. Nessuno degli iperparametri viene vincolato ad assumere lo stesso valore iniziale durante la procedura iterativa

	Val. iniziali	Iter. 0	Iter. 1	Iter. 2	Iter. 3
$\hat{M}$	0,7	0,6945	0,6939	0,6936	0,6929
$\hat{Q}$	9,0	7,3970	9,2295	11,8341	14,8263
$\hat{R}$	625,0	137,1493	53,4147	24,8019	13,8035
EQMT		16,3531	29,1232	51,7595	74,8659
Veqmt		12,1109	21,0507	92,4262	261,4227
		Iter. 5	Iter. 7	Iter. 9	Iter. 11
$\hat{M}$		0,6913	0,6894	0,6860	0,6843
$\hat{Q}$		21,7471	33,5413	62,2611	78,2666
$\hat{R}$		6,5813	2,9270	0,2708	0,0000
EQMT		93,3493	100,0160	113,8857	122,0896
Veqmt		1189,8305	6687,5563	41562,2337	73450,9380

La valutazione della bontà dei risultati si ottiene elevando al quadrato la norma della differenza tra il vettore di stato stimato e quello teorico, queste lunghezze vengono sommate rispetto al tempo. Nella tabella questo indice viene identificato con la sigla *EQMT*, per richiamare il concetto di errore quadratico medio a cui questo indicatore si ispira. Con *Veqmt* si indica, invece, la varianza dell'indice stesso.

Analizzando la tabella 4, si osservi come la stima della varianza *R* si annulli alla 11<sup>a</sup> iterazione; ciò comporta la tendenza della stima della varianza *Q*, ad aumentare senza intaccare sostanzialmente la struttura autoregressiva del parametro variabile. La stima di *R* tende a zero grazie all'estrema *capacità di adattamento* del parametro variabile. Esso rende talmente flessibile l'iperpiano di regressione che l'interpolazione *tra* punti viene trasformata in una interpolazione *per* punti. L'adattamento si ottiene al costo della distorsione delle stime dei parametri del modello come dimostra l'indice *EQMT* e la sua varianza. Il punto di minima distanza e di minima variabilità della stessa si ha in corrispondenza dell'iterazione zero che consegue dall'assegnazione dei veri valori dei parametri come valori iniziali.

In generale, può essere la stima della varianza *R* oppure *Q* a tendere a zero. Questa tendenza si accompagnerà alla distorsione della stima della varianza che non si annulla per effetto dell'assorbimento di variabilità ad essa non imputabile. La distorsione può interessare anche la stima del parametro *M*, si dovranno quindi scegliere come valori iniziali quelle stime relative ad iterazioni in cui si suppone che l'effetto di distorsione sulle varianze sia il più limitato possibile.

Una strategia per fissare i valori iniziali degli iperparametri può essere suggerita dal comportamento delle varianze appena descritto, e può essere riassunta come segue:

i) si fissa il valore iniziale di  $Q$  molto piccolo e costante ad ogni iterazione; in tal modo l'algoritmo porterà verosimilmente la stima di  $R$  a valori sempre più elevati facendo tendere a zero la stima di  $Q$ ;

ii) in modo analogo si fissa il valore iniziale di  $R$  molto piccolo e costante ad ogni iterazione; questa volta saranno le stime di  $Q$  ad avere valori sempre più elevati mentre tenderà a zero la stima di  $R$ .

Questi primi due passi dovrebbero portare ad ottenere i limiti massimi del valore delle stime di  $Q$  ed  $R$ , ed a suggerire un possibile valore iniziale per  $M$ . Successivamente:

iii) si fissano come valori iniziali i limiti massimi e, senza mantenerli costanti ad ogni iterazione, si permette all'algoritmo di trovare la combinazione *ottimale* tra le stime delle due varianze.

L'ottimalità della combinazione in questo esempio viene indicata ancora sulla base dell'indice  $EQMT$  e della sua varianza. Nel caso di dati non simulati tali indici non possono essere disponibili e si utilizza un altro indicatore, che sostituisce al valore teorico dello stato, la stima dello stato ottenuta nell'iterazione precedente. Gli esperimenti di simulazione hanno mostrato una discreta sensibilità di questo indice nel possedere un punto di minimo relativo in corrispondenza di valori iniziali di  $R$  e  $Q$  vicini a quelli teorici.

Attuando la procedura sopra indicata, si sono dapprima assunti i valori iniziali  $M = 0,69$ ,  $Q = 0,01$ ,  $R = 1$ , vincolando i primi due valori a rimanere costanti come valori iniziali per tutte le iterazioni dell'algoritmo ed utilizzando come valore iniziale di  $R$  la stima dell'iterazione precedente. La stima di  $R$  è cresciuta rapidamente raggiungendo i seguenti valori nelle prime tre iterazioni: 14451,7522, 54879,4124, 56750,0698. Successivamente i valori di  $R$  hanno oscillato di molto poco attorno a quest'ultimo valore. Parallelamente i valori stimati di  $Q$  sono risultati: 0,6024, 0,0009, 0,0001. L'ultimo numero è il valore su cui si è stabilizzata la stima di  $Q$ . Le stime di  $M$  sono stabili sul valore 0,6886.

Successivamente sono stati assunti i valori iniziali di  $M = 0,69$ ,  $Q = 1$ ,  $R = 0,01$ , vincolando  $M$  ed  $R$  ad essere costanti come valori iniziali per tutte le iterazioni. In questo caso non avendo vincolato il valore di  $Q$ , la procedura si arresta alla terza iterazione poiché la stima di  $R$  si annulla. Dalla prima iterazione si ottengono i valori di  $M = 0,684$  e  $Q = 79,6077$  che rimangono costanti nelle due iterazioni successive.

I valori massimi di varianza per  $R$  e  $Q$  sono dunque 56750 e 80, che possono essere assunti come valori iniziali per una nuova sequenza di iterazioni, mentre il valore di  $M$  può rimanere fissato a 0,69. Nella tabella 5 sono riportati i risultati ottenuti con questi valori:

TAB. 5. Valori degli iperparametri ottenuti con l'algoritmo *SRIFIS* partendo dai valori iniziali scritti nel primo riquadro. Solo il parametro *M* viene vincolato ad assumere lo stesso valore iniziale durante la procedura iterativa

	Val. iniziali	Iter. 0	Iter. 1	Iter. 2	Iter. 3
$\hat{M}$	0,69	0,6931	0,6934	0,6936	0,6937
$\hat{Q}$	80,00	3,8052	5,1155	6,8468	8,9538
$\hat{R}$	56750,00	1062,6888	447,7438	164,0732	61,7712
EQMT		19,7437	21,8487	21,1269	27,6307
$V_{eqmt}$		37,8089	21,7089	11,8608	16,0701
EQM			0,3218	0,2616	0,8330
$V_{eqm}$			0,1249	0,4121	1,4931
		Iter. 4	Iter. 5	Iter. 7	Iter. 9
$\hat{M}$		0,6935	0,6929	0,6913	0,6894
$\hat{Q}$		11,5024	14,4757	21,3523	32,7587
$\hat{R}$		27,5814	14,7632	6,7913	3,0773
EQMT		48,0039	72,5150	93,2753	99,8336
$V_{eqmt}$		72,3728	231,7850	1102,4752	6173,6364
EQM		3,1898	2,7796	0,4786	0,6585
$V_{eqm}$		5,4789	4,1160	7,2557	14,9498

Una corretta interpretazione dei risultati della tabella deve tener presente che, ad ogni iterazione, i valori iniziali per i parametri *Q* ed *R* sono forniti dalle stime degli stessi nell'iterazione precedente. La lettura della tabella porta ad alcune osservazioni:

1. L'indice EQMT e la sua varianza indicano la seconda iterazione preferibile alle successive, poiché in essa si ha uno scostamento dei valori veri dei parametri molto contenuto ed una variabilità minima di tale scostamento. All'aumentare delle iterazioni, successivamente alla seconda, l'indice e la sua varianza forniscono chiara evidenza della distorsione subita dai parametri costanti e variabili del modello.

2. Il passaggio dai valori della prima iterazione a quelli della seconda possono essere comparati con quelli della tabella 4 nel passaggio dai valori iniziali (veri) alle stime dell'iterazione zero. Si ottiene un'analogia riduzione nella stima della varianza *R*, mentre per *Q* vi è un recupero dell'effetto di schiacciamento della stima provocato dall'alta varianza iniziale di *R*.

3. Nella tabella 5 sono riportati anche i valori dell'indice EQM e della sua varianza ottenuti, come si è già accennato, sostituendo al valore teorico dello stato, la stima dello stato ottenuta nell'iterazione precedente. Si noti come tale indice abbia il suo minimo in corrispondenza dei risultati della seconda iterazione, mentre la sua varianza rimane molto contenuta. Va sottolineato che tale minimo è un minimo relativo, infatti all'aumentare delle iterazioni i parametri costanti e variabili del modello tendono ad appiattirsi sulla curva di

massima distorsione, l'indice quindi e la sua varianza tendono inevitabilmente a zero:

4. I valori degli iperparametri risultano sottostimati, anche se non di molto, e ciò influisce sulle statistiche per la valutazione della bontà degli stimatori sia dei parametri costanti che del parametro variabile.

Sembra qui opportuno riportare anche i valori delle stime dei parametri ottenuti. Si prenda in considerazione anzitutto le stime dei parametri costanti ed i relativi errori standard. Per dare una valutazione dei valori ottenuti e misurare, relativamente al campione utilizzato, l'incidenza delle stime del parametro variabile, si è ritenuto opportuno effettuare il confronto riportato nella successiva tabella 6. In essa, le stime dei parametri ottenute con Micro-TSP sono condizionali ai valori teorici del parametro variabile del modello, mentre quelle ottenute con SRIFIS sono condizionali ai valori stimati del parametro variabile.

Poiché Micro-TSP non stima modelli a parametri variabili, il risultato per il confronto si è ottenuto interpretando il modello a parametri variabili come un modello non lineare nelle variabili così definito:

$$y_t = \alpha x_t \beta_t + \gamma_0 + \sum_{j=1}^3 \gamma_j w_{jt} + \varepsilon_t$$

Successivamente, la non linearità indotta dal parametro variabile è stata agevolmente eliminata costruendo un nuovo regressore, dato dal prodotto tra lo stesso parametro variabile e la variabile esplicativa a cui tale parametro fa riferimento. Il parametro  $\alpha$  nel modello teorico deve necessariamente assumere valore unitario. Si è comunque preferito non imporre tale vincolo e stimare anche il parametro  $\alpha$  come ulteriore verifica di bontà di adattamento dei dati al modello teorico.

TAB. 6. Confronto tra stime dei parametri costanti del modello a parametri variabili. Le stime ottenute con Micro-TSP sono condizionali ai valori teorici del parametro variabile del modello. Le stime ottenute con SRIFIS sono condizionali ai valori stimati del parametro variabile. Tra parentesi sotto i nomi delle stime sono riportati i valori teorici dei parametri e sotto i valori delle stime gli errori standard

Metodi di stima	Coefficienti stimati							
	$\alpha$ (1)	$\gamma_0$ (5)	$\gamma_1$ (1)	$\gamma_2$ (0,5)	$\gamma_3$ (-2)	$\hat{M}$ (0,7)	$\hat{\delta}_0$ (10)	$\hat{\delta}_1$ (-1,5)
Micro TSP	0,9999 (0,0003)	7,4904 (1,8692)	1,0324 (0,03136)	0,5125 (0,0430)	-1,9829 (0,0241)	0,6937 (0,0027)	10,3450 (0,2641)	-1,5024 (0,0066)
SRIFIS	—	9,3808 (3,0660)	1,1378 (0,0446)	0,3657 (0,0771)	-1,9211 (0,0373)	0,6936 (0,0023)	10,5784 (0,1285)	-1,5047 (0,0040)

L'analisi dei risultati deve necessariamente tener conto del fatto che si è considerato un singolo insieme campionario ed uno specifico modello generatore dei dati, quindi vanno assolutamente evitate generalizzazioni che potrebbero apparire fuorvianti rispetto a indagini inferenziali più complete. Valutando, tuttavia, il loro parziale valore indiziario, è evidente la vicinanza delle stime per i parametri costanti ottenute con i due metodi. Gli errori standard ottenuti dallo SRIFIS per i parametri dell'equazione della variabile endogena sono più elevati di quelli ottenuti condizionatamente ai veri valori del parametro variabile, viceversa sono più piccoli quelli dell'equazione del parametro variabile.

L'ipotesi che i valori stimati con SRIFIS non differiscano significativamente dalle stime ottenute con Micro-TSP viene accettata per tutti i parametri, eccetto la stima  $\hat{\gamma}_1$  il cui valore  $t_{193} = 2,36$  supera il valore critico al livello di significatività del 5%. Per contro, mentre per le stime Micro-TSP si accetta l'ipotesi che esse non differiscano significativamente dai valori teorici, eccetto la stima di  $M$ , per l'algoritmo SRIFIS solamente le stime di  $\gamma_0$  e  $\delta_1$  non differiscono significativamente dai valori teorici.

Un'ulteriore conferma della presenza di un effetto distorsivo è data dall'applicazione del test  $F$  asintotico di Goldfeld e Quandt<sup>8</sup>, con il quale si verifica l'ipotesi di cambiamento strutturale dei parametri quando le varianze non sono uguali. L'ipotesi di non uguaglianza delle varianze è stata accettata verificando che il test  $F_{195,193} = 3,84$  supera il valore critico al livello di significatività del 5%. Il calcolo del test  $F$  asintotico di Goldfeld e Quandt ha fornito il valore  $F_{5,388} = 5,24$  che supera il valore critico al livello di significatività del 5%.

L'analisi dei residui del modello condotta, attraverso la lettura degli autocorrelogrammi, confermano una lieve presenza di autocorrelazione al ritardo uno per i residui dell'equazione della variabile endogena, il test di Box-Pierce sui primi otto ritardi fornisce il valore 22,036, superiore al valore critico al livello di significatività del 5%. Lo stesso test, applicato ai residui dell'equazione del parametro variabile, fornisce il valore 15,740, inferiore invece al valore critico e quindi mostra maggior conformità di tali residui ad una realizzazione di un processo stocastico puramente casuale.

L'effetto distorsivo, pur leggero, viene indotto dalle stime del parametro variabile. Come si vedrà in seguito, il numero di stime statisticamente non accettabili per il parametro variabile risulta superiore, anche se non di molto, a quello atteso e ciò è sufficiente per influenzare le stime dei parametri costanti del modello.

<sup>8</sup> Si veda Goldfeld e Quandt [1978] oppure Amemiya [1985, 35 e ss.].

Per quanto riguarda i parametri variabili, si può far riferimento alla seguente tabella:

Tab. 7. Valori teorici del parametro variabile ed intervalli di confidenza (coefficiente di confidenza  $\cong 0,95$ ) calcolati

$t$	$1(\beta_t)$	$\beta_t$	$u(\beta_t)$	$t$	$1(\beta_t)$	$\beta_t$	$u(\beta_t)$
8	<sup>a</sup> 0,4291	0,3337	2,6033	56	-50,3609	-48,2360	-45,3423
9	85,5068	87,0980	90,0233	57	-89,0825	-87,7720	-85,5707
10	127,0268	128,0600	132,2302	58	-127,9827	-126,0700	<sup>a</sup> -126,3273
11	143,8835	144,7800	145,0289	59	-38,1501	-35,8110	-35,5940
12	147,6333	148,2400	148,5467	60	46,1588	49,6760	50,0733
13	233,1209	234,4200	237,7977	61	66,4607	67,1630	67,6206
14	148,2264	150,3700	153,5948	62	84,9025	86,9420	<sup>a</sup> 86,4858
15	94,0665	96,8590	98,6501	63	34,0755	39,8430	<sup>a</sup> 38,8085
16	<sup>a</sup> 7,7580	9,0437	9,4612	64	-24,0781	-21,8930	-21,4141
17	-97,1069	-96,2780	-95,9973	65	-41,6276	-40,4730	-39,7514
18	-67,0377	-66,4610	-63,0966	66	<sup>a</sup> 3,7175	3,4468	4,8450
19	-47,9031	-47,6150	-45,2156	67	27,4005	27,5130	28,7510
20	-116,5433	-115,4700	-108,2664	68	-48,0441	-47,1550	-44,3473
21	<sup>a</sup> -92,0708	-93,3240	-84,9443	69	-82,2171	-80,1050	-78,5104
22	-121,3496	-121,3900	-116,7395	70	-145,7229	-144,4000	-141,2825
23	-80,2905	-77,9030	-76,0459	71	13,6451	14,8260	15,8148
24	42,5000	42,8640	44,0605	72	<sup>a</sup> 78,8445	78,2760	84,0068
25	<sup>a</sup> 24,1650	23,4330	28,2660	73	112,5335	114,2700	116,0145
26	87,3484	88,1450	88,5719	74	111,3748	111,9900	115,4879
27	107,9850	108,5600	108,6907	75	<sup>a</sup> 144,7974	144,7000	149,3668
28	-2,3964	-1,2517	<sup>a</sup> -1,5486	76	66,4663	67,2790	68,7708
29	-65,8004	-64,8600	-64,7511	77	74,5588	77,9310	82,5137
30	52,8147	53,5170	53,5441	78	<sup>a</sup> -48,0983	-48,2490	-42,6522
31	28,8840	29,7140	30,3333	79	-28,7097	-27,3760	-26,0095
32	80,2235	80,9100	81,7050	80	-16,3286	-14,0160	-13,4291
33	75,6152	77,9680	78,2780	81	-78,7007	-77,9680	-77,7426
34	46,7098	47,0630	47,7169	82	-114,0728	-113,4700	-113,1717
35	20,4782	21,1290	21,2706	83	-142,6226	-141,4900	-141,2484
36	38,2724	38,5740	39,2739	84	-153,1135	-151,5200	-148,7583
37	-67,3226	-65,8960	-64,4982	85	-24,5718	-21,0360	<sup>a</sup> -21,0367
38	-88,7751	-87,4160	-84,3544	86	40,5506	42,5980	43,6628
39	-109,2592	-108,1300	-106,2783	87	133,8116	134,8700	141,4161
40	-83,2100	-82,5930	-81,9590	88	206,0799	206,5300	207,2687
41	-70,4557	-70,7010	-65,7818	89	217,8960	222,0500	222,4239
42	7,0866	9,4240	9,6368	90	166,3670	166,6900	167,2729
43	145,3383	146,7100	150,5234	91	136,9937	138,5100	139,1310
44	143,4015	145,0500	146,6753	92	90,3834	91,3280	92,3266
45	166,3158	167,4900	168,1190	93	78,5974	82,9900	84,4824
46	75,8963	77,2960	77,9987	94	53,4165	54,0290	55,6597
47	18,6167	18,7420	19,3792	95	-27,5793	-27,4200	-24,2851
48	31,5514	31,8010	32,1587	96	10,5332	13,4680	15,6901
49	3,9502	4,3488	4,7984	97	26,5500	28,1880	29,2027
50	4,2039	5,2463	5,7309	98	21,3984	21,7040	23,3526
51	11,3351	13,8030	<sup>a</sup> 13,4417	99	82,0388	86,3420	<sup>a</sup> 85,7065
52	-49,4766	-43,0070	-42,3565	100	213,2429	213,7500	214,6243
53	-69,8276	-69,0710	-68,4190	101	185,2768	186,0300	186,2462
54	-42,4358	-41,8370	-41,4223	102	<sup>a</sup> 189,5840	188,6100	197,0925
55	5,7962	10,9520	13,0161	103	145,1197	146,8800	<sup>a</sup> 146,8188

Tab. 7. (segue)

$t$	$l(\hat{\beta}_t)$	$\beta_t$	$u(\hat{\beta}_t)$	$t$	$l(\hat{\beta}_t)$	$\beta_t$	$u(\hat{\beta}_t)$
104	147,0069	149,0600	150,3520	152	92,5681	94,0410	94,9014
105	143,1397	144,0100	145,9906	153	83,4322	85,8570	*85,3188
106	157,9643	161,6800	165,1589	154	107,8215	108,8600	109,4787
107	174,7757	175,4500	175,9576	155	136,0253	138,0300	139,4903
108	*187,8831	187,8100	196,1231	156	162,4727	164,7200	167,9078
109	135,0492	135,8200	142,5697	157	81,0768	86,9100	*85,8578
110	125,8252	127,8200	133,7845	158	73,8550	77,0510	79,5549
111	123,4341	123,9800	124,9399	159	64,8252	65,6310	66,3375
112	115,0481	115,7100	115,8835	160	-22,2957	-19,8100	-19,4414
113	80,0212	80,8720	81,1884	161	-81,3619	-80,1940	-79,6573
114	*89,2254	89,1780	89,9907	162	-54,0200	-52,6250	-45,5034
115	158,5079	159,0200	159,8820	163	-21,4388	-14,0240	-12,9152
116	132,1304	132,9000	133,2086	164	29,8094	33,9990	35,6080
117	59,8613	61,2510	61,8689	165	114,2761	114,6700	116,6727
118	-54,4740	-51,4690	-48,0707	166	177,9651	178,3100	178,8629
119	-53,4536	-51,9780	-49,4643	167	188,5634	189,2600	195,0676
120	-11,7106	-11,3920	-10,4270	168	163,3662	165,9100	166,3732
121	-9,2729	-8,6650	-8,0156	169	167,0875	167,9200	168,0861
122	33,0373	33,2050	33,8527	170	272,6945	273,0700	274,1892
123	12,4029	14,1410	15,8556	171	223,0353	223,6700	224,2073
124	55,2577	58,2460	61,8639	172	110,7530	111,9900	*111,7190
125	4,1598	6,5138	12,2001	173	57,8259	57,8940	58,7350
126	-3,8928	-3,5762	-2,4726	174	113,0718	114,6200	116,2190
127	27,1673	28,4170	28,7075	175	153,7275	153,9600	158,0316
128	7,4095	8,3902	8,8518	176	*109,1916	108,6400	111,1926
129	*86,7186	86,3830	87,7189	177	*104,2343	104,2300	106,3788
130	31,0847	31,4730	31,8862	178	42,9977	43,3060	45,7011
131	-42,2801	-41,6480	-40,7865	179	*45,3290	44,9060	46,4595
132	24,6000	27,0330	29,9616	180	40,6955	41,2270	42,4392
133	27,0037	28,8140	29,5705	181	76,8486	77,1300	78,0168
134	147,4112	148,6000	149,7455	182	112,0803	113,1300	113,6991
135	*33,5501	33,1850	34,9744	183	53,9030	54,3080	54,4944
136	20,2931	21,1690	22,1639	184	63,0405	69,1130	69,5763
137	99,8669	105,3500	107,7382	185	-16,7660	-15,8010	-14,9618
138	52,2265	56,0920	59,8630	186	27,1139	30,5610	34,2387
139	0,8191	1,2676	1,7733	187	1,6031	1,8816	2,6788
140	50,0692	50,8330	51,0586	188	36,8667	37,6320	38,5072
141	125,0276	128,7800	132,5724	189	36,7970	38,1130	39,0747
142	53,7188	55,2940	55,8926	190	56,2632	58,6600	62,3732
143	49,7414	49,9450	51,4330	191	67,9071	68,7340	69,1425
144	13,7813	14,5890	15,1777	192	69,7602	71,2360	72,1969
145	-51,8983	-50,6480	-50,1392	193	52,5765	54,4470	55,9713
146	-26,4225	-25,6230	-24,4497	194	65,6412	66,3270	66,9453
147	-7,9476	-6,8906	-1,2392	195	25,8820	27,0980	27,6234
148	48,0476	49,7130	54,3213	196	21,3974	23,8010	27,0463
149	64,8451	67,9980	68,3859	197	52,3261	55,1120	55,3100
150	96,9400	100,1500	101,6604	198	178,6719	181,7000	*181,4382
151	128,1275	129,8300	134,4261	199	187,7565	190,9700	194,8155
				200	202,7912	203,0400	203,4344

\* Valori degli estremi dell'intervallo che rimangono al di sopra o al di sotto del valore teorico non includendolo.

Nella tabella 7 sono riportati gli intervalli di confidenza della stima dei valori del parametro variabile e per facilitare la valutazione sulla *bontà* delle stime ottenute si sono inclusi i valori teorici dei parametri variabili. I casi in cui gli intervalli di confidenza non contengono il valore teorico sono 28, superiore alla decina di casi che ci si poteva aspettare sulla base del coefficiente di confidenza fissato. Ciò costituisce ulteriore evidenza di un leggero effetto distorsivo oppure un'indicazione che lo stimatore della varianza dello stato è sottostimata. In ogni caso, il risultato di stima è soddisfacente poiché in una rappresentazione grafica il profilo del parametro variabile stimato e quello teorico risultano praticamente indistinguibili.

La grande flessibilità parametrica del modello porta naturalmente a ritenere ugualmente ottimali valori stimati degli iperparametri in un intorno ampio dei veri valori che hanno generato i dati. Un aspetto rilevante della ricerca potrebbe riguardare i limiti entro cui le stime degli iperparametri possono variare congiuntamente senza creare distorsioni nelle stime dei parametri contenuti nello stato o nella sua matrice di covarianza.

### Riferimenti bibliografici

- Amemiya, T. (1985), *Advanced Econometrics*, Oxford, Basil Blackwell.
- Carraro, C. (1988), *Square Root Kalman Algorithms in Econometrics*, in «Computer Science for Economics and Management», n. 1, pp. 41-51.
- Carraro, C. e Sartore, D. (1984), *Iterative Kalman Filter: Theory and Application to Regression Models*, Note di lavoro N. 8408, Dipartimento di Scienze Economiche, Università di Venezia «Ca' Foscari», ottobre.
- (1987), *Square Root Iterative Filter: Theory and Applications to Econometric Models*, in «Annales d'Économie et de Statistique», n. 6/7, pp. 435-459.
- Chow, G. C. (1983), *Econometrics*, New York, McGraw-Hill Book Company.
- Engle, R. F., Hendry, D. F. e Richard, J. F. (1983), *Exogeneity*, in «Econometrica», 51, pp. 277-304.
- Goldfeld, S. M. e Quandt, R. E. (1978), *Asymptotic Tests for the Constancy of Regressions in the Heteroschedastic Case*, Research Memorandum no. 229, Econometric Research Program, Princeton University.
- Sartore, D. (1990), *Flexibility of Time Varying State Space Models*, Dipartimento di Scienze Economiche, Università di Venezia «Ca' Foscari», dattiloscritto.

Abstract

## **Some Specification and Estimation Problems in Multi-Equation Time Varying Parameter Models**

by Carlo Carraro and Domenico Sartore

After having reviewed appropriate solutions to the numerical problems arising when estimating the hyper-parameters of the state space representation of a time varying parameter model, this paper identifies two new problems. The first one, called «excess of structural flexibility» is shown to arise whenever the estimation procedure leads to a singular covariance matrix of the model errors. The second problem, called «excess of parameter sensitivity» consists in a high dependence of the estimated hyper-parameters from the initial values used in the iterative estimation procedure. Solutions to the problems are explored, and preliminary Monte-Carlo experiments are proposed to evaluate the performance of the solutions.