Rivista di Statistica Applicata, Vol. 17, N. 1 (Marzo) 1984

Analisi canonica e capacità previsiva per processi multivariati

Domenico Piccolo, Centro di Specializzazione e Ricerche Economico-Agrarie per il Mezzogiorno, Portici, Università di Napoli

La capacità previsiva di un modello lineare assume spesso il ruolo di criterio decisionale per una pluralità di obiettivi statistici: identificazione degli ordini di un modello, verifica sui parametri, analisi della causalità e così via. In tale ambito, questo articolo propone un indice multivariato per processi ARMA vettoriali di cui si illustrano significato e proprietà, soprattutto per la sua esplicita connessione con la trasformazione canonica di cui rappresenta una utile sintesi. Infine, il lavoro si conclude con una breve rassegna di alcune proposte complementari presenti nella letteratura.

1. Introduzione

Una importante utilizzazione della modellistica econometrica consiste nella previsione del comportamento futuro di variabili oggetto di politica economica sia per obiettivi di controllo che per stabilire nessi di causalità nel senso di Granger. Per questo, la capacità previsiva diviene molto spesso un criterio di giudizio della bontà di un modello e un criterio di scelta fra modelli alternativi, come si ritrova nelle proposte di Akaike e nelle innumerevoli varianti.

In questo tipo di analisi, tuttavia, occorrerebbe discriminare fra una capacità intrinseca ai dati ed un'altra di natura metodologica, derivante dalla bontà del modello prescelto e dalla sua flessibilità in rapporto alla variabilità delle osservazioni. In breve, se i dati costituiscono realizzazioni di numeri casuali non vi è capacità intrinseca per prevedere la successiva osservazione pur esistendo, metodologicamente, un previsore ottimale nel senso dei minimi quadrati.

Questo discorso, in termini di analisi di fenomeni economici, implica che alcune serie — sin dal grafico — evidenziano un elevato contenuto previsivo, anche con metodologie appena sufficienti: ciò è tipico di variabili evolutive, con stagionalità più o meno regolare, nell'ipotesi di accrescimento di popolazioni, e così via. Peraltro, esistono serie economiche nelle quali la irregolarità dei punti di svolta, la non costanza dei cicli, l'alternanza attorno al punto di equilibrio, ecc. non consentono di pervenire, anche con metodologie raffinate, a previsioni comparabili con quelle precedenti: tale situazione è tipica dei fenomeni stazionari o resi tali, dei tassi di variazione, dei residui da relazioni causali e così via.

Se quindi è compito dello statistico cercare, in ogni caso, la individuazione di previsori ottimali per una data struttura, occorre anche considerare quanta capacità previsiva i dati contengano intrinsecamente, al di là di qualsiasi sforzo metodologico per costruire strumenti efficienti.

Sul piano delle analisi univariate, per esempio, abbiamo evidenziato come la modellistica ARMA e/o qualche conveniente trasformazione consenta di pervenire alla più efficiente parametrizzazione di una serie storica (si veda Piccolo (1981) per questa discussione) ed abbiamo sottolineato come per tali strutture esista un indice della capacità previsiva che è determinato sia dai parametri sia dalla natura del modello da cui i dati sono, o si pensano siano, generati. Questa ricerca ha condotto ad una misura analoga all'indice di determinazione multipla R^2 della regressione, di cui in Piccolo (1982) abbiamo discusso proprietà e caratteristiche.

Questo articolo, che per certi aspetti costituisce una generalizzazione di quei risultati, si propone di approfondire la definizione di un indice analogo per $\mathbf{Z}_t \sim$ ARMA vettoriale: si vedrà come questa misura sia immediatamente riconducibile alla trasformazione canonica proposta da Box e Tiao e, come tale, operativamente utilizzabile nei packages usuali per l'analisi multivariata delle serie storiche. L'articolo si conclude confrontando questa nostra proposta con altre soluzioni presenti nella letteratura, essendo convinti che solo una pluralità di approcci può consentire uno studio accurato di una struttura multivariata, soprattutto quando trattasi di variabili economiche.

2. La trasformazione canonica

Anche in settori ben diversi dall'analisi di dati temporali, la complessità dei legami multipli ha sempre consigliato delle trasformazioni dei dati in modo da consentire — in spazi o metriche differenti — una migliore interpretazione delle informazioni possedute. Per l'analisi di serie multiple è fondamentale, al riguardo, il lavoro di Box e Tiao (1977) nonché i suggerimenti e le linee di ricerca contenute nel lavoro empirico di Tiao e altri (1979).

Rinviando a tali lavori, ci limitiamo qui ad un breve richiamo circa gli autovalori e gli autovettori connessi alla trasformazione in variabili canoniche.

$$\text{Siano } \Gamma_0 = \Gamma(0) = E(\mathbf{Z}_t \mathbf{Z}'_t) \quad \text{e} \quad \Gamma_1 = \Gamma_1(0) = E(\tilde{\mathbf{Z}}_{t-1}(1)\, \tilde{\mathbf{Z}}'_{t-1}(1))$$

le matrici di varianze e covarianze di $\mathbf{Z}_t$ e del previsore lineare per $\mathbf{Z}_t$ effettuato in $t - 1$, cioè

$$\mathbf{Z}_t = \tilde{\mathbf{Z}}_{t-1}(1) + \mathbf{a}_t, \text{ ove } \tilde{\mathbf{Z}}_{t-1}(1) = E(\mathbf{Z}_t \mid \mathbf{Z}_{t-1}, \mathbf{Z}_{t-2} \ldots), \; \mathbf{a}_t \sim \mathrm{WN}(\mathbf{O}, \tilde{\Sigma}) .$$

Per l'ortogonalità fra previsori e innovazioni si ha: $\Gamma_0 = \tilde{\Gamma}_1 + \tilde{\Sigma}$.

Quando $\mathbf{Z}_t \sim \mathrm{AR}(p)$ allora $\tilde{\mathbf{Z}}_{t-1}(1) = \Sigma_{j=1}^{p} \Phi_j \mathbf{Z}_{t-j}$; tuttavia, quanto diremo per $\mathbf{Z}_t \sim \mathrm{AR}(p)$ può essere agevolmente esteso per qualsiasi $\mathbf{Z}_t$, trasformato in modo da essere modellabile mediante una struttura ARMA (p, q) invertibile, per il quale è lecito scrivere:

$$\tilde{\mathbf{Z}}_{t-1}(1) = \sum_{j=1}^{\infty} \pi_j \mathbf{Z}_{t-j} .$$

Siano $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_m$ l'insieme degli autovalori della matrice $\Gamma_0^{-1}\tilde{\Gamma}_1 = \mathbf{I} - \Gamma_0^{-1}\tilde{\Sigma}$ e sia $\mathbf{M}' = (m_1, \ldots, m_m)$ la matrice degli autovettori corrispondenti. Allora, essendo $\mathbf{Z}_t = \tilde{\mathbf{Z}}_{t-1}(1) + \mathbf{a}_t$, la trasformazione

$$\mathbf{W}_t = \mathbf{M}\,\mathbf{Z}_t \Rightarrow \tilde{\mathbf{W}}_{t-1}(1) = \mathbf{M}\,\tilde{\mathbf{Z}}_{t-1} ; \quad \mathbf{b}_t = \mathbf{M}\,\mathbf{a}_t$$

opera la scissione

$$\mathbf{W}_t = \tilde{\mathbf{W}}_{t-1}(1) + \mathbf{b}_t .$$

Analogamente, in corrispondenza di $\Gamma_{0Z} = \tilde{\Gamma}_{1Z} + \tilde{\Sigma}_a$ (con ovvio significato dei simpoli), la trasformazione via autovettori opera la scissione

$$\Gamma_{0W} = \tilde{\Gamma}_{1W} + \tilde{\Sigma}_b$$

per la quale, come è dimostrato in Box e Tiao (1977; 356-357), risulta

$\Gamma_{0W} = \mathbf{M}\,\Gamma_{0Z}\,\mathbf{M}' = diag\,\{\sigma^2_{W1}, \ldots, \sigma^2_{Wm}\}$; $\sigma^2_{Wi} = Var\,(W_{it})$, $i = 1, \ldots, m$.

$\tilde{\Gamma}_{1W} = \mathbf{M}\,\tilde{\Gamma}_{1Z}\,\mathbf{M}' = diag\,\{\sigma^2_{\tilde{W}1}, \ldots, \sigma^2_{\tilde{W}m}\}$; $\sigma^2_{\tilde{W}i} = Var\,(\tilde{W}_{i(t-1)}(1))$, $i = 1, \ldots, m$

$\tilde{\Sigma}_b = \mathbf{M}\,\tilde{\Sigma}\,\mathbf{M}' = diag\,\{\sigma^2_{b1}, \ldots, \sigma^2_{bm}\}$; $\sigma^2_{bi} = Var\,(b_{it})$, $i = 1, \ldots, m$.

Esiste cioè, per ciascuna variabile canonica W_{it}, $i = 1, \ldots, m$ una decomposizione della varianza

$$Var\,(W_{it}) = Var\,(\tilde{W}_{i(t-1)}(1)) + Var\,(b_{it}),\ \mathrm{i} = 1, \ldots, m .$$

In tal modo, una misura della capacità previsiva della serie i-ma W_{it} è

$$0 \leq s_i = \frac{Var\,(\tilde{W}_{i(t-1)}(1))}{Var\,(W_{it})} \leq 1 .$$

Ora, per noti teoremi di algebra lineare, si ha:

$$\Lambda = diag\,(\lambda_1, \ldots, \lambda_m) = \mathbf{M}'^{-1}\,\Gamma_{0Z}^{-1}\tilde{\Gamma}_{1Z}\,\mathbf{M}'\,;$$

$$\Gamma_{0Z} = \mathbf{M}^{-1}\,\Gamma_{0W}\,\mathbf{M}'^{-1}\,; \quad \tilde{\Gamma}_{1Z} = \mathbf{M}^{-1}\,\tilde{\Gamma}_{1W}\,\mathbf{M}'^{-1}$$

e, quindi,

$$\Lambda = \mathbf{M}'^{-1}\,(\mathbf{M}'\,\Gamma_{0W}^{-1}\mathbf{M})\,(\mathbf{M}^{-1}\,\tilde{\Gamma}_{1W}\,\mathbf{M}'^{-1})\,\mathbf{M}' = \Gamma_{0W}^{-1}\tilde{\Gamma}_{1W}$$

$$= diag\,\{\sigma^{-2}_{\tilde{W}1}, \ldots, \sigma^{-2}_{\tilde{W}m}\} \cdot diag\,\{\sigma^2_{W1}, \ldots, \sigma^2_{Wm}\}$$

$$= diag\,\{s_1, \ldots, s_m\}\,.$$

Pertanto gli autovalori λ_i, $i = 1, \ldots, m$ coincidono e rappresentano ordinatamente misure di capacità previsive s_i, $i = 1, \ldots, m$ delle singole serie trasformate W_{it}, $i = 1, \ldots, m$. Su questo aspetto ritorneremo nel paragrafo successivo. Tali autovalori (variando tra 0 e 1) consentono una graduazione delle componenti, in termini di capacità previsiva, in tre insiemi ben caratterizzati:

1. Autovalori vicino a 0 riflettono relazioni stabili contemporanee tra le originarie Z_{it} e quindi le variabili risultanti variano indipendentemente attorno alla loro media.
2. Autovalori compresi tra 0 e 1 indicano una effettiva e intermedia struttura stocastica fra le componenti che migliora — in parte — la capacità previsiva contenuta nelle serie univariate.
3. Autovalori vicino a 1 indicano componenti quasi non-stazionarie, altamente prevedibili perché $R^2 \to 1$ (v. Piccolo, 1982; pp. 3-11), il cui significato per serie economiche è di rendere conto del *trend*.

In conclusione, l'analisi canonica tramite Λ decompone lo spazio R^m di $\mathbf{Z}_t$ nei sottospazi R^{m_1}, R^{m_2}, R^{m_3} delle componenti sopra descritte, ove

$$R^m = R^{m_1} \cup R^{m_2} \cup R^{m_3}\,, \quad m_1 + m_2 + m_3 = m\,,$$

ed a cui possiamo associare i seguenti significati:

R^{m_1} : sottospazio a componenti (quasi) indipendenti;

R^{m_2} : sottospazio a componenti stazionarie;

R^{m_3} : sottospazio a componenti (quasi) non-stazionarie.

È evidente l'importanza — a fini interpretativi — dell'analisi canonica: esiste cioè una trasformazione $\mathbf{W}_t = \mathbf{M}\,\mathbf{Z}_t$, determinata a partire dagli autovalori di $\mathbf{I} - \Gamma_0^{-1}\,\tilde{\Sigma}$, mediante la quale una combinazione lineare delle serie originarie produce m nuove serie W_{it}, $i = 1, \ldots, m$, ordinate dalla meno prevedibile, contemporaneamente indipendenti così come indipendenti sono i previsori per 1 lag e le innovazioni residue.

Tuttavia, anche se l'applicazione che Box e Tiao (1977; pp. 359-362) conducono sui dati di Quenouille (1957) per il mercato suinicolo statunitense è convincente e particolarmente ben riuscita, è indubbio che le potenzialità di un simile approccio restano ancora da esplorare sia sul piano teorico che, soprattutto, sul piano dell'analisi econometrica.

3. Una misura di capacità previsiva multivariata

L'affermazione secondo cui l'usuale indice di determinazione R^2 è funzione esclusiva dei parametri e non della variabilità dei processi trovasi in Piccolo (1974a; 1974b) per i processi AR ed in Nelson (1976) per i processi lineari. Entrambi cautelano circa i rischi di una interpretazione di tale misura secondo i criteri usuali della regressione. Senza riferimenti a tale problematica, occorre ricordare comunque che misure e grafici di fatto equivalenti all'indice R^2 per modelli AR (2) si trovano già in Stone (1949) e in Stralkowski, Wu e DeVor (1970).

In Piccolo (1982), sulla base della ortogonalità fra previsore per 1 periodo ed innovazioni, viene proposto di misurare la capacità previsiva di qualsiasi processo univariato $Z_t \sim$ ARMA, stazionario ed invertibile, mediante

$$R^2 = 1 - Var\,(a_t)/Var\,(Z_t)\ ,$$

dimostrando poi, per tale indicatore, proprietà, teoremi e campi di variazione.

Sulla base della scissione prima esaminata per processi multivariati, una generalizzazione di tale misura ci sembra immediata tenendo conto che $|\tilde{\Sigma}|$ e $|\tilde{\Gamma}_0|$ sono le corrispondenti misure di variabilità multipla per $Var\,(a_t)$ e per $Var\,(Z_t)$, rispettivamente. Definiamo, quindi, misura della capacità previsiva del processo $\mathbf{Z}_t$ l'indice

$$F^2 = 1 - \frac{|E\,(\mathbf{a}_t\,\mathbf{a}_t')|}{|E\,(\mathbf{Z}_t\,\mathbf{Z}_t')|} = 1 - \frac{|\tilde{\Sigma}|}{|\Gamma_0|} = 1 - |\Gamma_0^{-1}\,\tilde{\Sigma}|\ .$$

Per i risultati del paragrafo precedente, se $\Lambda = diag\,\{\lambda_1, \ldots, \lambda_m\}$ è l'insieme degli autovalori di $\mathbf{I} - \Gamma_0^{-1}\,\tilde{\Sigma}$, allora gli autovalori di $\Gamma_0^{-1}\,\tilde{\Sigma}$ saranno contenuti in $\mathbf{I} - \Lambda = diag\,\{1 - \lambda_1, \ldots, 1 - \lambda_m\}$.

Pertanto, risulta subito che

$$F^2 = 1 - |\mathbf{I} - \Lambda| = 1 - (1 - \lambda_1) \ldots (1 - \lambda_m) \,,$$

il che dimostra come l'indice F^2 — oltre ad essere una misura della capacità previsiva multivariata — possa essere utilizzato anche per sintetizzare i risultati dell'analisi canonica.

Si noti che

$F^2 \to 0$ se $\lambda_i \to 0$ per tutti gli $i = 1, \ldots, m$;

$F^2 \to 1$ se almeno un $\lambda_i \to 1$, per $i = 1, \ldots, m$.

In tal senso, $F^2 \to 0$ è indicazione di una totale imprevedibilità delle m componenti costituenti il processo multivariato, mentre $F^2 \to 1$ è indicazione che almeno una di esse è totalmente e linearmente prevedibile a partire dalle altre (o da un sottoinsieme delle altre): in questa seconda ipotesi, occorre eliminare tale componente (o tali componenti) e ricalcolare F^2 per l'insieme delle componenti rimanenti, e così via.

Particolarizziamo, ora, qualcuno dei risultati precedenti, dimostrando, anzitutto, alcuni lemma:

Lemma 1.

$\tilde{\Gamma}_0 - \tilde{\Sigma}$ è definita positiva.

Dim. Difatti, seguendo Dhrymes (1970; pag. 582-583), gli autovalori di $\Gamma_0^{-1} \tilde{\Sigma}$ sono inferiori ad 1. Alternativamente, possiamo osservare che, per ogni processo stazionario,

$$\Gamma_0 - \tilde{\Sigma} = \sum_{j=1}^{\infty} \psi_j \tilde{\Sigma} \psi_j'$$

è certamente convergente e, quale somma di matrici definite positive, è definita positiva.

Lemma 2.

Se $\mathbf{A} - \mathbf{B}$ è definita positiva, allora $|\mathbf{A}| > |\mathrm{B}|$ quando $\mathbf{A}$ e $\mathbf{B}$ sono definite positive.

Dim. Si veda Dhrymes (1970, pag. 584).

Lemma 3.

Siano $\mathbf{X}_t \sim$ AR (1) e $\mathbf{Y}_t \sim$ MA (1) due processi multivariati, rispettivamente, stazionari e invertibili, con identiche matrici $\Phi_1 = \Theta_1 = \varphi$; Σ.

Allora, si ottiene:

$$\Gamma_{0X} = \tilde{\Sigma} + \varphi \, \Gamma_{0X} \, \varphi' \; ; \quad \Gamma_{0Y} = \tilde{\Sigma} + \varphi \, \tilde{\Sigma} \, \varphi'$$

ove $\quad \tilde{\Sigma} = E\,(\mathbf{a}_t\,\mathbf{a}_t')\;,\; \Gamma_{0Y} = E\,(\mathbf{Y}_t\,\mathbf{Y}_t')\;;\quad \Gamma_{0X} = E\,(\mathbf{X}_t\,\mathbf{X}_t')\,.$

Dim. Segue immediatamente dalle definizioni.

Teorema

Siano F_A^2, F_M^2 gli indici calcolati, rispettivamente, per $\mathbf{X}_t \sim AR\,(1)$, $\mathbf{Y}_t \sim MA\,(1)$ sopra definiti. Allora,

$$F_A^2 > F_M^2 .$$

Dim. È immediato dedurre che

$$F_A^2 > F_M^2 \Leftrightarrow |\tilde{\Sigma} + \varphi\,\Gamma_{0X}\,\varphi'| > |\tilde{\Sigma} + \varphi\,\tilde{\Sigma}\,\varphi'|\,,$$

che è quanto cercheremo di dimostrare. Per il lemma 1, $\Gamma_{0X} - \tilde{\Sigma}$ è definita positiva, così come $\varphi\,(\Gamma_{0X} - \tilde{\Sigma})\,\varphi'$. Quindi,

$$|\varphi\,\Gamma_{0X}\,\varphi' - \varphi\,\tilde{\Sigma}\,\varphi'| = |(\varphi\,\Gamma_{0X}\,\varphi' + \tilde{\Sigma}) - (\varphi\,\tilde{\Sigma}\,\varphi' + \tilde{\Sigma})| > 0\,.$$

Ma entrambe le matrici in parentesi nel determinante sono definite positive; quindi, per il lemma 2, sarà

$$|\varphi\,\Gamma_{0X}\,\varphi' + \tilde{\Sigma}| > |\varphi\,\tilde{\Sigma}\,\varphi' + \tilde{\Sigma}|\,. \qquad \text{c.v.d. (*)}$$

4. Confronti e sviluppi ulteriori

Ci sembra utile esporre qui una breve rassegna di lavori complementari rispetto alla problematica di questo articolo, essendo le tematiche qui solo accennate bisognevoli della integrazione con approcci alternativi.

Un quadro di riferimento significativo per collocare il tema delle previsioni nel contesto dei modelli ARMA è offerto dal recente lavoro di Mazzali (1982) il quale, confermando ed approfondendo nell'ipotesi lineare alcuni risultati generali di Priestley (1981), dimostra come « la dimensione dello spazio di previsione di un modello ARMA (p, q) è $r = \max\,(p, q)$ ». Non a caso, Mazzali dimostra che l'indice R^2 da noi proposto è proprio il primo coefficiente di correlazione canonica della matrice che contiene le informazioni rilevanti sulle previsioni univariate. A noi sembra assai probabile che tale risultato possa essere generalizza-

(*) Una dimostrazione alternativa deriva da $\Gamma_{0x} = \tilde{\Sigma} + \varphi\,\Gamma_{0x}\,\varphi' = \tilde{\Sigma} + \varphi\,(\tilde{\Sigma} + \varphi\,\Gamma_{0x}\,\varphi')\,\varphi' = \tilde{\Sigma} + \varphi\,\tilde{\Sigma}\,\varphi' + \varphi^2\,\Gamma_{0x}\,(\varphi^2)' = \Gamma_{0y} + \varphi^2\,\Gamma_{0x}\,(\varphi^2)'$, dal che si deduce che $\Gamma_{0x} - \Gamma_{0y}$ è definita positiva, e quindi $F_A^2 > F_M^2$.

to nel caso multivariato confermando le connessioni fra F^2, trasformazione canonica e capacità previsiva.

Naturalmente, la proposta qui sviluppata è suscettibile di varianti in più direzioni sia sul piano generale dell'analisi econometrica che su quello più statistico dell'analisi delle serie storiche.

Circa il primo aspetto, potrebbero applicarsi al processo $\mathbf{Z}_t \sim$ ARMA gli indici tipici della regressione multivariata e della teoria delle equazioni simultanee, ampiamente illustrati in Dhrymes (1970). Per esempio, volendo sintetizzare la variabilità multipla mediante la traccia (e non mediante il determinante, come in F^2), si potrebbe considerare l'indice

$$G^2 = 1 - tr\,(\tilde{\Sigma})/tr\,(\Gamma_0)\,,$$

il quale, pur se di agevole calcolo, non è influenzato dalla mutua correlazione delle componenti di $\mathbf{Z}_t$.

Nell'ambito dei suggerimenti provenienti dagli studi sulle serie storiche segnaliamo, invece, i contributi di Pierce (1979) e Geweke (1982), di Vitale (1982) e di Battaglia (1983).

Il lavoro di Pierce (1979) si pone in termini espliciti di molteplici misure di R^2 mentre quello di Geweke (1982) sviluppa l'analisi in termini indiretti mediante metriche del tipo $F = -\log(1 - R^2)$. Entrambi esplorano il grado di dipendenza fra serie multivariate pervenendo a diverse generalizzazioni dell'indice di determinazione, allo scopo di tenere comunque conto dei legami univariati tra le serie componenti e delle differenti tipologie causali: legame contemporaneo, unidirezionale, feedback, ecc. La nostra analisi si connette ad entrambi i lavori pur essendo più esplicitamente rivolta all'aspetto previsivo, da un lato per gli ovvi legami con l'indice di determinazione, dall'altro perché è connessa — come in Geweke — al rapporto di due determinanti, sintesi di variabilità multipla. Occorre tuttavia segnalare che, anche se questi due lavori sono notevoli per le connessioni che evidenziano fra analisi causale, relazioni tra spettri incrociati e capacità previsiva, molti dei risultati ivi dedotti sono già presenti in lavori di Parzen (si veda Parzen, 1982, per una elencazione) e in Granger (1963; 1969).

La proposta di Vitale (1982) media tre aspetti di un modello ARIMA che, secondo l'Autore, determinano la « capacità previsiva »: l'inerzia, misurata da R^2; l'evoluzione polinomiale, misurata dal grado d della differenza presente nel modello; la invertibilità, misurata dalla distanza delle radici *MA* dalla regione di non-invertibilità. Tale indice non è fondato su una relazione statica e dipende da una aggregazione soggettiva di aspetti non indipendenti; tuttavia, la proposta merita di essere sottolineata perché è forse l'unica misura che consente di pervenire ad un indice descrittivo per modelli *ARIMA*. Andrebbe quindi particolarizzata per modelli stagionali e generalizzata nel caso multivariato.

Il lavoro di Battaglia (1983), che estende un precedente risultato univariato, tratta delle matrici di autocovarianze inverse in tema di interpolazione ottimale di processi lineari multivariati. In tale contesto, l'Autore definisce un indice di determinismo lineare multiplo mediante

$$A_M = 1 - |\Gamma_i^{-1}(0)| / |\Gamma_0|$$

essendo $\Gamma_i(0)$ la matrice di autocovarianza inversa al *lag* $k = 0$. L'indice A_M « esprime la proporzione di varianza generalizzata di $\mathbf{Z}_t$ spiegabile mediante una combinazione lineare delle variabili $\mathbf{Z}_{t-k}$, $k \neq 0$ » fornendo, quindi, « una misura del grado in cui il processo tende a soddisfare un vincolo lineare deterministico » (ivi, p. 7).

Pertanto, mentre la nostra proposta F^2 tiene conto della storia « passata » di $\mathbf{Z}_t$ (essendo $\mathbf{Z}_{t-k}$, $k > 0$ l'unico insieme di informazioni utili per le previsioni), l'indice A_M di Battaglia tiene conto della storia « passata e futura » di $\mathbf{Z}_t$ (essendo $\mathbf{Z}_{t+k}$, $k \neq 0$ un insieme utile e necessario nei problemi di interpolazione dai quali A_M è derivato). In generale, tra i due indici sussistono le relazioni:

$$F^2 = 1 - (1 - A_M) \, |\tilde{\Sigma} \Gamma_i(0)| \longleftrightarrow A_M = 1 - (1 - F^2) \, |\tilde{\Sigma} \Gamma_i(0)|^{-1}$$

È agevole dimostrare che $F^2 = A_M$ se e solo se $\mathbf{Z}_t \sim WN$ multivariato, essendo questa l'unica situazione in cui $\tilde{\Sigma} \Gamma_i(0) = \mathbf{I}$.

Questi brevi note mostrano la necessità di un più unitario quadro di riferimento dei rapporti fra analisi canonica, teoria statistica delle previsioni ed analisi dei processi lineari ma confermano, a maggior ragione, la urgenza di convertire nella prassi della costruzione dei modelli tali acquisizioni teoriche. In tale direzione intendiamo sviluppare la ricerca successiva, sulla scìa di alcune generalizzazioni ed applicazioni già presenti nei lavori pionieristici di Akaike (1976), Robinson (1973), Tsay e Tiao (1983).

Una sintesi dei risultati qui esposti è stata presentata alla Giornata di studio organizzata dalla Commissione SIS per lo « Studio delle Serie Storiche », svoltasi a Roma il 14 dicembre 1982. Il lavoro è stato completato mentre l'Autore era Visiting Fellow presso il Department of Statistics, Madison, University of Wisconsin, grazie ad un parziale contributo CNR-NATO Senior Fellowship Scheme, nel periodo giugno-settembre 1983.

Riferimenti bibliografici

Akaike H., Canonical correlation analysis of time series and the use of an information criterion, in « *System Identification: Advances and Case Studies* » (eds. R.K. Mehra & D.G. Lainiotis), Academic Press, N.Y., 1976, 27-96.

Battaglia F., Le covarianze inverse di una serie temporale multivariata, *Pubblic. Istituto di Calcolo delle Probabilità*, Università di Roma, 1983, 1/83.

Box G.E.P. & Tiao G.C., A Canonical Analysis of Multiple Time Series, *Biometrika*, **64**, 1977, 355-365.

Dhrymes P.J., *Econometrics. Statistical Foundations and Applications*, Harper & Row, New York, 1970.

Geweke J., Measurement of Linear Dependence and Feedback Between Multiple Time Series, *JASA*, **77**, 1982, 378, 304-324 (with Discussion).

Granger C.W.J., Economic Processes Involving Feedback, *Information and Control*, **6**, 1963, 28-48.

Granger C.W.J., Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, **37**, 1969, 424-438.

Mazzali A., Gli spazi di previsione dei modelli ARMA, *Rivista di Statistica Applicata*, **15**, 4, 1982, 295-308.

Nelson C.R., The interpretation of R^2 in Autoregressive-Moving Average Time Series Models, *The American Statistician*, **30**, 4, 1976, 175-180.

Parzen E., Comment to: « Measurement of Linear etc.... » by Geweke (1982), *JASA*, **77**, 1982, 378, 320-322.

Piccolo D., Processi stocastici lineari per l'analisi delle serie storiche, *Produzione Animale*, **13**, 1974a, 107-121.

Piccolo D., *Analisi delle serie temporai: i processi AR(2)*, Giannini ed., Napoli, Centro di Specializzazione e Ricerche, Portici, 1974b.

Piccolo D., Per una interpretazione dei modelli ARIMA, in: « *Analisi moderna delle serie storiche* » (ed. D. Piccolo), Atti Convegno Nazionale 1981, F. Angeli ed., Milano, 1983.

Piccolo D., Capacità previsiva dei modelli stocastici lineari, in: « *Atti XXXI Riunione Scientifica SIS* », Torino, vol. II, 1982, 489-503.

Pierce D.A., R^2 Measures for Time Series, *JASA*, **74**, 1979, 901-910.

Priestley M.B., *Spectral Analysis and Time Series*, voll. I, II, Academic Press, New York, 1981.

Quenouille M.H., *The Analysis of Multiple Time Series*, Griffin, London, 1957.

Robinson P.M., Generalized canonical analysis for time series, *Journal of Multivariate analysis*, **3**, 1973, 141-160.

Stone J.R.N., Prediction from autoregressive schemes and linear stochastic difference sistems, *Econometrica*, Suppl. II, 1949, 29-37.

Stralkowski C.M., Wu S.M. & DeVor R.E., Charts for the Interpretation and Estimation of the Second Order Autoregressive Model, *Technometrics*, **12**, 3, 669-685.

Tiao G.C., Box G.E.P., Grupe M.R., Hudak G.B., Bell W.R. & Chang G.I., *The Wisconsin Multiple Time Series (WMTS-1) Program: A Preliminary Guide*, Department of Statistics, University of Wisconsin, Madison, 1979.

Tsay R.S. & Tiao G.C., *Use of Canonical Analysis in Time Series Model Identification*, Techn. Rep. 283, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA, 1983.

Vitale C., Capacità previsiva e non invertibilità nei modelli ARIMA, in: « *Atti XXXI Riunione Scientifica SIS* », Torino, vol. II, 1982.

Summary

The forecastability of a linear model becomes very often a decision rule for many statistical objectives: identification of model orders, parameters checking, causality analysis, and so on. In this context, we propose a multivariate index for vector ARMA processes and we discuss its meaning and properties. Moreover, we show that this measure is closely related to the canonical transformation analysis of which it may be an useful synthesis. A brief comparison with some other proposed indexes concludes the work.