Analisi di intervento (*Intervention Analysis)*

Matteo Pelagatti

SOMMARIO

1 Introduzione 1
2 Caso univariato 2
2.1 Ipotesi sul modello stocastico in assenza di intervento 2
2.2 Modello d'intervento 2
2.3 Residui 6
2.4 Stima congiunta dei parametri 7
3 Caso multivariato 8
3.1 Ipotesi sul modello stocastico in assenza di intervento 8
3.2 Modello d'intervento 8
3.3 Stima congiunta dei parametri 9
3.4 Stima a due stadi dei parametri 10
Bibliografia 11

1 INTRODUZIONE

Spesso capita di dovere valutare l'effetto di un *intervento* esterno sull'andamento di una serie storica, dove per intervento si intende l'azione di un qualsiasi fenomeno che ad un dato periodo T - o da un dato periodo T - influenza la serie stessa. Esempi di simili interventi possono essere l'effetto di una campagna pubblicitaria sulla serie storica delle vendite di un certo prodotto, i risultati di un provvedimento contro l'inquinamento da gas di scarico sulla serie storica della quantità di Ozono nell'aria (Box e Tiao, 1975), le conseguenze dell'obbligatorietà delle cinture di sicurezza sulla serie storica delle morti per incidente stradale (Bhattacharyya e Layton, 1979; Abraham, 1987), ecc.

Da un punto di vista metodologico, ciò che differenzia l'Analisi d'intervento dalle procedure più generali di identificazione e stima degli outliers è la conoscenza a priori della posizione e della natura degli outliers. Questo tipo di analisi è anche utile quando si siano identificati degli outliers per mezzo di test di discordanza e se ne voglia indagare la struttura e le conseguenze sul modello stocastico che ha generato la serie. Inoltre, come risulterà evidente dal seguito di questo capitolo, l'Intervention Analysis può essere vista come un caso particolare nell'ambito dei modelli di *funzione di trasferimento*: infatti se z_t (per $t = 1, 2, 3, ..., n$) è una realizzazione finita di un processo stocastico, e p_t è una serie tale che, supposto essere T ($< T < n$) l'istante dell'intervento, per $t \neq T \leftrightarrow p_t = 0$ e per $t = T \leftrightarrow p_t = 1$, l'analisi d'intervento consiste nella ricerca e nella stima della funzione di trasferimento che pone in relazione la serie p_t con la serie z_t.

2 CASO UNIVARIATO

2.1 Ipotesi sul modello stocastico in assenza di intervento

Sia z_t $(t = 1, 2, 3, ..., n)$ una serie storica generata dal processo stocastico stazionario Z_t con valore atteso nullo e rappresentabile attraverso la modellistica ARIMA:

$$\varphi(B)Z_t = \theta(B)a_t$$

dove:

1. B è l'operatore *backshift* tale che $By_t = y_{t-1}$;
2. $...a_{t-1}, a_t, a_{t+1}, ...$ è una serie di variabili casuali indipendentemente ed identicamente distribuite con media nulla e varianza σ_a^2;
3. $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q$, $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - ... - \varphi_p B^p$ sono i polinomi "media mobile" e "autoregressivo" in B di grado q e p rispettivamente;
4. le radici di $\theta(B)$ sono esterne al cerchio unitario e quelle di $\varphi(B)$ sono non interne al cerchio unitario.

2.2 Modello d'intervento

Si supponga che al tempo T ($< T < n$) avvenga un intervento, e che quindi z_t sia rappresentabile attraverso un modello ARMA solamente per $t \in [1, (T-1)]$, ma non per il resto della serie. Box e Tiao (1975), nel lavoro genitore della Intervention Analysis, proposero il seguente modello d'intervento:

$$Y_t = \Im_t + Z_t = \frac{\omega(B)}{\partial(B)} I_t + \frac{\theta(B)}{\varphi(B)} a_t$$

dove

1. $\Im_t$ è la funzione di trasferimento, che trasferisce gli effetti dell'*input* I_t sull'*output* Y_t. $\Im_t$ è data dalla soluzione dell'equazione lineare alle differenze:

   $$\delta(B)\Im_t = \omega(B)I_t$$

2. I_t è una variabile indicatore, detta *variabile d'intervento*, che assume solo valori nulli o unitari, e che può configurare l'intervento in due modi utili:
   attraverso una variabile "scalino" che assume valore unitario dall'istante dell'intervento in poi e valore nullo altrove:

   $$I_t = S_t^{(T)} = \begin{cases} 0, & \mathrm{t} < \mathrm{T} \\ 1, & \mathrm{t} \geq \mathrm{T} \end{cases}$$

   o attraverso una variabile "impulso" che assume valore unitario nell'istante dell'intervento e valore nullo altrove:

   $$I_t = P_t^{(T)} = \begin{cases} 0, & \mathrm{t} \neq \mathrm{T} \\ 1, & \mathrm{t} = \mathrm{T} \end{cases}$$

   tra le due variabili sussiste la seguente relazione

   $$(1 - B)S_t^{(T)} = P_t^{(T)}$$

   per la quale un modello d'intervento può essere rappresentato indifferentemente per mezzo dell'una o dell'altra variabile;
3. δ(B) e ω(B) sono polinomi in B di grado *r* e *s* rispettivamente:

$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \ldots - \delta_r B^r$$

$$\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \ldots - \omega_s B^s .$$

Nel caso in cui l'intervento provochi un salto di livello di ampiezza ω, ritardato di k periodi, la funzione di trasferimento adatta è (fig.2.1 dove $k = 1$):

$$\Im_t = \omega B^k S_t^{(T)} .$$

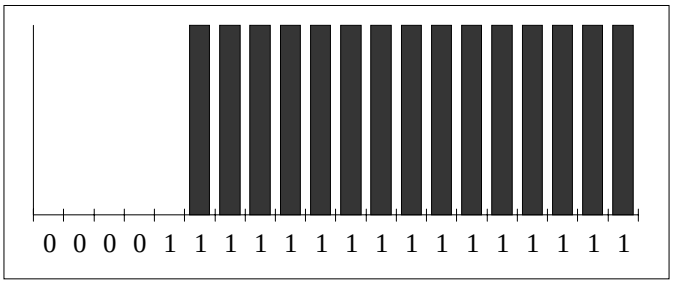

fig.2.1

Il medesimo *output* si sarebbe potuto ottenere con la funzione:

$$\Im_t = \frac{\omega B^s}{1 - B} P_t^{(T)}$$

dove l'*input* è la variabile impulso.

Alcune volte, invece, il cambio di livello può avvenire gradualmente a partire dal periodo $T + 1$; in questo caso un'appropriata funzione di trasferimento è:

$$\Im_t = \frac{\omega B}{1 - \delta B} S_t^{(T)} \quad 0 < \delta < 1,$$

che tende a $\dfrac{\omega}{1 - \delta}$ per $t \to \infty$ (fig.2.2).

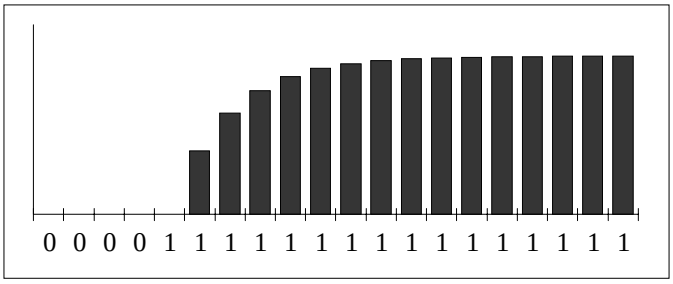

fig.2.2

Quando $\delta = 1$ la stessa funzione di trasferimento produce una "rampa" che tende ad infinito per $t \to \infty$ (fig.2.3).

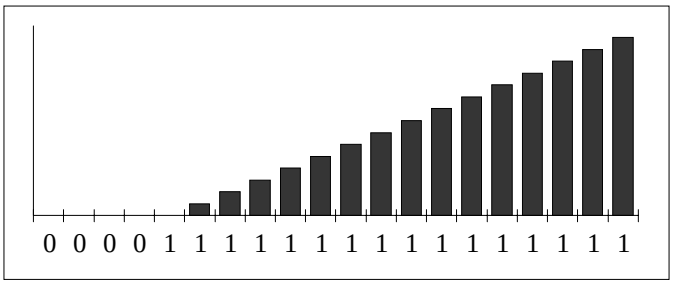

fig.2.3

Se si desidera rappresentare una situazione in cui l'intervento (per esempio di una campagna pubblicitaria) provoca un balzo verso l'alto della serie, che rientra gradualmente sino a tornare al livello di partenza, la funzione appropriata è (fig.2.4):

$$\Im_t = \frac{\omega B}{1-\delta B} P_t^{(T)}.$$

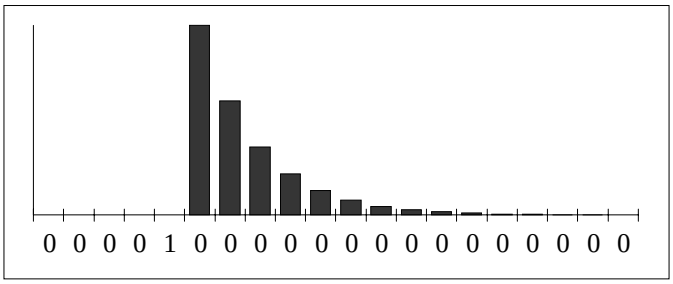

fig.2.4

Se invece la serie, dopo il salto iniziale, ritorna gradualmente verso un livello più alto di quello pre-intervento, una funzione adatta è:

$$\Im_t = \frac{\omega_1 B}{1-\delta B} P_t^{(T)} + \omega_2 B S_t^{(T)}$$

o equivalentemente

$$\Im_t = \left(\frac{\omega_1 B}{1-\delta B} + \frac{\omega_2 B}{1-B} \right) P_t^{(T)}$$

dove il livello iniziale è $(\omega_1 + \omega_2)$, mentre quello tendenziale, al crescere di t, è ω_2 (fig.2.5).

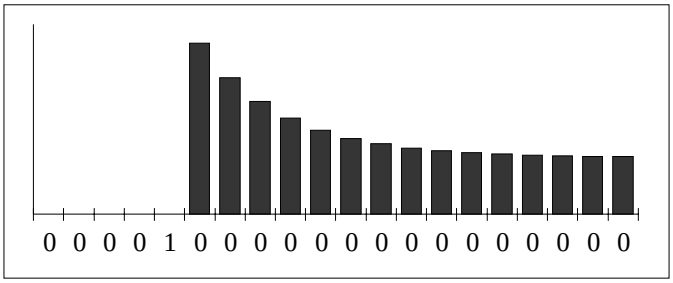

fig.2.5

Box e Tiao (1975) suggeriscono di utilizzare diverse variabili impulso per rappresentare situazioni in cui l'intervento si estende per diversi periodi (ad esempio una campagna pubblicitaria prolungata); tuttavia in alcuni casi, per parsimonia di parametri, si potrebbe utilizzare una variabile $R_t^{(T_1 \ldots T_2)}$ tale che:

$$R_t^{(T_1 \ldots T_2)} = \begin{cases} 0, & \text{per } t < T_1 \wedge t > T_2 \\ 1, & \text{per } T_1 \le t \le T_2 \end{cases}$$

Una situazione in cui il livello della serie si incrementa linearmente sino a raggiungere una nuova stazionarietà può essere rappresentata per mezzo della variabile descritta, con la seguente funzione (fig.2.6):

$$Y_t = \frac{\omega}{1-B} R_t^{(T_1 \ldots T_2)}$$

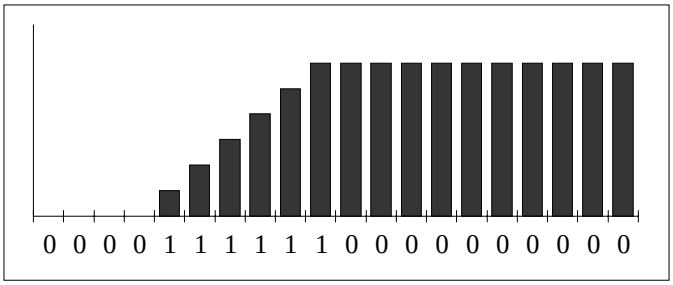

fig.2.6

dove ω è l'incremento periodico tra $T_1 - 1$ e T_2.

Il caso in cui la serie subisce un incremento graduale a differenze decrescenti in $[T_1, T_2]$, per poi ritornare asintoticamente verso il livello di partenza da $T_2 + 1$ è ben rappresentato da (fig.2.7):

$$\Im_t = \frac{\omega}{(1-\delta B)} R_t^{(T_1 \ldots T_2)}$$

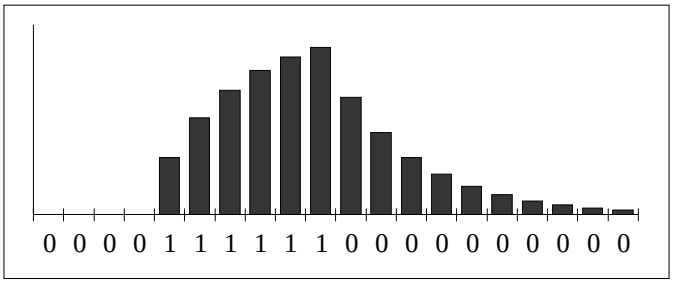

fig.2.7

dove tra T_1 e T_2 $\Im_t \to \frac{\omega}{1-\delta}$ al crescere di *t*, mentre da $T_2 + 1$ $\Im_t \to 0$ per *t* sufficientemente grande. Sommando questa funzione di trasferimento con $\omega_2 S_t^{(T_1)}$ o con $\omega_2 S_t^{(T_2+1)}$ si ottengono funzioni di trasferimento simili, ma che dopo T_2 tendono a ω_2 al crescere di *t* (figg.2.8 e 2.9).

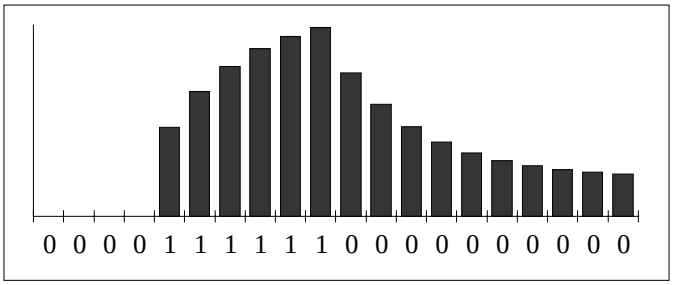

fig.2.8

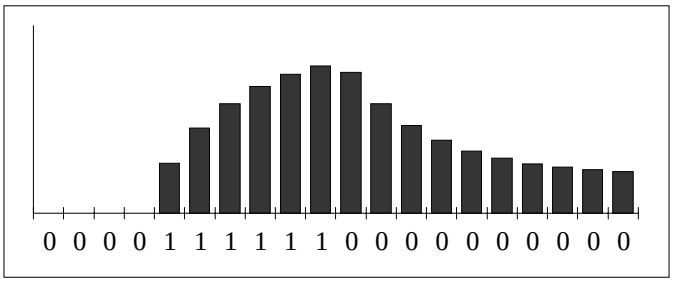

fig.2.9

È importante notare che queste funzioni di trasferimento implicano l'uguaglianza del coefficiente di crescita e del coefficiente di caduta. Perciò un loro utilizzo migliora l'efficienza

delle stime solo quando questa ipotesi è plausibile; altrimenti l'intrinseco vincolo sui parametri che deriva dall'utilizzo di tali funzioni provoca gravi distorsioni degli stimatori.

Un caso particolarmente interessante si presenta quando

$$\Im_t = \frac{\theta(B)}{\varphi(B)} \omega P_T \ .$$

Questa funzione di trasferimento rappresenta un intervento che nella letteratura sugli outliers è usualmente chiamato *Innovational Outlier* (IO). La particolarità dell'IO è che esso agisce additivamente sulla innovazione[1] a_T :

$$Y_t = \frac{\theta(B)}{\varphi(B)} (a_t + \omega P_T) \ .$$

L'estensione del modello al caso di *k* interventi è immediata:

$$Y_t = \sum_{j=1}^{k} \Im_{tj} + Z_t = \sum_{j=1}^{k} \frac{\omega_j(B)}{\delta_j(B)} I_{tj} + \frac{\theta(B)}{\varphi(B)} a_t \ .$$

2.3 Residui

É interessante esaminare l'effetto di un intervento sui residui della serie rispetto al solo modello ARMA appropriato (senza la funzione di trasferimento e le variabili d'intervento).

Mantenendo la stessa simbologia dei paragrafi precedenti, la serie dei residui è data da

$$e_t = \frac{\varphi(B)}{\theta(B)} Y_t \ .$$

Ma dato che

$$Y_t = \frac{\omega(B)}{\partial(B)} I_t + \frac{\theta(B)}{\varphi(B)} a_t \ ,$$

sostituendo si ottiene

$$e_t = \frac{\theta(B)}{\varphi(B)} \frac{\omega(B)}{\partial(B)} I_t + a_t \ ,$$

il cui valore atteso è

$$E[e_t] = \frac{\theta(B)}{\varphi(B)} \frac{\omega(B)}{\partial(B)} I_t \ .$$

Ne consegue che, stimato il modello ARMA, per modelli non troppo complessi ci si può fare un'idea della struttura dell'intervento osservando il grafico dei residui. In questo caso le stime del modello ARMA incontaminato possono essere ottenute utilizzando la porzione di dati che precede l'intervento, o per mezzo degli *stimatori robusti* che saranno presentati nel prossimo capitolo. Un siffatto esame può essere di qualche utilità quando attraverso metodi di identificazione degli outliers si venga a conoscenza della posizione di un valore anomalo e se ne voglia indagare la struttura.

Nel caso particolare dell'IO i residui sono dati da

---

[1]La serie di variabili casuali $[a_t]$ è detta serie delle innovazioni.

$$e_t = \omega P_T + a_t \ ,$$

quindi questo tipo di intervento è particolarmente facile da riconoscere attraverso l'osservazione della serie dei residui.

2.4 Stima congiunta dei parametri

Sia

$$y_t = \sum_{j=1}^{k} \Im_{tj} + Z_t = \sum_{j=1}^{k} \frac{\omega_j(B)}{\delta_j(B)} I_{tj} + \frac{\theta(B)}{\varphi(B)} a_t$$

il modello in cui si vogliono stimare i parametri del processo ARMA e della funzione di trasferimento, ferme restando le ipotesi precedentemente fatte sul processo Z_t e sulle funzioni $\Im_{tj}$.

Sia $\mathbf{w}$ un vettore il cui t-esimo elemento è[2]

$$w_t = y_t - \sum_{j=1}^{k} \Im_{tj}$$

che conseguentemente segue il modello

$$w_t = \frac{\theta(B)}{\varphi(B)} a_t \ .$$

Indicando con $\boldsymbol{ß}$ il vettore di tutti i parametri da stimare (del modello ARMA e delle funzioni di trasferimento), la funzione di verosimiglianza può essere così scritta:

$$L(\boldsymbol{ß}, \sigma_a^2 | \mathbf{y}) = (2\pi\sigma_a^2)^{-\frac{n}{2}} |\mathbf{M}|^{\frac{1}{2}} \exp(-S(\boldsymbol{ß})/2\sigma_a^2) \ ,$$

dove $\mathbf{M}^{-1}\sigma_a^2$ è la matrice di covarianza del vettore $\mathbf{w}$ e

$$S(\boldsymbol{ß}) = \mathbf{w}'\mathbf{M}\mathbf{w} = \sum_{t=-\infty}^{n} \left[\mathrm{E}(a_t | \mathbf{y}, \boldsymbol{ß})\right]^2 .$$

Se nessuna delle radici di $\theta(B)$ e di $\varphi(B)$ è prossima al cerchio unitario ed n è sufficientemente grande, i valori degli elementi di $\boldsymbol{ß}$ che minimizzano $S(\boldsymbol{ß})$ (minimi quadrati) sono una buona approssimazione dei valori di massima verosimiglianza. Dato che $\mathbf{w}$ è stazionario, $\mathrm{E}(a_t | \mathbf{y}, \boldsymbol{ß})$ sarà trascurabile per valori $t \le -Q$ dove Q è un intero positivo opportunamente scelto; perciò possiamo sostituire la somma infinita $S(\boldsymbol{ß})$ con la somma finita

---

[2]Per semplificare la notazione e non appesantire le formule si è supposto che il processo Z_t fosse stazionario con valor medio nullo. Ove queste condizioni non sussistessero, sarebbe più comodo sviluppare la funzione di trasferimento per la serie non differenziata, piuttosto che differenziare prima dell'analisi. In questo caso la formula corretta per $\mathbf{w}$ è:

$$\mathbf{w} = (1-B)^d (1-B^s)^D \{(y_t - \mu) - \sum_{j=1}^{k} \Im_{tj}\}$$

dove d è il grado della differenziazione, s è il numero di dati della serie per anno e D è il grado della differenziazione stagionale. Il valore di μ può essere stimato con uno stimatore di posizione robusto (cfr. cap. 3), altrimenti è un ulteriore parametro da stimare col metodo della massima verosimiglianza.

$$S_*(\text{ß}) = \sum_{t=-Q}^{n} \left[\mathrm{E}(a_t | \mathbf{y}, \text{ß}) \right]^2 .$$

In Box e Jenkins (1976) viene mostrato come i valori iniziali della sommatoria ed il loro numero Q possano essere ottenuti attraverso un processo di "previsione all'indietro".

3 CASO MULTIVARIATO

3.1 Ipotesi sul modello stocastico in assenza di intervento

Sia[3] $\mathbf{z}_t = (z_{1t}, ..., z_{mt})'$ (t = 1, 2, ..., n) un vettore di serie storiche generate dal vettore di processi stocastici stazionari $\underline{\mathbf{Z}}_t = (Z_{1t}, ..., Z_{mt})'$ con valori attesi nulli e rappresentabili attraverso la modellistica VARMA[4] (*vector autoregressive moving average models*):

$$\boldsymbol{\varphi}(B)\underline{\mathbf{Z}}_t = \boldsymbol{\theta}(B)\mathbf{a}_t$$

dove:

1. B è l'operatore *backshift* tale che $By_t = y_{t-1}$;
2. $\mathbf{a}_t = (a_{1t}, ..., a_{mt})'$, $t = ...,1,2,3,...$, è una sequenza di vettori di variabili casuali temporalmente indipendentemente distribuite, con vettore delle medie nullo e matrice di covarianza $\boldsymbol{\Sigma}$ non singolare;
3. $\boldsymbol{\theta}(B) = \mathbf{I}_m - \boldsymbol{\theta}_1 B - \boldsymbol{\theta}_2 B^2 - ... - \boldsymbol{\theta}_q B^q$, $\boldsymbol{\varphi}(B) = \mathbf{I}_m - \boldsymbol{\varphi}_1 B - \boldsymbol{\varphi}_2 B^2 - ... - \boldsymbol{\varphi}_p B^p$ sono matrici $m \times m$ polinomiali in B di grado q e p rispettivamente, con $\boldsymbol{\theta}_i$ e $\boldsymbol{\varphi}_j$ matrici $m \times m$ di parametri;
4. le radici di $|\boldsymbol{\theta}(B)|$ sono esterne al cerchio unitario e quelle di $|\boldsymbol{\varphi}(B)|$ sono non interne al cerchio unitario.

3.2 Modello d'intervento

Si supponga che nel periodo T avvenga un intervento. Tenuto conto delle nuove ipotesi, il modello d'intervento può essere così riscritto:

$$\underline{\mathbf{Y}}_t = \underline{\boldsymbol{\mathfrak{I}}}_t + \underline{\mathbf{Z}}_t = \boldsymbol{\delta}_d^{-1}(B)\boldsymbol{\omega}_d(B)\underline{\mathbf{I}}_T + \boldsymbol{\varphi}^{-1}(B)\boldsymbol{\theta}(B)\mathbf{a}_t$$

dove

1. $\underline{\boldsymbol{\mathfrak{I}}}_t$ è un vettore di funzioni di trasferimento;
2. $\underline{\mathbf{I}}_T = (I_{1T}, ..., I_{mT})'$ è un vettore di variabili d'intervento dello stesso genere di quelle già viste (variabile scalino, variabile impulso);

---

[3]Un chiarimento sulla notazione che verrà usata: dato che usualmente le variabili casuali vengono rappresentate con lettere maiuscole e le loro realizzazioni con lettere minuscole, ma è anche vero che le matrici vengono simboleggiate da lettere maiuscole in grassetto e i vettori con lettere minuscole in grassetto, in seguito verrà tenuta la distinzione maiuscolo-minuscolo per le matrici e i vettori mentre i vettori di variabili casuali verranno scritti con stile maiuscolo-sottolineato. Unica eccezione la variabile casuale white noise tradizionalmente rappresentata con la lettera minuscola *a*.

[4]I metodi di identificazione e stima dei modelli ARMAV sono stati trattati da Haug e Box (1977) Tiao e Box (1981), Tsay (1989).

3. $\boldsymbol{\delta}_d(B)$ e $\boldsymbol{\omega}_d(B)$ sono matrici diagonali $m \times m$ i cui rispettivi elementi (sulla diagonale) sono:

$$\delta_{ii}(B) = 1 - \delta_{1ii} B - \delta_{2ii} B^2 - \ldots - \delta_{r_i ii} B^{r_i},$$

$$\omega_{ii}(B) = \omega_{0ii} - \omega_{1ii} B - \omega_{2ii} B^2 - \ldots - \omega_{s_i ii} B^{s_i}.$$

Per estendere il modello al caso di k interventi si può procedere come sopra, sommando k funzioni di trasferimento, oppure mantenendo la pura forma matriciale:

$$\underline{\mathbf{Y}}_t = \sum_{i=1}^{k} \underline{\mathfrak{I}}_{it} + \underline{\mathbf{Z}}_t = \boldsymbol{\Phi}(B)\mathbf{i}_t + \boldsymbol{\varphi}^{-1}(B)\boldsymbol{\theta}(B)\mathbf{a}_t,$$

dove:

1. $\boldsymbol{\Phi}(B)$ è una matrice $m \times k$ il cui generico elemento è $\phi_{ij} = \dfrac{\omega_{ij}(B)}{\delta_{ij}(B)}$ $(i = 1,\ldots,m;\ j = 1,\ldots,k)$, con $\delta_{ij}(B) = 1 - \delta_{1ij} B - \delta_{2ij} B^2 - \ldots - \delta_{r_{ij}\,ij} B^{r_{ij}}$, e $\omega_{ij}(B) = \omega_{0ij} - \omega_{1ij} B - \omega_{2ij} B^2 - \ldots - \omega_{s_{ij}\,ij} B^{s_{ij}}$.
2. $\mathbf{i}_t = (I_{T_1}, I_{T_2}, \ldots, I_{T_k})'$ è un vettore di variabili d'intervento. Si noti il diverso significato di questo vettore rispetto a $\underline{\mathbf{I}}_T = (I_{1T}, \ldots, I_{mT})'$ del caso precedente: ora ciascun elemento del vettore rappresenta una variabile che interviene in un proprio periodo T_i, in precedenza le variabili, elementi del vettore $\underline{\mathbf{I}}_T$, intervenivano tutte nello stesso periodo T, e per ogni singolo *output*, elemento del vettore $\underline{\mathbf{Y}}_t$, si poteva decidere la forma della variabile d'intervento (scalino o impulso). Questa rappresentazione, peraltro, non toglie nulla alla scelta della variabile d'intervento più opportuna, infatti, come si è visto, ogni variabile scalino è trasformabile in una variabile impulso e viceversa.

3.3 Stima congiunta dei parametri

Si prenda in considerazione il modello

$$\mathbf{y}_t = \sum_{i=1}^{k} \underline{\mathfrak{I}}_{it} + \underline{\mathbf{Z}}_t = \boldsymbol{\Phi}(B)\mathbf{i}_t + \boldsymbol{\varphi}^{-1}(B)\boldsymbol{\theta}(B)\mathbf{a}_t,$$

che rappresenta il caso più generale, dove la variabile a sinistra dell'uguale è ora un vettore di realizzazioni di m processi stocastici. Risolvendo rispetto ad $\mathbf{a}_t$ si ottiene

$$\mathbf{a}_t = \boldsymbol{\theta}^{-1}(B)\boldsymbol{\varphi}(B)\left[\mathbf{y}_t - \boldsymbol{\Phi}(B)\mathbf{i}_t\right].$$

Si indichi con **ß** il vettore di tutti i parametri da stimare (del modello VARMA e delle funzioni di trasferimento) e con $\boldsymbol{\Sigma}$ la matrice di covarianza di $\mathbf{a}_t$. La funzione di densità di probabilità di $\mathbf{a}_t$, supposti gaussiani i *white noise*, è

$$f(\mathbf{a}_t) = (2\pi)^{-\frac{m}{2}} \left|\boldsymbol{\Sigma}\right|^{-\frac{1}{2}} \exp(-\tfrac{1}{2}\mathbf{a}'_t \boldsymbol{\Sigma}^{-1}\mathbf{a}_t).$$

Essendo i vettori $\mathbf{a}_t$ temporalmente indipendenti e identicamente distribuiti, risulta

$$p(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n) = \prod_{t=1}^{n} f(\mathbf{a}_t) = (2\pi)^{-\frac{mn}{2}} \left|\boldsymbol{\Sigma}\right|^{-\frac{n}{2}} \exp(-\tfrac{1}{2}\sum_{t=1}^{n} \mathbf{a}'_t \boldsymbol{\Sigma}^{-1}\mathbf{a}_t).$$

La funzione di verosimiglianza[5] condizionata ai valori iniziali $\mathbf{y}_*$ e $\mathbf{a}_*$ sarà

$$L(\boldsymbol{\beta}|\mathbf{y}, \mathbf{y}_*, \mathbf{a}_*) = (2\pi)^{-\frac{mn}{2}} |\boldsymbol{\Sigma}|^{-\frac{n}{2}} \exp\left[-\tfrac{1}{2}\,\mathrm{tr}\,\boldsymbol{\Sigma}^{-1}\mathbf{S}(\boldsymbol{\beta}|\mathbf{y}, \mathbf{y}_*, \mathbf{a}_*)\right]$$

dove $\mathbf{S}(\boldsymbol{\beta}|\mathbf{y},\mathbf{y}_*,\mathbf{a}_*)= \sum_{t=1}^{n} \mathbf{a}_t\mathbf{a}'_t$.

Se nessuna delle radici di $|\boldsymbol{\theta}(B)|$ e di $|\boldsymbol{\varphi}(B)|$ è prossima al cerchio unitario ed n è sufficientemente grande, i valori degli elementi di **ß** che minimizzano $\mathbf{S}(\boldsymbol{\beta})$ (minimi quadrati) sono una buona approssimazione dei valori di massima verosimiglianza.

3.4 Stima a due stadi dei parametri

Abraham (1980) ha proposto una procedura di stima a due stadi, nel primo dei quali vengono stimati i parametri VARMA, e nel secondo dei quali si stimano quelli delle funzioni di trasferimento. Il vantaggio di questo metodo è che, diminuendo significativamente il numero di parametri da stimare simultaneamente, semplifica la massimizzazione della funzione di verosimiglianza. Vi sono però anche svantaggi e restrizioni: la procedura richiede, infatti, che esista un numero di osservazioni pre-intervento sufficiente per potere identificare e stimare il modello VARMA, ma anche in questo caso più fortunato vi è una perdita di informazione nella stima dei parametri, e conseguentemente le proprietà dello stimatore a due stadi sono meno buone di quelle dello stimatore congiunto[6].

Nel primo stadio si identifica e stima il modello VARMA utilizzando le prime $T-1$ osservazioni della serie. Fatto questo si può riscrivere il modello generale nel seguente modo:

$$\mathbf{y}_t = \boldsymbol{\Phi}(B)\mathbf{i}_t + \overline{\boldsymbol{\varphi}}^{-1}(B)\overline{\boldsymbol{\theta}}(B)\mathbf{a}_t$$

dove il segno orizzontale sopra gli operatori significa che essi sono stati già stimati.

Ponendo

$$\mathbf{y}'_t = \overline{\boldsymbol{\theta}}^{-1}(B)\overline{\boldsymbol{\varphi}}(B)\mathbf{y}_t \quad \text{e} \quad \boldsymbol{\Phi}'(B) = \overline{\boldsymbol{\theta}}^{-1}(B)\overline{\boldsymbol{\varphi}}(B)\boldsymbol{\Phi}(B)$$

la formula del modello generale diventa:

$$\mathbf{y}'_t = \boldsymbol{\Phi}'(B)\mathbf{i}_t + \mathbf{a}_t$$

Il secondo stadio, cioè la stima dei parametri delle funzioni di trasferimento è del tutto simile a quella di una regressione multivariata non lineare.

---

[5]Si noti che $\sum_{t=1}^{n} \mathbf{a}'_t\boldsymbol{\Sigma}^{-1}\mathbf{a}_t = tr\left[\boldsymbol{\Sigma}^{-1}\mathbf{A}\right]$ dove $\mathbf{A} = \sum_{t=1}^{n} \mathbf{a}_t\mathbf{a}'_t$.

[6]La situazione qui descritta è simile a quella dei minimi quadrati a due stadi in econometria, per le proprietà di questi stimatori si rimanda ad un manuale di econometria (per es. Theil, 1971, pp. 497-500).

BIBLIOGRAFIA

Box, G. E. P., Tiao G. C. (1975) - "Intervention Analysis with applications to Economics and Enviromental Problems", *JASA* 70, 349 pp. 70-9.

Bahattacharyya M. N., Layton A. P. (1979) - "Road Toll, an Australian Case Study in Intevention Analysis", *JASA* 74, 596-603.

Abraham B. (1987) - "Applacation of Intervention Analysis to a Road Fatality Series in Ontario", *Journal of Forecasting*, 6, 211-219.

Abraham B. (1980) - "Intervention Analysis and Multiple Time Series", *Biometrika*, 67, 73-78.

Box, G. E. P., Tiao G. C. (1981) - "Modelling multiple time series with applications", *JASA* 76, 802.816.

Fox, A. J. (1972) - "Outliers in Time Series", *Journal of the Royal Statistical Society*, Ser. B, 43, 350-363.

Chung C., Lon-Mu L. (1993) - "Joint Estimation of Model Parameters and Outlier Effects in Time Series", *JASA*, 88, 421, 284-297.

Murry J. P. Jr., Stam A. ,Lastovicka J. L. (1993) - "Evalueting an Anti-Drinking and Driving Advertising Campaign With a Sample Survey and Time Series Intervention Analysis", *JASA* Vol.88, No.421, 50-56.

Tiao G. C. (1985) - "Autoregressive Moving Average Models, Intervention Problems and Outlier Detection in Time Series", in Hannan E. J., Krishnaiah P. R., Rao M. M. (Eds). *Handbook of Statistics*, Vol. 5. Elsevier Science Publishers, pp. 85-118.