

STIMA PSEUDOLINEARE DI MODELLI ARMA MULTIVARIATI

Carlo Grillenzoni

Dip. di Economia Politica, Univ. di Modena, Dip. di Scienze Statistiche, Univ. di Padova.

L'obiettivo di definire algoritmi di stima semplificati per modelli non lineari, tende a rispondere ad una molteplicità di esigenze: di agibilità ed adattabilità dello stimatore, di semplicità dei programmi di calcolo, di soluzione del problema dei valori iniziali, ecc.. In questo lavoro viene definito un algoritmo per modelli ARMA multivariati che si basa sulla approssimazione del gradiente, relativo allo stimatore classico dei minimi quadrati non-lineari, con le quantità input-output del sistema. Ciò è permesso dalla particolare struttura degli ARMA, la cui non linearità è relativa ai parametri ma non alle variabili. In termini intuitivi, disponendo di stime iniziali dei parametri e/o generando i "residui" (input del sistema), questi possono essere trattati come regressori osservabili (pseudolineari); l'algoritmo si sviluppa applicando iterativamente il metodo dei minimi quadrati ordinari.

1. INTRODUZIONE

La genesi della struttura dei modelli autoregressivi a media mobile multivariati $ARMA_m(p, q)$ risale ai lavori di Whittle (1953) e Quenouille (1957); il problema della loro stima lo si è cominciato ad affrontare sistematicamente con i lavori di Whittle (1961) e Hannan (1970), con particolare rilievo allo studio delle proprietà degli stimatori nel dominio delle frequenze.

Nel corso degli anni '70 il problema della stima è stato sviluppato nel dominio del tempo soprattutto nel contesto della massima verosimiglianza con i lavori di Wilson (1973), Nicholls (1976), Osborn (1977), Phadke-Kedem (1978) e Hillmer-Tiao (1979); l'obiettivo principale è stato la ricerca, nel contesto gaussiano, della funzione di verosimiglianza esatta.

Negli anni '80 si sono finalmente messi a punto algoritmi e pacchetti organici di stima. Data la difficoltà a massimizzare funzioni di verosimiglianza così complicate e/o ad esplicitarne lo stimatore relativo, la soluzione è stata trovata negli algoritmi dei minimi quadrati non lineari, opportunamente integrati da procedure di *backforecasting* per il calcolo incondizionato della somma dei residui al quadrato (funzione obiettivo), si veda Tiao-Box (1981), Jenkins-Alavi (1981).

I risultati ottenuti non sono comunque soddisfacenti. Per ordini (p, q) e dimensioni (m) superiori a 2, le routine di stima possono bloccarsi per problemi di valori iniziali dei parametri, matrici da invertire per il calcolo del gradiente, problemi generali di convergenza. Il problema di definire algoritmi *semplificati* tende prioritariamente a rispondere a esigenze minimali di agibilità dello stimatore, ma è suscettibile di importanti sviluppi metodologici connessi con la soluzione generale del problema dei valori iniziali.

In questo lavoro viene definito un algoritmo estremamente semplice e naturale che estende a livello *iterativo e multivariato* gli stimatori *ricorsivi* definiti a Regressione Pseudo-Lineare, già affrontati dal punto di vista algoritmico e di analisi delle proprietà asintotiche da Solo (1981), Ljung-Soderstrom (1983) per il caso di modelli *ARMAX* univariati. Il metodo di stima è fondamentalmente quello dei minimi quadrati non-lineari ed è caratterizzato dalla approssimazione del gradiente con le quantità input-output del sistema *ARMA*_{*m*}. Sarà la bontà di questa approssimazione, caratterizzabile in termini di funzione di trasferimento del sistema, che determinerà le proprietà asintotiche dello stimatore.

2. LA STRUTTURA DEGLI *ARMA*_{*m*}

In questa sezione vengono prese in rassegna e brevemente commentate le principali ipotesi e proprietà che stanno alla base dei modelli *ARMA*_{*m*}. Si consideri un processo stocastico vettoriale Gaussiano, con media nulla, stazionario in covarianza ed ergodico (sommabile in media quadratica). In notazioni

$$\underline{z}_t \sim N_m(\underline{0}, \Gamma_k), \quad \Gamma_k = E[\underline{z}_t, \underline{z}'_{t-k}], \quad \sum_k |\Gamma_k| < \infty \quad (2.1)$$

dove

$$\underline{z}_t = [z_{1t}, z_{2t}, \dots, z_{mt}], \quad \Gamma_k = \{\gamma_{ij}(k)\}, \quad |\Gamma_k| = \left(\sum_i \sum_j |\gamma_{ij}(k)|^2 \right)^{1/2}$$

e la sua rappresentazione lineare *ARMA*_{*m*} di ordine (p, q) :

$$\underline{z}_t = \sum_{k=1}^p \Phi_k \underline{z}_{t-k} - \sum_{k=1}^q \Theta_k \underline{e}_{t-k} + \underline{e}_t, \quad \underline{e}_t \sim IN_m(\underline{0}, \Sigma) \quad (2.2)$$

dove $\underline{e}_t = [e_{1t}, \dots, e_{mt}]$ e $\Phi_k = \{\phi_{ijk}\}$, $\Theta_k = \{\theta_{ijk}\}$ sono matrici *m.m* di coefficienti.

Introducendo l'operatore ritardante *B* (dove $B^k \underline{z}_t = \underline{z}_{t-k}$), l'espressione (2.2) può essere scritta in forma polinomiale $\Phi(B)\underline{z}_t = \Theta(B)\underline{e}_t$, vale a dire

$$(I - \Phi_1 B - \dots - \Phi_p B^p) \underline{z}_t = (I - \Theta_1 B - \dots - \Theta_q B^q) \underline{e}_t \quad (2.3)$$

sulla quale assumiamo le condizioni di *stazionarietà invertibilità ed identificabilità parametrica*

$$\Phi(B), \Theta(B) \text{ invertibili} \quad (2.4a)$$

$$\text{Det}\Phi(B) \neq 0 \text{ in } |B| \leq 1 \quad (2.4b)$$

$$\text{Det}\Theta(B) \neq 0 \text{ in } |B| \leq 1 \quad (2.4c)$$

$$\Phi(B), \Theta(B) \text{ prime} \quad (2.5a)$$

$$\text{Rango} [\Phi_p, \Theta_q] = m \quad (2.5b)$$

$$\Phi(0) = \Theta(0) = I_m, \quad \Sigma = \Sigma' > 0 \quad (2.5c)$$

con B intesa come variabile complessa.

Le condizioni di *stazionarietà* (2.4 a, b) consentono di esprimere il modello in forma MA_m razionale, attraverso sviluppo in serie, in forma MA_m infinita

$$\underline{z}_t = \Phi(B)^{-1} \Theta(B) \underline{e}_t = \sum_{k=1}^{\infty} \Psi_k \underline{e}_{t-k} + \underline{e}_t \quad (2.6)$$

Procedendo come nella analisi univariata (si veda Box-Jenkins (1970), l'algoritmo di calcolo della sequenza $\{\Psi_k\}$ è definito da

$$\Psi_k = \sum_{j=1}^p \Phi_j \Psi_{k-j} - \Theta_k \quad k = 1, 2, 3, \dots$$

(dove $\Psi_k = 0, k < 0$ e $\Theta_k = 0, k > q$), per cui la sequenza tende a zero asintoticamente solo se $\text{Det} \Phi(B)$ ha radici all'esterno del cerchio unitario (condizione (2.4 b)). Osserviamo infatti che $\{\Psi_k\}$ può anche essere ottenuta in modo *disaggregato*, sviluppando in serie i singoli polinomi $\Psi(B)$

$$\Psi_{ij} = \tilde{\phi}(B) / \text{Det}\Phi(B) = \sum_{k=0}^{\infty} \Psi_{ijk} B^k$$

Analogamente alla analisi univariata, le proprietà del secondo ordine delle rappresentazioni (2.2), (2.6) sono date rispettivamente da

$$\Gamma_k = \sum_{j=1}^p \Phi_j \Gamma_{k-j} - \sum_{j=0}^q \Theta_j \Sigma \Psi'_{j+k} = \sum_{j=0}^{\infty} \Psi_j \Sigma \Psi'_{j+k} \quad (2.7)$$

allora $\{\Gamma_k\}$ tende a zero asintoticamente (cioè $\{\underline{z}_t\}$ è ergodico solo se $\{\Psi_k\}$ è convergente (cioè le condizioni di stazionarietà sono rispettate).

Le condizioni di *invertibilità* (2.4 a, c) consentono di rappresentare il sistema in forma AR_m razionale e quindi in forma AR_m infinita

$$\Theta(B)^{-1} \Phi(B) \underline{z}_t = \Pi(B) \underline{z}_t = \underline{e}_t, \quad \underline{z}_t = \sum_{k=1}^{\infty} \Pi_k \underline{z}_{t-k} + \underline{e}_t \quad (2.8)$$

l'algoritmo di calcolo della sequenza $\{\Pi_k\}$ è definito da

$$\Pi_k = \sum_{j=1}^q \Theta_j \Pi_{k-j} - \Phi_k$$

(in cui $\Pi_k = 0, k < 0$ e $\Phi_k = 0, k > p$), che converge se $\text{Det } \Theta(B)$ ha radici all'esterno del cerchio unitario (2.4c). A livello interpretativo, la rappresentazione (2.8) permette di collegare le ipotesi (2.1) e (2.2); infatti per la linearità della regressione in ambito gaussiano abbiamo

$$\hat{z}_{t-1} = E\left[\underline{z}_t | \underline{z}_{t-1}, \underline{z}_{t-2}, \dots\right] = \sum_{k=1}^{\infty} \Pi_k^* \underline{z}_{t-k}$$

che corrisponde a $(\underline{z}_t - \underline{e}_t)$, inoltre in condizioni di identificazione $\Pi_k^* = \Pi_k$.

La definizione di rappresentazioni AR_m e MA_m infinite e delle relative proprietà di invertibilità e stazionarietà non risponde solo a requisiti formali. La prima, esplicitando la rappresentazione \underline{e}_t (non osservabile) in termini di \underline{z}_t , è determinante in fase di stima poiché il gradiente è funzione di $\Theta(B)^{-1}$; la seconda ha una validità in fase di previsione in quanto la varianza del previsore è data da (si veda Jenkins-Alavi (1981)) $\hat{\Sigma}(l) = \Sigma + \sum_{k=1}^{l-1} \Psi_k \Sigma \Psi_k'$.

Accenniamo ora alle condizioni di identificabilità parametrica. Utilizzando la (2.7) l'espressione generatrice delle covarianze diventa

$$\Gamma(B) = \sum_{k=-\infty}^{+\infty} \Gamma_k B^k = \Psi(B) \Sigma \Psi(B^{-1})' \quad (2.9)$$

Daltronde, per il teorema di fattorizzazione spettrale multivariato di Rozanov (1967) p. 47, è vero il converso; data cioè la espressione parametrica di $\Gamma(B)$, sotto condizioni generali di regolarità, esiste un algoritmo di calcolo che fattorizza $\Gamma(B)$ come in (2.9). Ora poiché in ambito gaussiano gli stimatori dei parametri del sistema (Ψ_k) sono funzioni dei momenti del secondo ordine del processo (Γ_k) , il sistema (2.2) è *identificato* (cioè univocamente stimabile) se la fattorizzazione (2.9) è *unica*. Hannan (1969, 1979) ha mostrato per sistemi stazionari, che condizioni sufficienti a questo fine sono fornite dalle (2.5).

La condizione cruciale è data dalla (2.5b). È importante però rimarcare che per sistemi *specializzati* $ARMA_m(P, Q)$ con $P=\{p_{ij}\}$, $Q=\{q_{ij}\}$ matrici di ordini dei polinomi $\phi_{ij}(B)$, $\theta_{ij}(B)$, la (2.5b) diventa $\text{Rango}[\Phi(\max p_{ij}), \Theta(\max q_{ij})] = m$, per cui se i massimi sono unici ed $m > 2$, la condizione non è *mai* verificata.

3. STIMA PSEUDOLINEARE ITERATIVA

Gli algoritmi di regressione pseudolineare (*PLR*), anche noti con la definizione di massima verosimiglianza approssimata (*AML*) o minimi quadrati estesi (*ELS*), rappresentano una classe di stimatori *ricorsivi* ampiamente utilizzati in campo ingegneristico per la identificazione–stima di sistemi *ARMA*, *ARMAX*, *TF* (funzioni di trasferimento); si veda Ljung–Soderstrom (1983). La loro derivazione, risalente agli anni '60, è ottenuta dagli stimatori dei minimi quadrati non–lineari (*NLS*); in questa sezione, con riferimento ai sistemi *ARMA*_{*m*}, proponiamo un'estensione dell'approccio all'ambito *multivariato – iterativo*.

Per semplificare la derivazione, considereremo come sistema di riferimento un *ARMA*_{*m*} (1,1) utilizzando una convenzione di segni leggermente diversa

$$\begin{aligned} z_t &= \Phi z_{t-1} + \Theta e_{t-1} + e_t \\ (I - \Phi B) z_t &= (I + \Theta B) e_t \end{aligned} \quad (3.1)$$

I risultati ottenuti saranno comunque generalizzabili, attraverso una complicazione dei calcoli, al caso generale di ordini (p, q) > 1.

Vettorizziamo la matrice dei parametri $B' = [\Phi, \Theta]$, $\beta = \text{Vec} B'$, e consideriamo lo sviluppo in serie multivariato di $\{e_t\}$ attorno ad una stima iniziale $\hat{\beta}$

$$\begin{aligned} e_t(\beta) &\approx \Xi'_t(\hat{\beta})(\beta - \hat{\beta}) + e_t(\hat{\beta}) \\ \Xi'_t(\beta) &= \left[\frac{\partial e_t(\beta)}{\partial \beta'} = \left\{ \frac{\partial e_t(\beta)}{\partial \beta_j} \right\} \right]_{m \cdot 2m^2} \end{aligned} \quad (3.2)$$

dove, dalla rappresentazione (2.8), è facile verificare che ogni *residuo* e_t dipende, direttamente e indirettamente, da ogni parametro del sistema. Il corrispondente stimatore *NLS* diventa allora

$$\hat{\beta}(k+1) = \hat{\beta}(k) + \left[\sum_{t=1}^n \hat{\Xi}'_t(k) \hat{\Xi}_t(k) \right]^{-1} \sum_{t=1}^n \hat{\Xi}'_t(k) \hat{e}_t(k) \quad \text{NLS} \quad (3.3)$$

che tende a minimizzare iterativamente la varianza generalizzata di $\{e_t\}$.

Utilizzando le regole di derivazione delle matrici (si veda Wilson 1973, Spliid 1983), in base all'espressione $e_t = \Theta(B)^{-1} \Phi(B) z_t$ ed alla convenzione di segni prefissata, colonne tipiche di $\Xi'_t(\cdot)$ saranno

$$\frac{\partial e_t}{\partial \phi_{ij}} = -\Theta(B)^{-1} H_{ij} z_{t-1}; \quad \frac{\partial e_t}{\partial \theta_{ij}} = -\Theta(B)^{-1} H_{ij} e_{t-1}$$

dove H_{ij} sono matrici con 1 in posizione ij e 0 altrove. Il calcolo del gradiente ha quindi luogo come filtraggio di quantità input (e_t) ed output (z_t) attraverso il filtro $\Theta(B)^{-1}$

la cui funzione di trasferimento è caratterizzata da $1/\text{Det } \Theta(B)$ (evidentemente se il sistema non è invertibile la stima non può avere luogo).

In questo contesto, la procedura tipica di derivazione dell'algoritmo *PLR* consiste nell'evitare questo filtraggio, ed approssimare il gradiente come

$$\underline{\xi}'_i(\underline{\beta}) = \left[\frac{\partial \theta_i}{\partial (\underline{\phi}'_i, \underline{\theta}'_i)} \right] \approx [-\underline{z}'_{t-1}, -\underline{e}'_{t-1}] = -\underline{x}'_t(\underline{\beta}) \quad i = 1, 2, \dots, m$$

dove $\underline{\phi}'_i = (\phi_i, \dots, \phi_{i_m})$, $\underline{\theta}'_i = (\theta_i, \dots, \theta_{i_m})$. Procedendo in questo modo si avrebbe un *unico* tipo di gradiente approssimato (per riga), corrispondente al vettore dei regressori (pseudolineari) del modello (3.1); il che consente di riscrivere la (3.2) come

$$\underline{e}'_t(B) \approx -\underline{x}'_t(\hat{B})(B - \hat{B}) + \underline{e}'_t(\hat{B})$$

con $\hat{B}' = [\hat{\Phi}, \hat{\Theta}]$, $\underline{x}'_t(\hat{B}) = \underline{x}'_t(\hat{\beta})$ per convenienza posta in forma matriciale e trasposta.

Moltiplicando allora entrambi i membri dello sviluppo per $\underline{x}_t(\cdot)$, è possibile derivare lo stimatore

$$\hat{B}(k+1) = \hat{B}(k) + \left[\sum_{t=1}^n \hat{x}_t(k) \hat{x}'_t(k) \right]^{-1} \sum_{t=1}^n \hat{x}_t(k) \hat{e}'_t(k) \quad (3.4)$$

in cui $\hat{x}'_t(k) = [\underline{z}'_{t-1}, \hat{e}'_{t-1}(k)]$ è generabile al passo (k) attraverso

$$\hat{e}_t(k) = \underline{z}_t - \hat{\Phi}(k) \underline{z}_{t-1} - \hat{\Theta}(k) \hat{e}_{t-1}(k) \quad t = 1, 2, \dots, n \quad (3.5)$$

disponendo di un valore iniziale $\hat{e}_0(k)$.

La forma finale dell'algoritmo *PLR*-iterativo viene allora ricavata sostituendo la (3.5), riscritta come $[\hat{e}'_t(k) = \underline{z}'_t - \hat{x}'_t(k) \hat{B}(k)]$, nella (3.4), ottenendo lo stimatore compatto di regressione multipla

$$\hat{B}(k+1) = \left[\sum_{t=1}^n \hat{x}_t(k) \hat{x}'_t(k) \right]^{-1} \sum_{t=1}^n \hat{x}_t(k) \underline{z}'_t \quad \text{PRL} \quad (3.6)$$

Si noti come lo stimatore sia completamente caratterizzato dalle quantità osservabili \underline{z}_t e facilmente generabili $\underline{e}_t(k)$, disponendo della stima $\hat{B}(k)$.

Riferendosi al sistema (3.1) il procedimento iterativo è il seguente

0) stima di un AR_m [$p^* > (p+q)$]

$$\text{generazione di } \hat{e}_t(0) = \underline{z}_t - \hat{\Phi}_1(0) \underline{z}_{t-1} - \dots - \hat{\Phi}_{p^*}(0) \underline{z}_{t-p^*}$$

1) stima di $ARMA_m(1,1)$: $\underline{z}_t = \Phi(1) \underline{z}_{t-1} + \Theta(1) \hat{e}_{t-1}(0) + \bar{e}_t(1)$

$$\text{generazione di } \hat{e}_t(1) = \underline{z}_t - \hat{\Phi}(1) \underline{z}_{t-1} - \hat{\Theta}(1) \hat{e}_{t-1}(1)$$

2) stima di $ARMA_m(1,1)$: $\underline{z}_t = \Phi(2)\underline{z}_{t-1} + \Theta(2)\hat{\underline{e}}_{t-1}(1) + \tilde{\underline{e}}_t(2)$
 generazione di $\hat{\underline{e}}_t(2) = \dots$

Il dettaglio dell'algoritmo consente le seguenti precisazioni

- 1) il passo (0) è giustificato dal fatto che per inizializzare l'algoritmo occorrerebbe disporre dei valori iniziali dei parametri $\hat{\Phi}(0), \hat{\Theta}(0)$, che nel contesto $ARMA_m$ sono difficilmente ottenibili. Per superare l'ostacolo si genera *direttamente* la sequenza $\hat{\underline{e}}_t(0)$, di cui si ha realmente bisogno per procedere, attraverso una autoregressione di ordine $p^* > p+q$, che al più richiede la specificazione di valori iniziali per \underline{z}_t . Questa *soluzione generale* al problema dei valori iniziali che la metodologia pseudolineare è in grado di fornire, pone la stessa su un piano di preferenza rispetto ad altri metodi, non solo da un punto di vista pratico; ricordiamo infatti che la scelta dei valori iniziali può essere cruciale per la ricerca del minimo globale. In fase di *generazione* il valore iniziale $\hat{\underline{e}}_0(k)$ può essere ottenuto con procedure di *backforecasting* (Jenkins–Alavi, 1981) o semplicemente posto uguale a zero.
- 2) Con riferimento ai *passi* precedenti è spontaneo pensare ad una ulteriore semplificazione dell'algoritmo sostituendo alla stima del processo residuale $\hat{\underline{e}}_t(k)$, il residuo di regressione $\tilde{\underline{e}}_t(k) = \underline{z}_t - \hat{\Phi}(k)\underline{z}_{t-1} - \hat{\Theta}(k)\hat{\underline{e}}_{t-1}(k-1)$. Ciò tuttavia non è metodologicamente corretto ($\tilde{\underline{e}}_t(k)$ non è una approssimazione diretta di $\hat{\underline{e}}_t(k)$) e di fatto introduce una inefficienza che rallenta la convergenza. In questo contesto, per garantire la convergenza numerica (per campioni finiti) verso valori di minimo della funzione obiettivo ($Det \hat{\Sigma}(k)$) è conveniente introdurre un parametro di *stepsize* $0 < \varepsilon(k) < 1$, sostituendo $\hat{B}(k)$ con

$$\hat{B}(k) = \varepsilon(k)\hat{B}(k) + [1 - \varepsilon(k)]\hat{B}(k-1), \quad \varepsilon(k) \rightarrow Det \tilde{\Sigma}(k) \leq Det \hat{\Sigma}(k)$$

Per fini inferenziali, uno stimatore approssimato della matrice di dispersione dell'algoritmo può essere dedotto dalla espressione vettorizzata (3.3), adattando la statistica corrispondente dei minimi quadrati non lineari

$$\hat{\Omega}_n(k) = \left[\sum_t^n \hat{E}_t(k) \hat{\Sigma}(k)^{-1} \hat{E}_t'(k) \right]^{-1}, \quad \hat{\Sigma}(k) = \left[\frac{1}{n} \sum_t^n \hat{\underline{e}}_t(k) \hat{\underline{e}}_t'(k) \right] \quad (3.7)$$

dove $E_t'(\cdot)$ è una matrice ottenuta dalla $E_t'(\cdot)$ approssimandone le colonne con le quantità $[\underline{z}_{t-1}, \underline{e}_{t-1}]$, cioè evitando il filtraggio con $-\Theta(B)^{-1}H_{ij}$.

Benché l'algoritmo *PLR* coincida di fatto con la versione iterativa dello stimatore *OLS* multivariato del sistema (3.1) (si veda Spliid, 1983), la sua derivazione è stata dedotta da quello *NLS*; è proprio a partire da ciò che possono essere definite le sue principali *limitazioni*. In via generale, approssimando il gradiente non

si è più in grado di garantire le proprietà ottimali di *efficienza* e *consistenza* possedute dagli stimatori dei minimi quadrati non-lineari. Come è stato sottolineato recentemente da Hannan-McDougall (1988):

- i) la matrice di dispersione limite tenderà ad avere determinante superiore a quella dei metodi *NLS* o *MLE*, e (3.7) non è uno stimatore consistente per essa;
- ii) la convergenza è assicurata solo quando il polinomio $|\Theta(B)|$ appartiene ad un sottinsieme della regione di invertibilità in cui si comporta come un filtro *passivo*, vale a dire in cui è reale positivo $\Re\{\Theta(B)\} > 0, |B| = 1$.

Se da un lato lo stimatore (3.6) è preferibile a (3.3) per la sua semplicità computazionale e la soluzione del problema dei valori iniziali; dall'altro richiede un'azione continua di monitoraggio.

4. UN ESEMPIO EMPIRICO

Il problema economico considerato per l'analisi empirica verte sull'analisi delle fonti estere dell'inflazione in Italia. Il minimo di teoria economica necessaria a priori per i modelli $ARMA_m$ porta alla scelta delle variabili

\$ = tasso di cambio Lira/Dollaro

PI = indice dei prezzi all'ingrosso

PX = indice dei prezzi all'esportazione

PM = indice dei prezzi all'importazione

B = saldo della bilancia commerciale

t = dati mensili nel periodo 1973.01 – 1985.12.

Dall'analisi dei grafici, dei correlogrammi e delle serie differenziate, si è stabilito che la stazionarietà nei livelli può essere raggiunta, per tutte le serie, con una differenziazione di ordine uno.

Suggeriamo ora una procedura semplificata di specificazione degli ordini (p, q) che ci sembra coerente con la logica sottostante al modello $ARMA_m$. Questo strumento non è l'unico possibile per serie storiche multivariate; un approccio alternativo è fornito dai sistemi di funzioni di trasferimento simultanee (*STF*, si veda Liu (1987)), nei quali attraverso due matrici a polinomi razionali viene fornita una trattazione *specializzata* delle relazioni *auto* e *cross* interne al processo vettoriale. Il modello $ARMA_m$, basandosi su una semplice estensione matriciale dei modelli univariati, non distingue tra le due, e di fatto finisce per considerare la correlazione incrociata una semplice proiezione della autocorrelazione. Ciò può essere giusti-

ficato col fatto che le autorelazioni sono solitamente più potenti e significative, in particolare per le serie economiche. In base a queste considerazioni è possibile definire l' $ARMA_m$ un modello *a priorità di autocorrelazione* per il quale uno schema coerente di specificazione è

$$p = \max_i(p_i) \quad q = \max_i(q_i) \quad i = 1, 2, \dots, m$$

dove (p_i, q_i) sono gli ordini dei modelli univariati delle serie $\{z_{it}\}$.

L'identificazione può quindi avvenire attraverso l'indagine delle sole funzioni di autocorrelazione campionarie, riportate nella Tavola 1. Per ragioni di completezza (analisi delle funzioni di autocorrelazione parziale) e di confronto con i metodi comunemente utilizzati (basati sui criteri di informazione scalari), riportiamo in Tavola-2 anche le matrici delle funzioni di correlazione e correlazione parziale (auto e cross), in termini di segni dei coefficienti significativi (95%). Per una lettura dei valori della Tavola 2 rimandiamo per brevità al lavoro di Tiao-Box (1981), ed al manuale di uso del package WMTS (1980).

Tab. 1: Funzioni di Autocorrelazione Campionarie delle Serie Differenziate.

LAG	\$	PI	PX	PM	B
0	1.00	1.00	1.00	1.00	1.00
1	0.24	0.62	-0.16	-0.09	-0.38
2	0.07	0.41	-0.02	0.08	-0.11
3	0.05	0.28	0.01	-0.18	0.07
4	0.08	0.23	0.04	0.04	-0.09
5	0.12	0.20	0.00	0.09	0.08
6	0.00	0.14	0.11	0.12	-0.11
7	0.01	0.11	-0.01	0.14	-0.04
8	0.06	0.09	-0.03	-0.18	0.16
9	-0.05	0.00	-0.04	-0.04	0.02
10	-0.06	-0.03	0.04	-0.12	-0.21
11	-0.13	0.00	-0.10	0.15	0.03
12	-0.06	0.08	0.09	0.10	0.21
13	0.05	0.02	0.03	-0.06	-0.07
14	0.14	0.01	0.07	-0.05	0.00
15	0.03	-0.04	-0.01	-0.03	-0.12
16	0.08	-0.01	0.00	0.06	0.14
17	-0.03	0.00	0.05	-0.02	-0.01
18	0.02	0.02	-0.05	0.12	-0.15
19	-0.03	0.00	0.04	-0.05	0.05
20	-0.03	0.04	0.01	-0.04	0.06

Tab. II: Matrici delle Funzioni di Correlazione e Corr. Parziale.

LAG		LAG		CHI-SQ TEST
1	+ + + + •	1	+ • • • •	154.41
	• + • + •		+ + • • •	
	• + - • •		+ • - • •	
	• + • • -		• + • - +	
	• • • + -		• • • - -	
2	• + + + •	2	• • • • •	69.52
	• + • • •		- • + • •	
	• + • • •		+ • - + +	
	• + • • •		• • • • •	
	• • • • •		• • • • -	
3	• + • • •	3	• • • • •	38.97
	• + + • •		• • • • +	
	• • • • •		• + • • •	
	• + • - •		• • • - +	
	• • • • •		• • • • -	
4	• • • • •	4	• • • - -	40.43
	• + • • •		• • • • •	
	• • • • •		- • • • •	
	• • • • +		• • • - -	
	- • • - •		• + • • •	
5	• • • • •	5	• • - • -	39.46
	• + • • •		• • • • •	
	• • • • -		• • • • •	
	• • • • •		• • • • •	
	• • • • •		• • • • •	
6	• • • • •	6	+ • - • •	34.38
	• • • • •		• • • + •	
	• • • • +		• • + • •	
	• • • • •		• • • • •	
	• • • • •		• • + • •	

Utilizzando le ben note procedure univariate possiamo individuare

$$\begin{aligned}
 S_t &\sim ARMA(0,1) \\
 PI_t &\sim ARMA(1,0) \\
 PX_t &\sim ARMA(0,0) \quad \Rightarrow \underline{z}_t \sim ARMA_5(1,1) \\
 PM_t &\sim ARMA(0,0) \\
 B_t &\sim ARMA(0,1)
 \end{aligned}$$

Notiamo invece che adottando il metodo *aggregato* basato sul test chi-quadro della Tav. 2 (il cui valore discriminante è $\chi_{0.1}^2(25) = 44,3$), identificheremmo al più un $AR_5(2)$.

L'implementazione dell'algoritmo pseudolineare iterativo è avvenuta con un programma in TSP (1983), stimando la sequenza iniziale $\hat{e}_t(0)$ attraverso un $AR_5(3)$. In una prima fase di stima l'algoritmo non ha avuto convergenza, conseguenza probabilmente della forte correlazione simultanea (PI , PX , PM sono sinonimi), dell'elevato numero di parametri da stimare ($25+25+15$) e della scarsa significatività statistica di molti coefficienti. Questo insieme di situazioni ha probabilmente determinato la non-passività del sistema, benché lo specifico fattore responsabile della divergenza sia stato $\hat{\theta}_{55}(k) \rightarrow 1$ con conseguente non-invertibilità e quindi non-stimabilità. Una semplificazione del modello, ottenuta con l'eliminazione di tutti i coefficienti di regressione non significativi, a partite dalla 3^a iterazione (i valori sono riportati nella Tavola 3), ha risolto il problema. Il modello conseguente (specializzato) è risultato

$$\begin{bmatrix} S_t \\ PI_t \\ PX_t \\ PM_t \\ B_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & \phi_{13} & 0 & 0 \\ 0 & \phi_{22} & 0 & 0 & 0 \\ 0 & 0 & \phi_{33} & 0 & 0 \\ \phi_{41} & 0 & 0 & 0 & \phi_{45} \\ 0 & \phi_{52} & 0 & \phi_{54} & 0 \end{bmatrix} \begin{bmatrix} S_{t-1} \\ PI_{t-1} \\ PX_{t-1} \\ PM_{t-1} \\ B_{t-1} \end{bmatrix} + \begin{bmatrix} \theta_{11} & 0 & \theta_{13} & 0 & 0 \\ 0 & 0 & 0 & \theta_{24} & 0 \\ \theta_{31} & 0 & \theta_{33} & 0 & 0 \\ 0 & 0 & 0 & \theta_{44} & 0 \\ 0 & \theta_{52} & 0 & 0 & \theta_{55} \end{bmatrix} \begin{bmatrix} S_{t-1} \\ PI_{t-1} \\ PX_{t-1} \\ PM_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} S_t \\ PI_t \\ PX_t \\ PM_t \\ b_t \end{bmatrix}$$

La specializzazione compiuta ha comunque giovato della disponibilità dei valori nelle iterazioni successive alla 3^a (nella stima che non ha avuto convergenza), in cui ad esempio il parametro $Q154$ diventava persistentemente non significativo (da ciò la sua esclusione). La fluttuazione dei parametri e della loro significatività, induce quindi a suggerire nella pratica, l'uso di un procedimento di *eliminazione stepwise*, partendo cioè da un livello iniziale *minimo* di significatività (ad esempio $|T| > 1$) da elevare gradualmente, ed eliminando quei coefficienti che in due iterazioni consecutive risultassero non-significativi. In questo contesto, è importante rimarcare che essendo l'algoritmo *PLR* asintoticamente normale, ma non disponendo esplicitamente di uno stimatore per la sua matrice di dispersione, le statistiche T avranno solo approssimativamente una distribuzione di Student.

In 10 iterazioni, fissando $\varepsilon(k) = 1/2$ costante, l'algoritmo sul modello specializzato ha avuto convergenza, i risultati, relativi alla 1^a, 5^a e 10^a iterazione, sono riportati nella Tavola 4. In essa si può notare la buona velocità di convergenza e la buona significatività dei parametri identificati, il che legittima la specificazione degli ordini fatta. È facile inoltre verificare che il modello finale soddisfa alle condizioni di rango (2.5b), in quanto $[\theta_{11}, \phi_{22}, \phi_{33}, \theta_{44}, \theta_{55}]$ sono non-nulli. Nella Tavola 5 vengono infine riportate le matrici di correlazione seriale dei residui per accertare la bontà di adattamento; il fatto che correlazioni incrociate rimangano significative

Tab. III: Valori delle Stime dei ϕ_{ij} (P) e θ_{ij} (Q), alla 3^a iterazione.

PARAMETER	ESTIMATE	STANDARD ERROR	T-STATISTIC
P111	-0.2625453	0.3933353	-0.6674847
P112	2.471131	3.420052	0.7225420
P113	1.233431	0.7934644	1.554488
P114	-0.1315666	0.3160321	-0.4163088
P115	-0.8965824	5.593858	-0.1602798
Q111	0.5749170	0.4003108	1.436177
Q112	-3.173847	4.277372	-0.7420085
Q113	-1.573466	0.8635085	-1.822178
Q114	0.2581107E -01	0.3384544	0.7626158
Q115	-3.998247	5.994725	-0.6669609
P121	0.2023812E -01	0.1260133E -01	1.606030
P122	0.5944460	0.1095686	5.425331
P123	-0.3585186E -01	0.2542031E -01	-1.410363
P124	-0.1177723E -01	0.1012473E -01	-1.163214
P125	-0.2945693	0.1792111	-1.643700
Q121	-0.8464405E -02	0.1282481E -01	-0.6600026
Q122	-0.1127942	0.1370347	-0.8231072
Q123	0.3246641E -01	0.2766433E -01	1.173584
Q124	0.1663317E -01	0.1084311E -01	1.533986
Q125	0.2164691	0.1920537	1.127128
P131	-0.6681422E -01	0.1046720	-0.6383199
P132	-0.4945058E -01	0.9101235	-0.5433392
P133	0.5654177	0.2111519	2.677777
P134	0.1747287	0.8410030E -01	2.077622
P135	0.8560256	1.488603	0.5750528
Q131	0.1747172	0.1065283	1.640102
Q132	0.4975757	1.138268	0.4371339
Q133	-1.076564	0.2297916	-4.684958
Q134	-0.1702902	0.9006744E -01	-1.890697
Q135	0.8989216	1.595280	0.5634884
P141	0.2020696	0.2346927	0.8609966
P142	1.270493	2.040654	0.6225913
P143	0.1385765	0.4734391	0.2927019
P144	0.1633257	0.1885674	0.8661397
P145	3.732422	3.337706	1.118260
Q141	-0.7978778E -01	0.2388548	-0.3340430
Q142	2.519905	2.552194	0.9873486
Q143	-0.9545996E -01	0.5152326	-0.1852755
Q144	-0.4212411	0.2019468	-2.085902
Q145	3.215437	3.576893	0.8989471
P151	0.3897181E -02	0.9183232E -02	0.4243801
P152	0.2413262	0.7984824E -01	3.022311
P153	-0.1471306E -01	0.1852508E -01	-0.7942238
P154	-0.2804179E -01	0.7378407E -02	-3.800520
P155	-0.1522038	0.1306003	-1.165417
Q151	-0.3385519E -02	0.9346090E -02	-0.3622390
Q152	-0.3886034	0.9986417E -01	-3.891319
Q153	0.2325946E -01	0.2016041E -01	1.153720
Q154	0.1831951E -01	0.7901925E -02	2.318361
Q155	-0.5219941	0.1399593	-3.729612

Tab. IV: Stima Finale: Valori alla 1^a, 5^a e 10^a iterazione.

PARAMETER	ESTIMATE	STANDARD ERROR	T-STATISTIC
P113	0.6999944	0.4394516	1.592882
Q111	0.1499801	0.7774836E -01	1.929045
Q113	-1.093421	0.4898954	-2.231947
P122	0.5422503	0.5638126E -01	9.617562
Q124	0.7482799E -02	0.3893003E -02	1.922115
P133	0.2373285	0.1202359	1.973857
Q131	0.6316854E -01	0.2119968E -01	2.979693
Q133	-0.6320811	0.1339712	-4.718036
P141	0.1244213	0.4655212E -01	2.672731
P145	5.085024	1.643155	3.094671
Q144	-0.2273655	0.7571182E -01	-3.003039
P152	0.2870674	0.7712463E -01	3.722124
P154	-0.1250301E -01	0.3183808E -02	-3.927062
Q152	-0.3414334	0.9619338E -01	-3.549448
Q155	-0.4959794	0.7597014E -01	-6.528609
P113	1.249484	0.6207360	2.012907
Q111	0.1319519	0.7842688E -01	1.682483
Q113	-1.659007	0.6810712	-2.435879
P122	0.5031315	0.5818727E -01	8.646762
Q124	0.1007712E -01	0.4069126E -02	2.476482
P133	0.3088814	0.1714068	1.802037
Q131	0.7048449E -01	0.2165049E -01	3.255561
Q133	-0.6857782	0.1881870	-3.644131
P141	0.1242337	0.4815341E -01	2.579956
P145	5.845867	1.649796	3.543389
Q144	-0.1666308	0.7932896E -01	-2.100504
P152	0.1431977	0.7762229E -01	1.844801
P154	-0.1149832E -01	0.3017989E -02	-3.809928
Q152	-0.2177271	0.9833216E -01	-2.214201
Q155	-0.6372842	0.8009182E -01	-7.956920
P113	1.290406	0.6371753	2.025198
Q111	0.1299334	0.7827337E -01	1.659994
Q113	-1.700987	0.7013333	-2.425362
P122	0.5041210	0.5829823E -01	8.647279
Q124	0.1018488E -01	0.4063065E -02	2.506699
P133	0.3358384	0.1750174	1.918886
Q131	0.7284189E -01	0.2149836E -01	3.388254
Q133	-0.7202368	0.1927422	-3.736788
P141	0.1246992	0.4835870E -01	2.578630
P145	5.900052	1.652223	3.570979
Q144	-0.1586370	0.7939901E -01	-1.997972
P152	0.1330060	0.7801474E -01	1.704882
P154	-0.1085531E -01	0.2982047E -02	-3.64-220
Q152	-0.2113040	0.9869105E -01	-2.141066
Q155	-0.6637098	0.8028398E -01	-8.267027

RIFERIMENTI BIBLIOGRAFICI

- Box G.E.P., G.M. Jenkins (1970), *Time Series Analysis: forecasting and control*. San Francisco: Holden-Day.
- Chen H.F., L. Guo (1987), Optimal Adaptive Control and Consistent Parameter Estimates for ARMAX Model with Quadratic Cost. *SIAM J. on Control and Optim.* 25, 845–867.
- Fuller W.A. (1976), *Introduction to Statistical Time Series*. New York: Wiley.
- Hannan E.J. (1969), The Identification of Vector Mixed ARMA Systems. *Biometrika* 56, 223–225.
- Hannan E.J. (1970), *Multiple Time Series*. New York: Wiley.
- Hannan E.J. (1979), The Statistical Theory of Linear Systems; in Krishnaiah P.R. (ed.), *Development in Statistics 2*. New York: Academic Press.
- Hannan E.J., E.D. McDougall (1988), Regression Methods for ARMA Estimation. *J.A.S.A.* 83, 875–892.
- Hillmer S.C., G.C. Tiao (1979), Likelihood Function of Stationary Multiple ARMA Models, *J.A.S.A.* 74, 652–660.
- Jenkins G.M., A.S. Alavi (1981), Some Aspects of Modeling and Forecasting Multivariate Time Series. *J.T.S.A.* 2, 1–49.
- Lijng L., T. Soderstrom (1983), *Theory and Practice of Recursive Identification*. Cambridge (Mass.): M.I.T. Press
- Liu L.M. (1987), Sales Forecasting Using Multi-equation Transfer Function Models. *Jour. of Forecasting* 6, 223–238.
- Nicholls D.F. (1976), The Efficient Estimation of Vector Linear Time Series Models. *Biometrika* 63, 381–390.
- Osborn D.R. (1977), Exact and Approximate Maximum Likelihood Estimators for Vector MA Processes. *J.R.S.S. – B* 39, 114–118.
- Phadke M.S., G. Kedem (1978), Computation of the Exact Likelihood Function of Multivariate MA Models. *Biometrika* 65, 511–520.
- Quenouille M.H. (1957), *The Analysis of Multiple Time Series*. London: Griffin.
- Rozanov Y.A. (1967), *Stationary Stochastic Processes*. San Francisco: Holden-Day.
- Solo V. (1981), The Second Order Properties of a Time Series Recursion. *Ann. of Statistics* 9, 307–317.
- Spliid H. (1983), A Fast Estimation Method for vector ARMAX Models. *J.A.S.A.* 78, 843–849.
- Tiao G.C., G.E.P. Box (1981), Modeling Multiple Time Series with Application. *J.A.S.A.* 76, 802–816.
- TSP (1983), *Time Series Processor Package, User's Guide*. San Francisco: TSP Inc.
- Whittle P. (1953), The Analysis of Multiple Time Series. *J.R.S.S. – B* 14, 125–139.
- Whittle P. (1961), Gaussian Estimation in Stationary Time Series. *Bull. Int. Stat. Inst.* 33, 1–26.
- Wilson T.G. (1973), The Estimation of Parameters in Multivariate Time Series Models. *J.R.S.S. – B* 35, 76–85.
- WMTS (1980), *Wisconsin Multiple Time Series Package, User's Guide*. Madison: Univ. of Madison, Dep. of Statistics.

SUMMARY

Pseudolinear Estimation of Multivariate ARMA Models – The definition of simplified estimation algorithms for non-linear models tends to solve many problems: adaptability of the algorithm, speed of calculation, simplicity of the computer programs, solution of the initial value problem, and so on. In this work we propose an algorithm for multivariate ARMA models which approximate the gradient, in the classical non-linear least squares estimator, with the input-output quantities of the system. This may occur by the particular structure of the ARMA, which is non-linear in the parameters but not in the variables. Intuitively, given an initial estimate of the parameters and/or generating the "residuals" (the input of the system), these may act like observable regressors (pseudolinear); the algorithm arises, in this context, by applying iteratively the ordinary least squares methods.