

STUDI STATISTICI n. 8

MODELLI LINEARI DINAMICI

Pietro Muliere

marzo 1984

MODELLI LINEARI DINAMICI *

Pietro Muliere : Istituto di Scienze Economiche e Statistiche,
Università degli Studi di Pavia

RIASSUNTO

In questo lavoro vengono studiati i problemi statistici riguardanti i modelli lineari dinamici. Sono considerati i seguenti problemi: filtraggio, interpolazione, previsione. Questi tre problemi vengono risolti applicando il filtro di Kalman-Bucy. La determinazione del filtro di Kalman-Bucy è presentata nell'impostazione bayesiana.

* Lavoro eseguito nell'ambito del Gnafa-Cnr.
Sono grato al dott. C. Giannasca per il fondamentale contributo fornitomi nella stesura dei programmi.

I N D I C E

1. INTRODUZIONE	Pag. 1
2. SCHEMA DEL LAVORO	" 3
3. I MODELLI LINEARI DINAMICI	" 4
4. IL FILTRO DI KALMAN-BUCY: IMPOSTAZIONE BAYESIANA	" 7
5. PROPRIETA' ED OSSERVAZIONI SUL FILTRO DI KALMAN-BUCY	" 13
6. ALCUNI ESEMPI DI MODELLI LINEARI DINAMICI	" 20
6.1 Il modello stabile	" 20
6.2 Il modello di crescita lineare	" 26
6.2.1 Estensioni del modello di crescita	" 30
6.2.2 Il modello multi-stato	" 38
6.3 Il modello stagionale	" 50
6.4 Il modello stagionale additivo	" 51
6.5 Il modello stagionale di crescita lineare	" 52
6.6 Il modello lineare generale statico	" 59
6.7 Il modello di regressione dinamico	" 60
6.8 Il modello ARMA	" 64
7. UN METODO ALTERNATIVO PER DETERMINARE IL FILTRO DI KALMAN-BUCY.	" 65
8. UNA GENERALIZZAZIONE DEL MODELLO LINEARE DINAMICO	" 69
8.1 Definizioni del modello ed assunzioni di base	" 69
8.2 Equazioni ricorsive per il filtro	" 71
8.3 Alcuni semplici esempi	" 79

II

9. IL PROBLEMA DELL'INTERPOLAZIONE

Pag. 82

9.1 Il modello lineare dinamico semplice

" 82

9.2 Il modello lineare dinamico generalizzato

" 84

10. IL PROBLEMA DELLA PREVISIONE

" 89

10.1 Il modello lineare dinamico semplice

" 89

10.2 Il modello lineare dinamico generalizzato

" 92

BIBLIOGRAFIA

" 94

1. INTRODUZIONE.

Il tema della previsione ha da sempre affascinato gli statistici ed occupa un posto importante nella teoria dei processi stocastici e nell'analisi delle serie storiche. I primi lavori su tale argomento sono stati di Kolmogorov e Wiener. Kolmogorov (1941) ha affrontato il problema per processi aleatori stazionari a tempo discreto, mentre Wiener (1949) ha analizzato il caso di processi a tempo continuo. Kolmogorov, nel lavoro citato, utilizza una rappresentazione proposta da Wold (1938), Wiener invece riduce il problema della previsione alla soluzione dell'equazione integrale di Wiener-Hopf.

Kalman (1960), (1963) e Kalman-Bucy (1961) hanno esteso i risultati di Wiener e Kolmogorov ai processi aleatori non stazionari.

Sfortunatamente molta della letteratura sui filtri di Kalman-Bucy è pubblicata in riviste di ingegneria e con un linguaggio poco familiare agli statistici. Tale fatto ha portato alla scarsa applicazione in statistica di questi risultati. Il filtro di Kalman-Bucy è stato impiegato nel contesto di problemi di teoria del controllo e nell'analisi delle serie storiche. La generalità dei risultati e la facile programmazione su computer ha reso molto popolare il filtro di Kalman-Bucy. Tali risultati vengono utilizzati per una varietà di fini anche se la loro funzione di base è quella di stimare lo stato di un sistema. Per rendersi conto delle possibilità applicative di tale filtro basta scorrere le riviste degli ultimi venti anni. I lavori in cui viene impiegato sono numerosissimi e vanno dalla determinazione delle orbite del Voyager spacecraft alla stima della quantità di bestiame in Francia, da problemi di navigazione all'oceanografia, dalla teoria della credibilità alla teoria delle aspettative razionali in economia (si veda, come esempio, il numero speciale di IEEE Transactions on Automatic Control (1983) dedicato alle applicazioni del filtro di Kalman-Bucy).

Con questo lavoro ci proponiamo di:

- a) fornire ai lettori non iniziati le motivazioni per poter apprezzare ed applicare il filtro di Kalman-Bucy;
- b) fornire ai lettori con più esperienza su tali argomenti una rassegna delle tecniche e dei problemi di interesse corrente.

Il lavoro è evidentemente di natura essenzialmente descrittiva. Lo scopo è quello di mettere in evidenza come i risultati di Kalman-Bucy possono essere apprezzati dagli statistici qualora ci si ponga in un contesto bayesiano.

2. SCHEMA DEL LAVORO.

Le linee generali del lavoro sono le seguenti. Nel paragrafo 3 introdurremo i modelli lineari dinamici e definiremo i problemi statistici inerenti tali modelli. Presenteremo nel paragrafo 4 il procedimento ricorsivo (filtro di Kalman-Bucy) per poter effettuare inferenze sullo stato del sistema. La determinazione del filtro verrà presentata seguendo l'impostazione bayesiana. Nel paragrafo 5 verranno presentate alcune proprietà sul filtro. Il paragrafo 6 sarà dedicato alla presentazione di alcuni modelli lineari dinamici e analizzeremo la natura di tali modelli e le loro possibilità applicative.

Nel paragrafo 7 verrà determinato il filtro di Kalman -Bucy utilizzando il teorema della correlazione normale. Nel paragrafo 8 presenteremo una generalizzazione del modello lineare dinamico e verranno derivate le equazioni ricorsive per il filtro. Il paragrafo 9 servirà ad analizzare il problema dell'interpolazione, mentre il paragrafo 10 quello della previsione di più periodi.

3. I MODELLI LINEARI DINAMICI.

In questo numero ci proponiamo di introdurre una classe di modelli che permettono di:

- a) evitare nella costruzione del modello l'ipotesi di stazionarietà;
- b) costruire un modello che aiuti a spiegare la struttura stessa del fenomeno;
- c) introdurre informazioni di tipo soggettivo ad ogni stadio di costruzione del modello.

Come vedremo i modelli lineari dinamici (MLD) soddisfano ai "desiderata" elencati. Tali modelli hanno recentemente ricevuto una considerevole attenzione nella letteratura statistica (si veda Harrison-Stevens(1976)). Il modello lineare dinamico è descritto da due equazioni. La prima equazione mette in evidenza come vengono generate le osservazioni ed è detta "equazione delle osservazioni" :

$$Y_t = F_t \theta_t + v_t \quad (3.1)$$

dove

- t = indice del tempo ($t= 1,2,\dots$)
- Y_t = vettore ($m \times 1$) delle osservazioni del processo fatte al tempo t .
- θ_t = vettore ($n \times 1$) dei parametri del processo al tempo t .
- F_t = matrice ($m \times n$) di variabili indipendenti note al tempo t .
- v_t = vettore ($m \times 1$) di variabili aleatorie normali con media nulla e varianza nota al tempo t , data da $E(v_t v_t') = V_t$.

L'equazione (3.1) ricorda l'equazione del modello lineare generale. Le differenze derivano:

- a) dal fatto che il vettore delle osservazioni è definitivamente associato al tempo.
- b) dal fatto che i coefficienti della regressione non sono assunti costanti ma variano con il tempo.

L'aspetto dinamico dei parametri è incorporato nella "equazione del sistema" :

$$\theta_t = G_t \theta_{t-1} + w_t \quad (3.2)$$

dove

G_t = matrice ($m \times n$) che spiega come i parametri si sviluppano nel tempo.

w_t = indica lo sviluppo stocastico dei parametri nel tempo ed ha una distribuzione normale multivariata di media nulla e matrice di varianze e covarianze W_t .

In aggiunta alle usuali assunzioni del modello lineare inerenti i termini di errore, assumiamo che v_t e w_t siano stocasticamente indipendenti, ed inoltre che le matrici di varianze e covarianze V_t e W_t siano note.

Le equazioni (3.1) e (3.2) definiscono il MLD. Come si può notare tale modello è un processo markoviano formato da una equazione delle osservazioni che specifica come il processo Y_t è dipendente stocasticamente dal processo dei parametri θ_t ed una equazione del sistema che specifica come il processo dei parametri si evolve nel tempo.

Tali modelli sono detti anche "modelli stato - spazio" e si basano sulla proprietà di Markov. Tale proprietà implica l'indipendenza del futuro di un processo dal suo passato, dato il presente.

Ossia,

$$Y_t \perp\!\!\!\perp Y_{t-1} \mid \theta_t$$

(il simbolo $\perp\!\!\!\perp$ indica indipendenza stocastica)

In un tale modello lo stato contiene tutte le informazioni del passato che sono necessarie a prevedere il futuro.

Facciamo osservare che il modello presentato può, evidentemente, essere generalizzato considerando per v_t e w_t una distribuzione congiunta e valori attesi diversi da zero.

Nel prosieguo fermeremo la nostra attenzione al modello definito dalle equazioni (3.1) e (3.2).

I problemi statistici inerenti il MLD riguardano la determinazione di stime ottimali di θ_t sulla base delle osservazioni del processo.

Si possono distinguere tre problemi: filtraggio, *smoothing* interpolazione, previsione.

Per filtraggio si intende il problema della stima per ogni t del vettore non osservabile θ_t mediante le osservazioni $y_0, y_1, \dots, y_t$. Nel seguito indicheremo: $y^t = (y_0, y_1, \dots, y_t)$.

Per interpolazione si intende la stima del vettore θ_t per mezzo delle osservazioni $y^h = (y_0, y_1, \dots, y_h)$ con $h \geq t$.

Per previsione si intende la stima dei vettori θ_h e y_h mediante le osservazioni y^t con $h > t$.

RIFERIMENTI BIBLIOGRAFICI

Harrison-Stevens (1975), (1976), Smith (1980), Jazwinski (1970), Anderson-Moore (1979).

4. IL FILTRO DI KALMAN- BUCY : IMPOSTAZIONE BAYESIANA .

L'espressione " filtro di Kalman-Bucy " si riferisce ad un procedimento per effettuare inferenze su θ_t . Le inferenze su θ_t possono essere fatte mediante la diretta applicazione del teorema di Bayes, determinando la distribuzione finale:

$$p(\theta_t | y^t) = \frac{p(y_t | \theta_t, y^{t-1}) p(\theta_t | y^{t-1})}{\int_{\theta} p(y_t | \theta, y^{t-1}) p(\theta | y^{t-1}) d\theta} \quad (4.1)$$

dove $p(y_t | \theta_t, y^{t-1})$ indica la verosimiglianza e $p(\theta_t | y^{t-1})$ la distribuzione iniziale per θ .

Nell'approccio bayesiano, come è noto, la determinazione della distribuzione finale esaurisce il problema dell'inferenza. Esistono, tuttavia, situazioni in cui è utile determinare una conveniente sintesi della distribuzione finale. Una possibile determinazione di questa sintesi è suggerita dalla teoria delle decisioni. In tal caso il procedimento conduce ad attribuire un valore al parametro in modo che risulti minimo il rischio associato a questo tipo di decisione. Il valore che realizza tale proprietà è detto " stimatore bayesiano " .

Lo stimatore bayesiano di θ_t , in presenza di funzioni di danno quadratiche, è dato da:

$$m_t = \int \theta dH(\theta_t | y^t) \quad (4.2)$$

vale a dire il valore atteso della distribuzione finale. Infatti il rischio

$$R(p(\theta_t | y^t), d) = \int (\theta_t - d)^2 dH(\theta_t | y^t)$$

è minimo quando d coincide con il valore atteso di $p(\theta_t | y^t)$.

L'aspetto ricorsivo del procedimento è messo in luce dalle relazioni seguenti:

$$\begin{aligned} m_t &= E(\theta_t | Y^t = y^t) = \int \theta_t p(\theta_t | y^t) d\theta_t \\ &= \int \theta_t p(\theta_t | y^{t-1}, y_t) d\theta_t \\ &= \frac{\int \theta_t p(y_t | \theta_t, y^{t-1}) p(\theta_t | y^{t-1}) d\theta_t}{\int p(y_t | \theta_t, y^{t-1}) p(\theta_t | y^{t-1}) d\theta_t} \end{aligned}$$

Ponendo

$$\int p(y_t | \theta_t, y^{t-1}) p(\theta_t | y^{t-1}) d\theta_t = K_t^{-1}$$

si ha

$$m_t = K_t \int \theta_t p(y_t | \theta_t, y^{t-1}) p(\theta_t | y^{t-1}) d\theta_t \quad (4.3)$$

poichè y_t , dato θ_t , è indipendente da y^{t-1} si ha che la (4.3) è uguale a :

$$\begin{aligned} m_t &= K_t \int \theta_t p(y_t | \theta_t) \left[\int p(\theta_t, \theta_{t-1} | y^{t-1}) d\theta_{t-1} \right] d\theta_t \\ &= K_t \int \theta_t p(y_t | \theta_t) \left[\int p(\theta_t | \theta_{t-1}) p(\theta_{t-1} | y^{t-1}) d\theta_{t-1} \right] d\theta_t \quad (4.4) \end{aligned}$$

Il procedimento può evidentemente continuare.

Indichiamo ora la procedura per determinare la distribuzione finale di θ_t .

Determiniamo dapprima la distribuzione iniziale $p(\theta_t | y^{t-1})$

Facciamo osservare che il procedimento ricorsivo inizia al tempo $t=0$.

Assumiamo allora che $\theta_0 \sim N(m_0, C_0)$ (4.5)

Al tempo $t-1$, la nostra conoscenza su θ_{t-1} può, evidentemente, essere espressa dalla distribuzione finale di θ_{t-1} ,

$$(\theta_{t-1} | y^{t-1}) \sim N(m_{t-1}, C_{t-1}) \quad (4.6)$$

dove m_{t-1} e C_{t-1} sono rispettivamente il valore atteso e la matrice di varianze e covarianze di $(\theta_{t-1} | y^{t-1})$.

Prima di osservare y_t , θ_t è governata dall'equazione del sistema

combinazione lineare di θ_{t-1}

$$\theta_t = G_t \theta_{t-1} + w_t$$

pertanto la nostra conoscenza iniziale su θ_t è rappresentata da:

$$(\theta_t | y^{t-1}) \sim N(G_t m_{t-1}, G_t C_{t-1} G_t' + W_t) \quad (4.7)$$

La (4.7) fornisce la previsione di θ_t note le osservazioni fino al tempo $t-1$. Con funzione di danno quadratica la previsione di θ_t sarà data da :

$$E(\theta_t | y^{t-1}) = G_t m_{t-1}$$

mentre

$$\text{Cov}(\theta_t | y^{t-1}) = G_t C_{t-1} G_t' + W_t$$

D'ora innanzi indicheremo:

$$R_t = G_t C_{t-1} G_t' + W_t$$

Ricordando che y_t , dato θ_t , è indipendente da y^{t-1} e tenendo presente l'equazione delle osservazioni (3.1) si ha:

$$(y_t | \theta_t) \sim N(F_t \theta_t, V_t) \quad (4.8)$$

La distribuzione finale di $(\theta_t | y^t)$ (si veda Lindley (1965), Lindley-Smith (1972)) dato y^t , G_t , m_{t-1} , R_t è normale multivariata con valore atteso

$$m_t = (R_t^{-1} + F_t' V_t^{-1} F_t)^{-1} (F_t' V_t^{-1} y_t + R_t^{-1} G_t m_{t-1}) \quad (4.9)$$

e matrice di varianze e covarianze

$$C_t = (R_t^{-1} + F_t' V_t F_t)^{-1} \quad (4.10)$$

Applicando ora un classico risultato di algebra matriciale (si veda Rao(1973) pag.33 e Smith (1973) *) si ottiene facilmente:

$$m_t = G_t m_{t-1} + R_t F_t' (V_t + F_t R_t F_t')^{-1} (y_t - F_t G_t m_{t-1}) \quad (4.11)$$

$$C_t = R_t - R_t F_t' (V_t + F_t R_t F_t')^{-1} F_t R_t \quad (4.12)$$

Le espressioni (4.11) e (4.12) forniscono il filtro di Kalman-Bucy.

Dalla (3.1) e dalla (4.6) è agevole verificare che

$$y_t | y^{t-1} \sim N (F_t G_t m_{t-1} , F_t R_t F_t' + V_t) \quad (4.13)$$

Naturalmente la previsione di y_t , note le osservazioni fino al tempo $t-1$ (previsione di un passo) è:

$$\hat{y}_t = E(y_t | y^{t-1}) = F_t E(\theta_t | y^{t-1}) = F_t G_t m_{t-1} \quad (4.14)$$

L'errore che si commette nel prevedere y_t dal tempo $t-1$ è dato da:

$$e_t = y_t - \hat{y}_t = y_t - F_t G_t m_{t-1} \quad (4.15)$$

Si può facilmente mostrare che la distribuzione finale di $(\theta_t | y^t)$ può anche essere espressa in funzione dell'errore di previsione nel modo

$$(*) \quad (M^{-1} + A R^{-1} B)^{-1} = M - M A (R + B M A)^{-1} B M$$

seguinte:

$$p(\theta_t | y^t) = p(\theta_t | e_t, y^{t-1}) \propto p(e_t | \theta_t, y^{t-1}) p(\theta_t | y^{t-1}). \quad (4.16)$$

Nella (4.16) la distribuzione iniziale di θ_t è data dalla (4.7) mentre la distribuzione di $(e_t | \theta_t, y^{t-1})$ può essere facilmente ottenibile con le seguenti osservazioni.

Essendo

$$e_t = y_t - F_t G_t m_{t-1}$$

dalla (3.1) si ha

$$e_t = F_t (\theta_t - G_t m_{t-1}) + v_t$$

Dal fatto che

$$v_t \sim N(0, V_t)$$

si deduce che

$$(e_t | y^{t-1}, \theta_t) \sim N(F_t (\theta_t - G_t m_{t-1}), V_t) \quad (4.17)$$

Applicando ora la formula di Bayes ed il lemma matriciale di inversione citato in precedenza si ottengono i risultati dati nelle espressioni (4.11) e (4.12)

La natura ricorsiva dell'algoritmo è importante, poiché, come è stato messo in evidenza, la distribuzione finale può essere calcolata mediante il valore della osservazione al tempo t e la distribuzione $(\theta_{t-1} | y^{t-1})$. Per meglio sottolineare la ricorsività dell'algoritmo converrà scrivere la (4.11) e la (4.12) nel modo seguente:

$$m_t = G_t m_{t-1} + A_t e_t \quad (4.18)$$

$$C_t = R_t - A_t B_t A_t' \quad (4.19)$$

dove

$$A_t = R_t F_t' B_t^{-1} \quad (4.20)$$

$$B_t = F_t R_t F_t' + V_t \quad (4.21)$$

$$R_t = G_t C_{t-1} G_t' + W_t \quad (4.22)$$

$$e_t = y_t - \hat{y}_t = y_t - F_t G_t^m m_{t-1} \quad (4.23)$$

In conclusione il procedimento da seguire nelle applicazioni è il seguente:

- (1) selezionare il modello scegliendo F_t , G_t , V_t , W_t
- (2) esprimere la distribuzione iniziale di θ_0 scegliendo m_0 e C_0
- (3) prevedere y_{t+1} utilizzando la (4.14)
- (4) osservare y_{t+1} ed aggiornare la stima utilizzando la (4.18)
- (5) ritornare al punto (3).

RIFERIMENTI BIBLIOGRAFICI

Harrison- Stevens (1971), (1975), (1976) ; Lindley (1965) ; Meinhold-Singpurwalla (1983) ; Rao (1973) ; Smith (1980) ; Verral (1983)
 Lindley-Smith (1972) ; Wegman (1982) ; Jazwinski (1970) ; Aoki (1967)
 Anderson-Moore (1979) ; Leoni (1979).

5. PROPRIETA' ED OSSERVAZIONI SUL FILTRO DI KALMAN-BUCY.

5.1 La matrice guadagno.

Il filtro di Kalman -Bucy è stato ottenuto originariamente mediante la stima dei minimi quadrati. La formulazione bayesiana da noi presentata porta evidentemente agli stessi risultati ma fornisce la possibilità di effettuare inferenze mediante la distribuzione di probabilità di θ_t dato y^t . Il procedimento fornisce la stima bayesiana ottenuta in forma "sequenziale". Al tempo $t=0$ si ha che la stima è determinata interamente dall'opinione iniziale; dopo aver osservato la prima osservazione, tale valore viene inserito nella stima di θ , ossia in m_1 e così via.

Si può dire pertanto che il processo determina in maniera iterativa la distribuzione iniziale $p(\theta_t | y^{t-1})$ e quella finale $p(\theta_t | y^t)$. Il filtro di Kalman è pertanto un metodo esplicito per adattare le stime e le previsioni man mano che aumentano le informazioni.

L'adattamento della stima di θ_t avviene mediante la "matrice guadagno"

$$A_t = R_t^{-1} F_t' (V_t + F_t R_t^{-1} F_t')^{-1}.$$

Tale matrice non dipende dalle osservazioni y^t . Se A_t è "piccola" vuol dire che si assegna poco peso alle osservazioni correnti. Dalla definizione si può vedere che A_t è determinata mediante F_t , R_t , V_t . Un valore elevato di V_t porta ad una matrice A_t "piccola" e pertanto le osservazioni correnti hanno poco peso. Tale circostanza molto intuitiva, infatti se y_t è molto variabile allora incide poco sulla stima di θ_t .

5.2 LA PROPRIETA' INNOVATIVA.

Sia

$$\eta_t = \theta_t - E(\theta_t | y^{t-1})$$

l'errore di stima del parametro θ .

L'errore che si commette nel prevedere y_t da $t-1$ è:

$$\begin{aligned} e_t &= y_t - F_t E(\theta_t | y^{t-1}) \\ &= y_t - F_t G_t m_{t-1} \end{aligned}$$

Dalla (3.1) si ha

$$\begin{aligned} e_t &= F_t (\theta_t - G_t m_{t-1}) + v_t \\ &= F_t \eta_t + v_t. \end{aligned}$$

Mostriamo ora che e_t è un processo gaussiano "white noise"

$$E(e_t e_k') = E\{(F_t \eta_t + v_t)(F_k \eta_k + v_k)'\}$$

Per $t > k$, v_t è indipendente da η_k e v_k pertanto

$$\begin{aligned} E(e_t e_k') &= E\{F_t \eta_t (F_k \eta_k + v_k)'\} \\ &= E\{F_t \eta_t (y_k - F_k G_k m_{k-1})'\} \end{aligned}$$

Il principio di ortogonalità stabilisce che η_t è ortogonale a $(y_j, j < t)$

Poiché $E(\theta_k | y^{k-1})$ dipende solamente da $(y_j, j < k)$ possiamo concludere che

$$E(e_t e_k') = 0 \quad \text{per } t > k$$

In maniera analoga

$$E(e_t e_k') = 0 \quad \text{per } t < k$$

Per $t = k$

$$E(e_t e_t') = F_t R_t F_t' + V_t = V(\theta_t | y^{t-1})$$

Come si vede e_t è una combinazione lineare di variabili aleatorie normali pertanto si distribuirà anch'essa secondo una normale.

e_t viene detto PROCESSO INNOVATIVO. Tale processo rappresenta la nuova informazione contenuta in y_t . Kailath (1968) ha dimostrato che e_t e y_t contengono la stessa informazione statistica e sono pertanto equivalenti. Schweppe(1965) ha inoltre dimostrato che e_t può essere ottenuta da y_t mediante il procedimento di ortogonalizzazione di Gram-Schmidt.

Per concludere facciamo notare che

$$\eta_t = \theta_t - E(\theta_t | y^{t-1})$$

non è un processo innovativo, dato che

$$E(\theta_t | y^{t-1}) \neq E(\theta_t | \theta_{t-1}, \dots, \theta_0)$$

5.3 STABILITA'

Kalman(1963) ed altri autori hanno dimostrato che il filtro di Kalman-Bucy possiede la proprietà della stabilità globale asintotica per un modello completamente controllabile ed osservabile.

Diremo che il MLD è completamente osservabile se e solo se la matrice $(n \times mn)$

$$[F' ; G'F' ; \dots ; (G^{n-1})'F']$$

ha rango n .

Diremo che il MLD è completamente controllabile se e solo se la matrice $n \times m^2$

$$[I ; G ; G^2 ; \dots ; G^{n-1}]$$

ha rango n .

L'interpretazione fisica della proprietà di osservabilità è la seguente.

Si consideri una formulazione deterministica del modello con $v_t=0$ e $w_t=0$ per ogni t ed una successione di osservazioni $y(1), y(2), \dots, y(n)$. Allora l'osservabilità può essere descritta:

" θ_t è osservabile se la conoscenza di y^{t-1} determina completamente θ_t "

La controllabilità può essere pensata come segue:

" θ_t è controllabile se per tutti i possibili θ_0 esiste un vettore finito di controllo che permette di raggiungere θ_t al tempo t "

Il significato pratico di queste due condizioni è che la matrice di varianze e covarianze del filtro, C_t , e la matrice guadagno, A_t , portano a valori costanti di θ_t , indipendentemente dalle condizioni iniziali su θ .

Le condizioni di completa controllabilità ed osservabilità assicurano la stabilità globale asintotica del filtro (veda: Kalman-Bucy (1961), Deyst-Price(1968)).

I concetti di osservabilità e di controllabilità assumono notevole rilevanza nella teoria della realizzazione stocastica.

parzenisky pg 240

Abbiamo messo in evidenza in precedenza che per poter applicare il filtro di Kalman-Bucy è necessario avere un modello scritto nella forma di modello lineare dinamico. Molto spesso invece ci si trova di fronte ad un modello input-output e pertanto il problema preliminare è quello di costruire un modello lineare dinamico equivalente. Nel contesto dei processi stocastici il problema diviene il seguente:

Sia $y(t)$, $t \in T \subseteq \mathbb{R}$ un dato processo stocastico con spazio dei risultati Y . Si desidera costruire uno spazio X , un processo di Markov $x(t)$, $t \in T \subseteq \mathbb{R}$ con spazio dei risultati X , ed una funzione $h: X \rightarrow Y$ tale che $h(x(t))$ sia in qualche senso equivalente al processo originario $y(t)$. Tutto ciò è detto "problema della realizzazione stocastica".^{Ricordiamo} L'esposizione classica del problema limitatamente al caso dei processi stazionari a tempo discreto:

Dato un processo gaussiano stazionario Y_t diciamo che un processo X_t è una rappresentazione markoviana di Y_t se Y_t verifica il modello espresso nelle equazioni (3.1) e (3.2) dove F e G sono matrici costanti di dimensioni opportune e con F asintoticamente stabile. Si dice inoltre che X è una rappresentazione stocastica minimale allorchè la sua dimensione è minimale. La teoria della realizzazione stocastica è dovuta a Doob (1944) anche se è stato Kalman (1965) ad aver posto chiaramente le basi della teoria. L'interpretazione probabilistica è stata introdotta da Akaike (1974), (1975) che ha studiato due particolari rappresentazioni markoviane minimali. Il lavoro iniziato da Akaike è stato proseguito in diverse direzioni da Lindquist-Picci (1979), Picci (1976), Rozanov (1976) e Ruckebusch (1977).

5.4 ALCUNE ULTERIORI CONSIDERAZIONI.

I lavori di Wiener e Kolmogorov si basavano sulle seguenti assunzioni:

- (a) la serie osservata è stazionaria o può essere ricondotta alla stazionarietà con opportune trasformazioni;
- (b) i previsori sono funzioni lineari delle osservazioni passate;
- (c) il criterio per verificare l'ottimalità è quello dell'errore quadratico medio.

Nei paragrafi precedenti abbiamo presentato l'approccio bayesiano al filtro di Kalman - Bucy. In tale impostazione non ha alcuna rilevanza l'ipotesi di stazionarietà. Si ottiene inoltre la distribuzione finale di $(\theta_t | y^t)$ e solo se si vuole sintetizzare tale distribuzione si può scegliere una opportuna funzione di danno e minimizzare il rischio ad esso associato. E' immediato notare che anche nel filtro di Kalman i previsori sono funzioni lineari delle osservazioni.

Le assunzioni di base per poter applicare il filtro di Kalman sono molteplici, la principale è quella di avere un modello lineare dinamico, tuttavia in molti problemi tali modelli non sono disponibili, di qui lo sforzo di ottenere tali modelli per poter applicare la metodologia proposta da Kalman. Val la pena di sottolineare che in alcune circostanze è molto più facile procedere direttamente che non ricercare di rappresentarle come modello lineare dinamico (si veda a tal proposito Kailath (1981)).

I modi per arrivare alle equazioni (4.18) (4.19) che definiscono il filtro sono svariati. La ragione che ha spinto molti autori a ricercare queste diverse procedure è che nella pratica questi risultati, sebbene teoricamente equivalenti, hanno differenti proprietà numeriche.

Il filtro di Kalman-Bucy, come abbiamo visto, richiede una esatta conoscenza della matrice di varianze e covarianze sia degli errori delle osservazioni (V_t) sia degli errori del sistema (W_t).

In molte situazioni pratiche però V_t e W_t non sono note. Heffes(1966),

Nishimura (1967) hanno analizzato le conseguenze di questo fatto sul filtro. Molti altri autori hanno proposto schemi per identificare V_t e W_t (alcuni di questi li analizzeremo nel prossimo paragrafo) un procedimento interessante è dovuto a Mehra(1970) e si basa sull'assunzione che il sistema sia costante, completamente controllabile ed osservabile.

Concludiamo questo paragrafo mettendo in evidenza quelli che, a nostro avviso, sono i vantaggi che si hanno applicando la metodologia proposta:

- (1) permette di considerare informazioni a priori strutturali e numeriche;
- (2) permette di considerare parametri variabili nel tempo;
- (3) permette di non assumere la stazionarietà;
- (4) permette di individuare eventuali cambiamenti discontinui nella serie;
- (5) permette di effettuare i conti senza grossi oneri;
- (6) permette, in alcune circostanze, di non assumere la normalità delle distribuzioni.

RIFERIMENTI BIBLIOGRAFICI

Anderson-Moore(1979) ; Heffes(1966) ; Jazwinski (1970) ; Kailath (1968),(1981)
Mehra(1970),(1979); Nishimura(1967) , Prestley (1976), Smith(1980).
Kalman-Bucy(1961) ; Deyst-Price(1968); Kalman (1963), (1965) , Doob (1944)
Akaike(1974),(1975); Lindquist-Picci(1979); Picci (1976); Rozanov(1976);
Ruckebusch(1977).

6. ALCUNI ESEMPI DI MODELLI LINEARI DINAMICI.

In questo paragrafo presenteremo alcuni modelli lineari dinamici che differiscono tra loro per la scelta di F_t e di G_t . Analizzeremo la natura di tali modelli e le loro possibilità applicative.

Assumiamo che il modello sia completamente definito da: F_t, G_t, V_t, W_t .

6.1. Il modello stabile (steady model)

Tale modello si utilizza allorché i dati mostrano nessuna consistenza ad un trend (crescente o decrescente) ma variano aleatoriamente attorno ad un livello costante.

Un esempio può essere dato dalla domanda individuale, non stagionale, di un prodotto. Il modello stabilisce che la quantità venduta sia ragionevolmente stabile e che il consumo, in un dato periodo di tempo, sia uguale a quello del periodo precedente con una leggera variazione aleatoria in più o in meno. La domanda osservata rappresenta il vero bisogno dei consumatori con un errore che riflette le variazioni nelle ordinazioni (tali variazioni sono assunte aleatorie).

Tale modello può essere rappresentato matematicamente nel modo seguente:

$$\text{(equ. osservazioni)} \quad y_t = \mu_t + v_t \quad v_t \sim N(0, V_t)$$

$$\text{(equ. sistema)} \quad \mu_t = \mu_{t-1} + w_t \quad w_t \sim N(0, W_t)$$

dove y_t è la domanda osservata al tempo t e μ_t può essere pensato come il bisogno dei consumatori mentre v_t e w_t sono variabili aleatorie distribuite normalmente con media zero e matrice di varianze e covarianze V_t e W_t . Inoltre v_t e w_t sono tra loro indipendenti.

μ_t è il vero livello non noto del processo al tempo t , v_t è l'errore di osservazione al tempo t e w_t è il livello di disturbo al tempo t .

v_t e w_t sono assunte note.

Assumiamo che al tempo $t=0$ si abbia una conoscenza che ci permetta di definire la nostra opinione iniziale

$$\mu_0 \sim N(m_0, C_0)$$

La distribuzione finale di μ_t dato $(y_0, y_1, \dots, y_t) = y^t$

è anch'essa normale :

$$\mu_t | y^t \sim N(m_t, C_t)$$

dove m_t e C_t sono ottenute ricorsivamente utilizzando il filtro di Kalman-Bucy.

$$\hat{y}_t = E(y_t | y^{t-1}) = m_{t-1}$$

$$e_t = y_t - \hat{y}_t$$

$$\text{var}(\mu_t | y^{t-1}) = C_{t-1} + W_t$$

$$\text{Var}(y_t | y^{t-1}) = v_t + (C_{t-1} + W_t)$$

$$m_t = m_{t-1} + A_t e_t$$

$$A_t = C_t v_t^{-1} \quad Q_t = A_t \text{ var } V_t$$

$$C_t^{-1} = v_t^{-1} + (C_{t-1} + W_t)^{-1}$$

Come si può facilmente verificare il modello stabile si ottiene dal modello generale (3.1) e (3.2) ponendo $F_t = G_t = 1$ per ogni t . Osserviamo inoltre che $A_t = C_t v_t^{-1}$ è l'ammontare per mezzo del quale il livello del sistema viene aggiornato alla luce dell'osservazione y_t . A_t (matrice guadagno) è noto come coefficiente di adattamento.

Illustriamo ora il metodo proposto utilizzando i dati riportati da Box-Jenkins (1970) (serie A pag. 525)

Assumiamo le seguenti informazioni iniziali:

$$\mu_0 \sim N(17, 1)$$

$$v_t \sim N(0, 0.25)$$

$$w_t \sim N(0, 25)$$

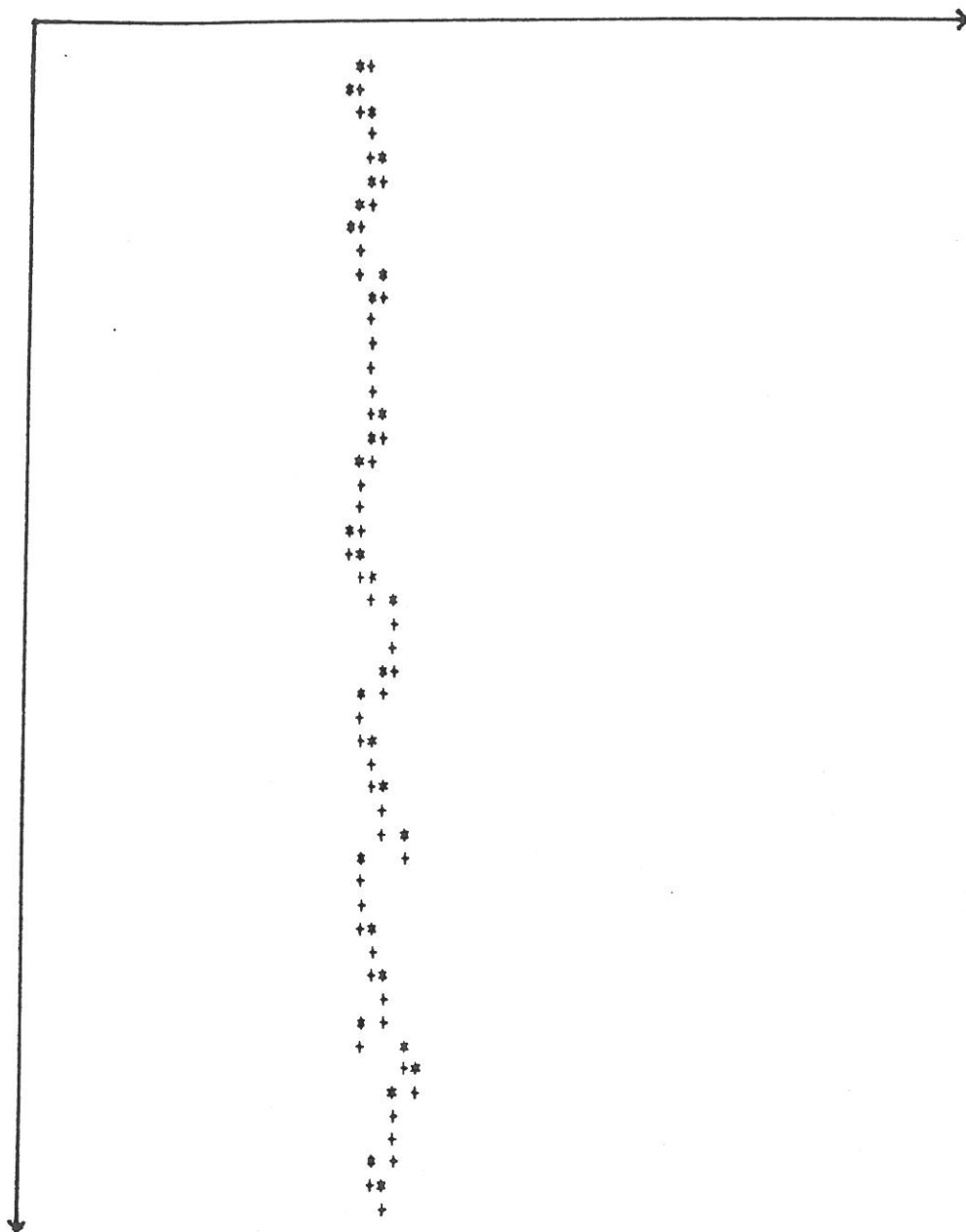
I risultati sono riportati nella tabella 1 e nella figura 1

NUM	Y1	M	NUM	Y1	M	NUM	Y1	M	NUM	Y1	M
***	****	*****	***	****	*****	***	****	*****	***	****	*****
1	17	17	51	17.6	17.296	101	16.5	16.898	150	16.8	17.001
2	16.6	17	52	16.9	17.597	102	17.2	16.504	151	17.2	16.802
3	16.3	16.604	53	16.7	16.907	103	16.4	17.193	152	17.2	17.196
4	16.1	16.303	54	16.8	16.702	104	17	16.408	153	17.4	17.2
5	17.1	16.102	55	16.8	16.799	105	17	16.994	154	17.2	17.398
6	16.9	17.09	56	17.2	16.8	106	16.7	17	155	16.9	17.202
7	16.8	16.902	57	16.8	17.196	107	16.2	16.703	156	16.8	16.903
8	17.4	16.801	58	17.6	16.804	108	16.6	16.205	157	17	16.801
9	17.1	17.394	59	17.2	17.592	109	16.6	16.596	158	17.4	16.998
10	17	17.103	60	16.6	17.204	110	16.5	16.6	159	17.2	17.396
11	16.7	17.001	61	17.1	16.606	111	16.6	16.501	160	17.2	17.202
12	17.4	16.703	62	16.9	17.095	112	16.6	16.599	161	17.1	17.2
13	17.2	17.393	63	16.6	16.902	113	17	16.6	162	17.1	17.101
14	17.4	17.202	64	18	16.603	114	17.1	16.996	163	17.1	17.1
15	17.4	17.398	65	17.2	17.986	115	17.1	17.099	164	17.4	17.1
16	17	17.4	66	17.3	17.208	116	16.7	17.1	165	17.2	17.397
17	17.3	17.004	67	17	17.299	117	16.8	16.704	166	16.9	17.202
18	17.2	17.297	68	16.9	17.003	118	16.3	16.799	167	16.9	16.903
19	17.4	17.201	69	17.3	16.901	119	16.6	16.305	168	17	16.9
20	16.8	17.398	70	16.8	17.296	120	16.8	16.597	169	16.7	16.999
21	17.1	16.806	71	17.3	16.805	121	16.9	16.798	170	16.9	16.703
22	17.4	17.097	72	17.4	17.295	122	17.1	16.899	171	17.3	16.898
23	17.4	17.397	73	17.7	17.399	123	16.8	17.098	172	17.8	17.296
24	17.5	17.4	74	16.8	17.697	124	17	16.803	173	17.8	17.795
25	17.4	17.499	75	16.9	16.809	125	17.2	16.998	174	17.6	17.8
26	17.6	17.401	76	17	16.899	126	17.3	17.198	175	17.5	17.602
27	17.4	17.598	77	16.9	16.999	127	17.2	17.299	176	17	17.501
28	17.3	17.402	78	17	16.901	128	17.3	17.201	177	16.9	17.005
29	17	17.301	79	16.6	16.999	129	17.2	17.299	178	17.1	16.901
30	17.8	17.003	80	16.7	16.604	130	17.2	17.201	179	17.2	17.098
31	17.5	17.792	81	16.8	16.699	131	17.5	17.2	180	17.4	17.199
32	18.1	17.503	82	16.7	16.799	132	16.9	17.497	181	17.5	17.398
33	17.5	18.094	83	16.4	16.701	133	16.9	16.906	182	17.9	17.499
34	17.4	17.506	84	16.5	16.403	134	16.9	16.9	183	17	17.896
35	17.4	17.401	85	16.4	16.499	135	17	16.9	184	17	17.009
36	17.1	17.4	86	16.6	16.401	136	16.5	16.999	185	17	17
37	17.6	17.103	87	16.5	16.598	137	16.7	16.505	186	17.2	17
38	17.7	17.595	88	16.7	16.501	138	16.8	16.698	187	17.3	17.198
39	17.4	17.699	89	16.4	16.698	139	16.7	16.799	188	17.4	17.299
40	17.8	17.403	90	16.4	16.403	140	16.7	16.701	189	17.4	17.399
41	17.6	17.796	91	16.2	16.4	141	16.6	16.7	190	17	17.4
42	17.5	17.602	92	16.4	16.202	142	16.5	16.601	191	18	17.004
43	16.5	17.501	93	16.3	16.398	143	17	16.501	192	18.2	17.99
44	17.8	16.51	94	16.4	16.301	144	16.7	16.995	193	17.6	18.198
45	17.3	17.787	95	17	16.399	145	16.7	16.703	194	17.8	17.606
46	17.3	17.305	96	16.9	16.994	146	16.9	16.7	195	17.7	17.798
47	17.1	17.3	97	17.1	16.901	147	17.4	16.898	196	17.2	17.701
48	17.4	17.102	98	17.1	17.098	148	17.1	17.395	197	17.4	17.205
49	16.9	17.397	99	16.7	17.1	149	17	17.103	198		17.398
50	17.3	16.905	100	16.9	16.704						

Tab. 1: Y1 = valori osservati ; M= previsioni un passo avanti.

Esempio modello stabile: dati Box-Jenkins.

Fig. 1. (*) valori osservati, (+) previsioni un passo avanti.



Vediamo ora alcune osservazioni sul modello stabile.

Osservazione 1.

Assumiamo che V_t e W_t siano costanti: V e W per ogni t . Si può osservare che C_t non dipende dalle osservazioni y^t ma solamente dal valore di t . Pertanto, quando $t \rightarrow +\infty$, $C_t \rightarrow C$ dove

$$C = 1/2 W (1 + 4(V/W))^{1/2} - 1)$$

Se al tempo t_0 poniamo $C_0 = C$ da cui $C_t = C$ allora

$$m_t = m_{t-1} + CV^{-1} (y_t - m_{t-1}).$$

Ponendo $k = 1 - CV^{-1}$ si ha

$$\begin{aligned} m_t &= (1-k)y_t + k m_{t-1} \\ &= (1-k) y_t + k((1-k) y_{t-1} + k m_{t-2}) \\ &= (1-k)(y_t + k y_{t-1} + k^2 y_{t-2} + \dots) \end{aligned}$$

cioè una media mobile esponenziale delle osservazioni. In questa forma limite il modello stabile è equivalente ad un modello ARIMA(0,1,1) (si veda Harrison (1975)).

Osservazione 2.

Il filtro di Kalman assume noto il modello delle varianze che, solitamente, nella pratica non sono note. Per poter applicare l'impostazione bayesiana questo è un problema cruciale. Harrison e Stevens (1974), (1976) hanno suggerito un approccio basato su un modello a più processi di classe I. Tale classe contiene molti modelli stabili M_{ij} caratterizzati dalle coppie di varianze (V_i, W_j) . Viene quindi specificata una distribuzione di probabilità dei modelli M_{ij} in corrispondenza dei punti (i,j) . Con il passare del tempo la probabilità

associata ai singoli modelli M_{ij} viene adattata mediante il teorema di Bayes e per effettuare le previsioni viene utilizzata una combinazione dei modelli ; allo stesso modo si procede per stimare le ultime varianze V e W . (Per una critica a tale metodo si veda O'Hagan(1976)).

Un diverso approccio è stato proposto da Cantarelis e Johnston (1979), (1982) . Questi autori propongono una impostazione che porta alle stime di massima verosimiglianza delle varianze, il cui adattamento avviene impiegando il filtro di Kalman. Si utilizza una classe di modelli che richiede un insieme di valori iniziali. Le varianze al limite sono indipendenti dalla scelta dei valori iniziali.

Osservazione 3.

Il modello stabile può essere anche definito per una classe molto ampia di processi al di fuori di quello normale (si veda Smith (1979). Per la classe dei modelli stabili , Souza e Harrison (1979), (1982) hanno proposto un metodo di previsione che generalizza quello di Harrison e Stevens(1976). Tale estensione è basata sul concetto di entropia e non richiede necessariamente l'assunzione di normalità.

6.2. Il modello di crescita lineare

Quando il livello del sistema cambia gradualmente, in altre parole quando il sistema ha una crescita allora conviene utilizzare un modello lineare di crescita dinamico.

Tale modello può essere rappresentato matematicamente nel modo seguente:

$$\begin{aligned} y_t &= \mu_t + \epsilon_t \\ \mu_t &= \mu_{t-1} + \beta_t + \delta \mu_t \\ \beta_t &= \beta_{t-1} + \delta \beta_t \end{aligned}$$

dove μ_t è il livello del processo al tempo t e β_t è l'incremento al tempo t .

Tale modello è espresso sotto forma di MLD come segue:

$$\begin{aligned} \text{equaz. osservazioni} \quad y_t &= (1, 0) \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} + v_t \\ \text{equaz. del sistema} \quad \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \delta \mu_t + \epsilon \beta_t \\ \delta \beta_t \end{pmatrix} \end{aligned}$$

L'equazione del sistema può scriversi anche nel modo seguente:

$$\theta_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \theta_{t-1} + w_t$$

dove

$$w_t \sim N \left[0, \begin{pmatrix} w_1 + w_2 & w_2 \\ w_2 & w_2 \end{pmatrix} \right]$$

con $w_1 = \text{Var}(\delta\mu_t)$, $w_2 = \text{Var}(\delta\beta_t)$ al tempo t .

Applichiamo ora i risultati del paragrafo 4.

In questo modello le matrici F_t e G_t assumono la forma:

$$F_t = (1, 0)$$

$$G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Assumiamo

$$\theta_0 \sim N\left(\begin{pmatrix} m_0 \\ b_0 \end{pmatrix}, C_0\right)$$

Avremo

$$(\theta_{t-1} | y^{t-1}) \sim N\left(\begin{pmatrix} m_{t-1} \\ b_{t-1} \end{pmatrix}, C_{t-1}\right)$$

e

$$(y_t | y^{t-1}) \sim N(m_{t-1} + b_{t-1}, B_t)$$

con

$$B_t = r_{11} + v_t$$

$$R_t = [r_{ij}] = G_t C_{t-1} G_{t-1}' + W_t$$

La previsione di un passo è evidentemente

$$\hat{y}_t = m_{t-1} + b_{t-1}$$

La distribuzione finale di θ_t dato y^t si ottiene facilmente ed è:

$$(\theta_t | y^t) \sim N\left(\begin{pmatrix} m_t \\ b_t \end{pmatrix}, C_t\right)$$

dove

$$m_t = m_{t-1} + b_{t-1} + A_{1,t} e_t$$

$$b_t = b_{t-1} + A_{2,t} e_t$$

$$C_t = R_t - A_t B_t A_t'$$

essendo

$$A_t = \begin{pmatrix} A_{1,t} \\ A_{2,t} \end{pmatrix} = \begin{pmatrix} r_{11}/B_t \\ r_{12}/B_t \end{pmatrix}$$

L'applicazione di tale modello è stata effettuata sull'indice generale dei prezzi al consumo per gli anni relativi al 1976 - 1982. (Fonte Istat)

I risultati sono esposti nella tabella 2 e nella figura 2.

Le assunzioni iniziali sono state le seguenti:

$$V_t = 25, m_0 = 200, b_0 = 0, C_0 = \begin{pmatrix} 100 & 5 \\ 5 & 5 \end{pmatrix}, W_0 = \begin{pmatrix} 1000 & 1 \\ 1 & 1 \end{pmatrix}$$

*****	YAC	Y	*****	YAC	Y	*****	YAC	Y
1	181.45	200	30	266.93	265.46	58	394.26	390.5
2	184.56	181.68	31	269.08	267.9	59	402.62	397.29
3	188.29	184.34	32	270.16	270.08	60	407.89	405.78
4	194.03	188.07	33	273.96	271.18	61	415.72	411.18
5	197.35	193.81	34	276.72	275	62	423.27	419.08
6	198.15	197.22	35	279.22	277.82	63	429.23	426.77
7	199.34	198.09	36	281.18	280.37	64	435.28	432.85
8	201.14	199.29	37	287.15	282.36	65	440.98	438.97
9	204.59	201.1	38	290.91	288.37	66	445.86	444.74
10	211.66	204.55	39	294.71	292.24	67	449.44	449.67
11	216.16	211.64	40	299.47	296.12	68	452.6	453.27
12	218.77	216.25	41	303.38	300.94	69	458.98	456.42
13	221.85	218.95	42	306.43	304.95	70	467.78	462.8
14	226.78	222.07	43	309.2	308.06	71	475.8	471.7
15	230.21	227.04	44	312.31	310.87	72	480.58	479.86
16	232.76	230.56	45	319.9	314.01	73	487.36	484.74
17	235.8	233.17	46	327.34	321.66	74	493.74	491.55
18	237.94	236.25	47	331.62	329.27	75	498.2	498.01
19	239.85	238.44	48	336.97	333.69	76	502.7	502.52
20	241.29	240.38	49	347.93	339.11	77	508.26	507.03
21	243.96	241.85	50	354.25	350.19	78	513.37	512.6
22	246.66	244.54	51	357.45	356.74	79	520.61	517.75
23	250.39	247.28	52	362.85	360.04	80	530.07	525.02
24	251.39	251.05	53	366.13	365.47	81	537.54	534.58
25	253.92	252.13	54	369.44	368.82	82	548.4	542.19
26	256.47	254.66	55	375.78	372.15	83	555.57	553.16
27	259.04	257.26	56	379.56	378.52	84	559.48	560.5
28	261.91	259.87	57	387.61	382.39	85		564.45
29	264.54	262.78						

Tab.2. YAC: valori osservati , Y: previsioni un passo avanti.

Esempio modello lineare di crescita:

Indice generale dei prezzi al consumo (1976- 1982)

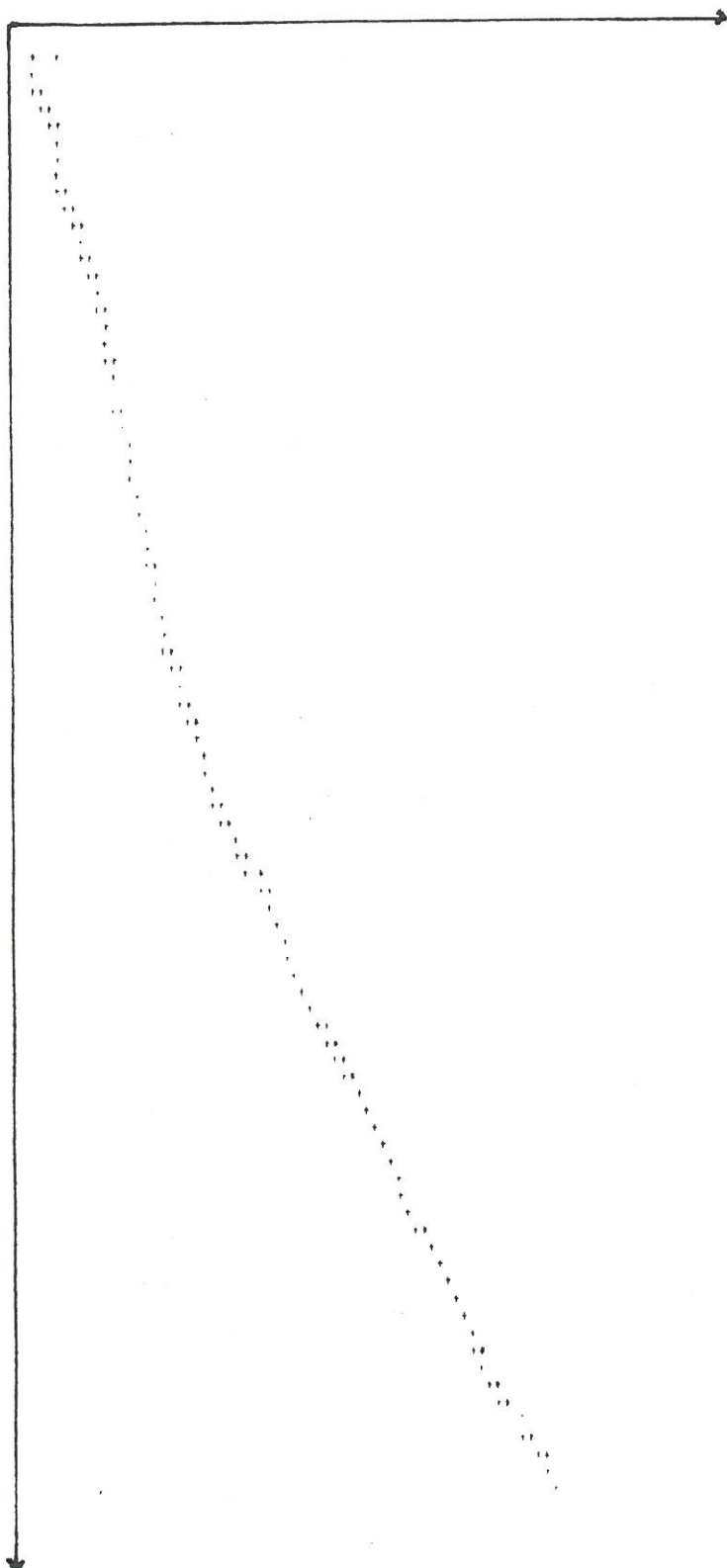


FIG. 2. (*) valori osservati , (+) previsioni un passo avanti.

6.2. 1. Estensioni del modello lineare di crescita.

Il modello introdotto si può estendere considerando alcune varianti.

Possiamo trovarci di fronte a quattro possibilità:

- 1) nessun cambiamento : i dati continuano ad essere generati da un modello in cui non avvengono cambiamenti nella crescita o nel livello.
- 2) presenza di un outlier : vi è una osservazione anomala, tuttavia non vi sono cambiamenti fondamentali nella natura dei dati.
- 3) cambio di livello : il sistema ha un cambiamento improvviso e dovrebbe continuare sulla nuova base con la stessa crescita.
- 4) cambio di crescita : vi è un cambiamento nella crescita del sistema.

Si possono così avere quattro modelli: M_1 , M_2 , M_3 , M_4 . Il modello M_1 è quello illustrato in precedenza.

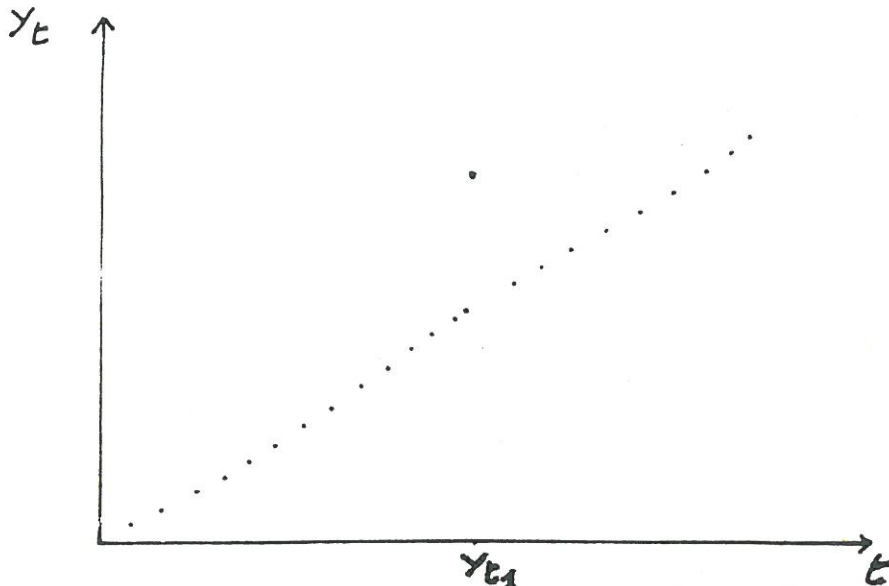
M_2 : Si può generare un outlier al tempo t_1 se y_{t_1} è un prodotto del seguente modello.

$$y_t = \mu_t + \Delta y_t + \delta y_t$$

$$\mu_t = \mu_{t-1} + \beta_t + \delta \mu_t$$

$$\beta_t = \beta_{t-1} + \delta \beta_t$$

L'impulso Δy_t fa sì che l'osservazione sia anomala. Se al tempo $t+1$ e successivi l'osservazione è generata secondo M_1 allora possiamo dire che si è prodotto un outlier.



E' immediato notare che i vari cambi di modelli si esprimono adeguatamente con cambi nelle varianze V e W .

M_2 dipende da un outlier pertanto da una varianza delle osservazioni molto grande: è necessario allora sostituire V con $V+V^0$ dove V^0 è una varianza molto grande.

Supponiamo che nel nostro esempio si sia prodotto un outlier alla 30 osservazione. Per risolvere il problema bisogna aumentare in corrispondenza di tale osservazione la varianza V . Noi abbiamo posto $V = 2500$. I risultati ottenuti si possono analizzare nella tavola 3 e nella fig. 3.

*****	YAC	Y	*****	YAC	Y	*****	YAC	Y
1	181.45	200	30	320	265.46	58	394.26	390.46
2	184.56	181.68	31	269.08	321.03	59	402.62	397.26
3	188.29	184.34	32	270.16	307.56	60	407.89	405.74
4	194.03	188.07	33	273.96	271.64	61	415.72	411.15
5	197.35	193.81	34	276.72	274.94	62	423.27	419.05
6	198.15	197.22	35	279.22	277.75	63	429.23	426.74
7	199.34	198.09	36	281.18	280.3	64	435.28	432.82
8	201.14	199.29	37	287.15	282.3	65	440.98	438.94
9	204.59	201.1	38	290.91	288.3	66	445.86	444.71
10	211.66	204.55	39	294.71	292.18	67	449.44	449.64
11	216.16	211.64	40	299.47	296.06	68	452.6	453.25
12	218.77	216.25	41	303.38	300.88	69	458.98	456.39
13	221.85	218.95	42	306.43	304.89	70	467.78	462.78
14	226.78	222.07	43	309.2	308.01	71	475.8	471.68
15	230.21	227.04	44	312.31	310.82	72	480.58	479.64
16	232.76	230.56	45	319.9	313.96	73	487.36	484.72
17	235.8	233.17	46	327.34	321.61	74	493.74	491.53
18	237.94	236.25	47	331.62	329.22	75	498.2	497.99
19	239.85	238.44	48	336.97	333.65	76	502.7	502.5
20	241.29	240.38	49	347.93	339.07	77	508.26	507.01
21	243.96	241.85	50	354.25	350.15	78	513.37	512.59
22	246.66	244.54	51	357.45	356.7	79	520.61	517.73
23	250.39	247.28	52	362.85	360	80	530.07	525
24	251.39	251.05	53	366.13	365.43	81	537.54	534.56
25	253.92	252.13	54	369.44	368.78	82	548.4	542.17
26	256.47	254.66	55	375.78	372.11	83	555.57	553.14
27	259.04	257.26	56	379.56	378.49	84	559.48	560.48
28	261.91	259.87	57	387.61	382.35	85		564.43
29	264.54	262.78						

Tab. 3. YAC= valori osservati, Y= previsioni un passo avanti

Esempio modello lineare di crescita con outlier.

Indice generale dei prezzi al consumo (1976-1982) con outlier alla 30 osservazione

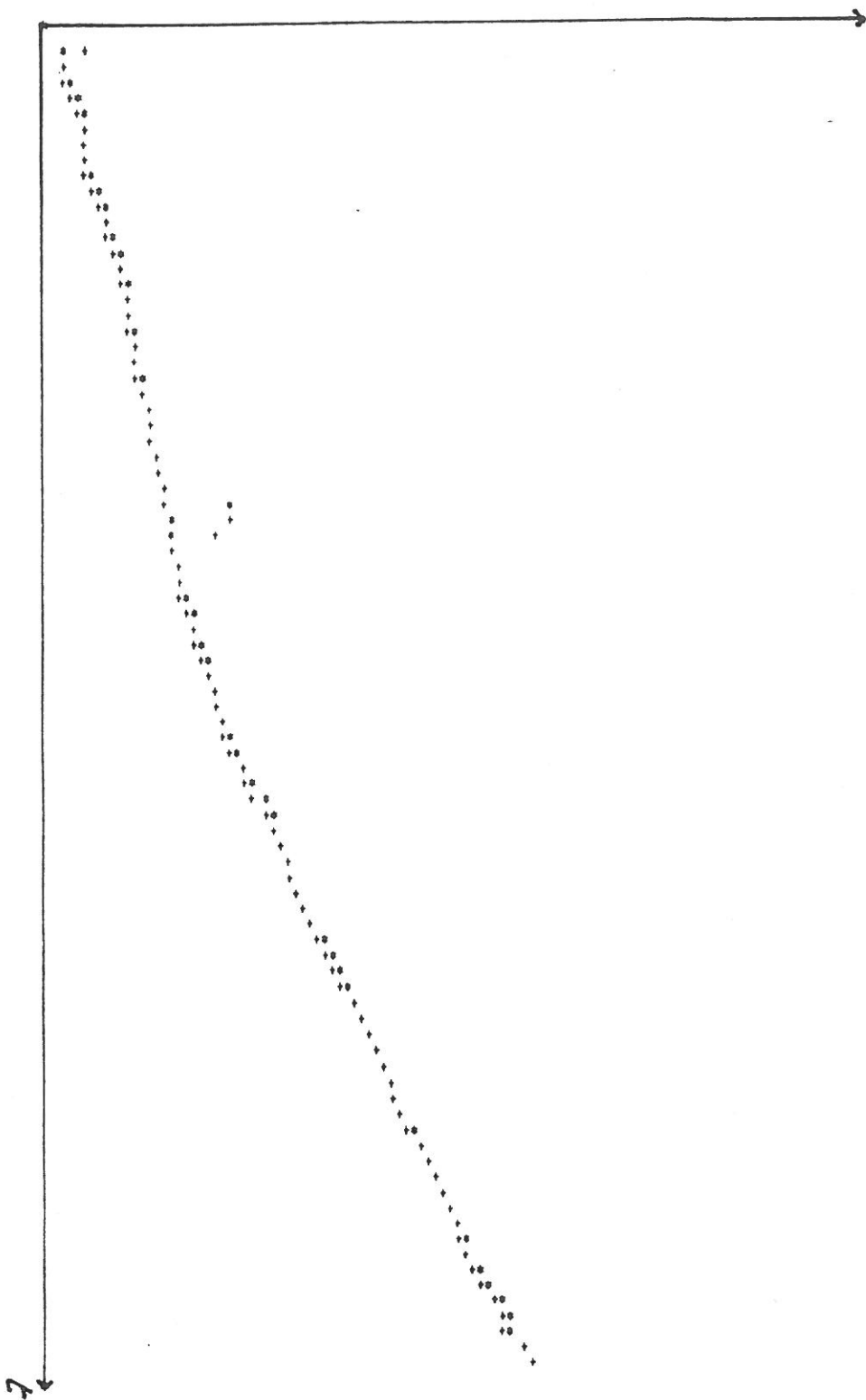


FIG. 3 . (*) valori osservati , (+) previsioni un passo avanti.

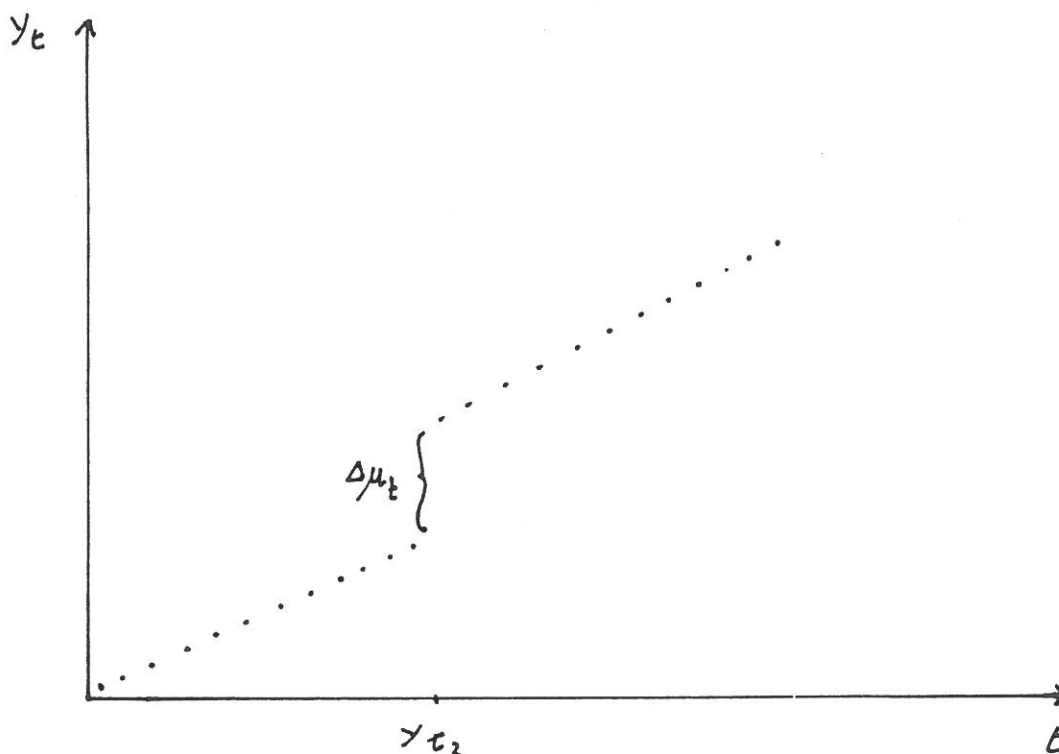
Il modello M_3 indica un cambiamento di livello al tempo t_2 e si ottiene se y_{t_2} è generata dal modello:

$$y_t = \mu_t + \delta y_t$$

$$\mu_t = \mu_{t-1} + \Delta\mu_t + \beta_t + \delta\mu_t$$

$$\beta_t = \beta_{t-1} + \delta\beta_t$$

Le osservazioni prima e dopo il tempo t sono ottenute mediante M_1 . M_3 dipende da un cambio nel livello del sistema e pertanto da una varianza grande per μ : è necessario sostituire w_1 con $w_1 + w_1^o$ dove w_1^o è una varianza molto grande.



Nel nostro esempio ipotizziamo un cambiamento di livello dopo la 50 esima osservazione. Assumiamo che la matrice di varianze e covarianze W venga così modificata (in corrispondenza della 50.esima oss.)

$$W = \begin{pmatrix} 50.000 & 1 \\ 1 & 1 \end{pmatrix}$$

I risultati sono esposti nella tav. 4 e nella FIG. 4.

****	****	Y	****	YAC	Y	****	YAC	Y
1	181.45	200	30	266.93	265.46	58	444.26	441.7
2	184.56	181.68	31	269.08	267.9	59	452.62	448.45
3	188.29	184.34	32	270.16	270.08	60	457.89	456.91
4	194.03	188.07	33	273.96	271.18	61	465.72	462.28
5	197.35	193.81	34	276.72	275	62	473.27	470.15
6	198.15	197.22	35	279.22	277.82	63	479.23	477.8
7	199.34	198.09	36	281.18	280.37	64	485.28	483.85
8	201.14	199.29	37	287.15	282.36	65	490.98	489.94
9	204.59	201.1	38	290.91	288.37	66	495.86	495.68
10	211.66	204.55	39	294.71	292.24	67	499.44	500.58
11	216.16	211.64	40	299.47	296.12	68	502.6	504.16
12	218.77	216.25	41	303.38	300.94	69	508.98	507.28
13	221.85	218.95	42	306.43	304.95	70	517.78	513.64
14	226.78	222.07	43	309.2	308.06	71	525.8	522.51
15	230.21	227.04	44	312.31	310.87	72	530.58	530.65
16	232.76	230.56	45	319.9	314.01	73	537.36	535.51
17	235.8	233.17	46	327.34	321.66	74	543.74	542.29
18	237.94	236.25	47	331.62	329.27	75	548.2	548.72
19	239.85	238.44	48	336.97	333.69	76	552.7	553.22
20	241.29	240.38	49	347.93	339.11	77	558.26	557.71
21	243.96	241.85	50	354.25	350.19	78	563.37	563.26
22	246.66	244.54	51	407.45	356.74	79	570.61	568.38
23	250.39	247.28	52	412.85	410.32	80	580.07	575.64
24	251.39	251.05	53	416.13	416.89	81	587.54	585.17
25	252.92	252.13	54	419.44	420.17	82	598.4	592.77
26	256.47	254.66	55	425.78	423.46	83	605.57	603.72
27	259.04	257.26	56	429.56	429.79	84	609.48	611.04
28	261.91	259.87	57	437.61	433.62	85		614.97
29	264.54	262.78						

Tab. 4. YAC = valori osservati, Y = previsioni un passo avanti.

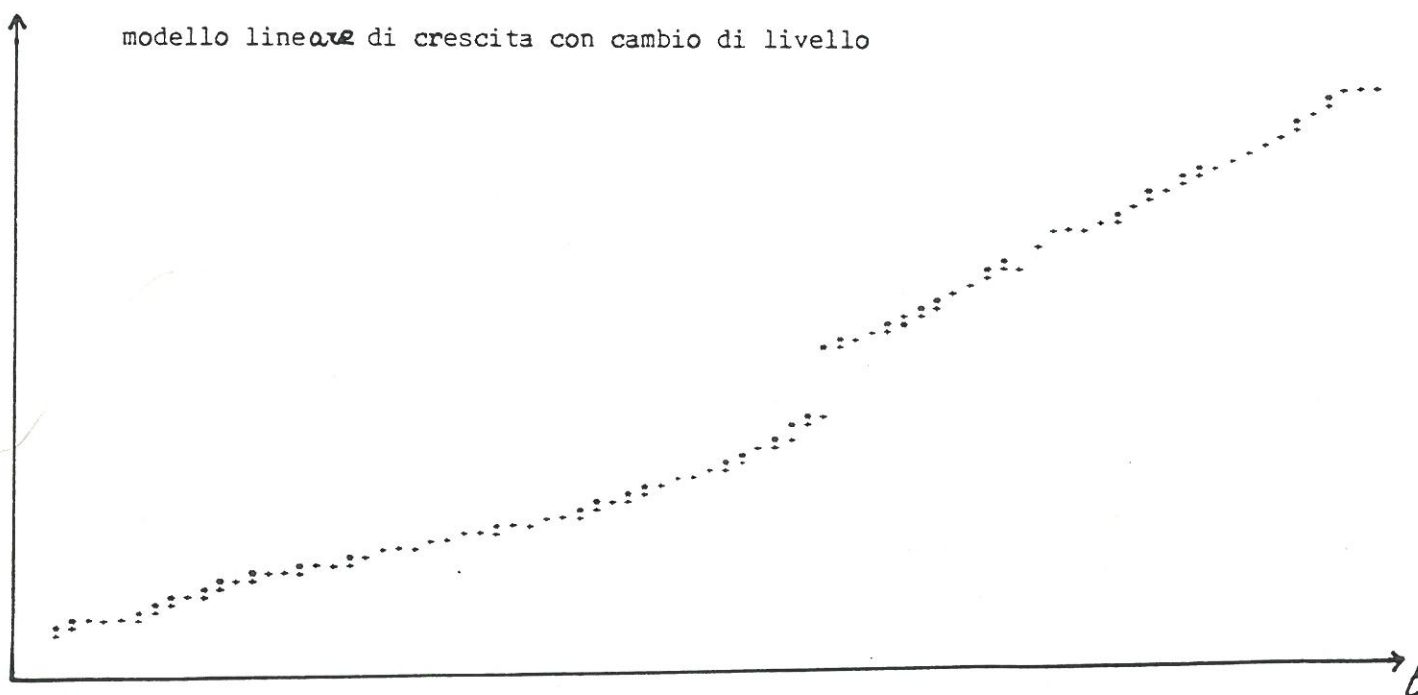


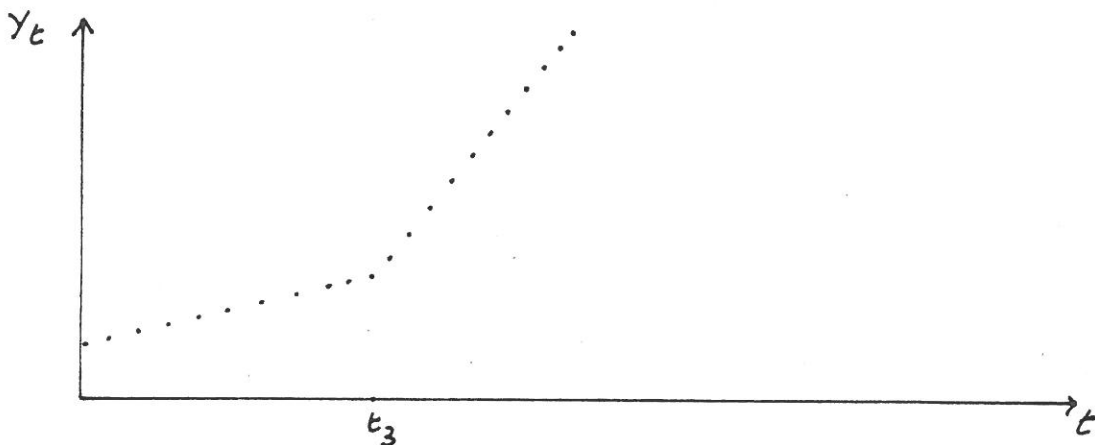
FIG.4. (*) valori osservati, (+) previsioni un passo avanti

M_4 esprime un cambiamento di pendenza all'istante t_3 e si ottiene se y_t è generata dal modello :

$$y_t = \mu_t + \delta y_t$$

$$\mu_t = \mu_{t-1} + \beta_t + \delta \mu_t$$

$$\beta_t = \beta_{t-1} + \Delta \beta_t + \delta \beta_t$$



M_4 dipende da un cambio nella crescita del sistema, pertanto da una varianza grande per β : è necessario sostituire w_2 con $w_2 + w_2^0$ dove w_2^0 è una varianza grande.

Nel nostro esempio abbiamo ipotizzato un cambio nella crescita dopo la 50.esima osservazione ed in tal punto abbiamo modificato la matrice W con

$$W = \begin{pmatrix} 6000 & 5000 \\ 5000 & 5000 \end{pmatrix}$$

I risultati sono dati nella tavola 5 e nella fig. 5:

INDEX	YAC	Y	INDEX	YAC	Y	INDEX	YAC	Y
*****	*****	*****	*****	*****	*****	*****	*****	*****
1	181.45	200	30	266.93	265.46	50	549.65	541.74
2	184.56	181.68	31	269.08	267.9	59	590.51	576.71
3	188.29	184.34	32	270.16	270.08	60	618.85	619.12
4	194.03	188.07	33	273.96	271.18	61	657.75	647.72
5	197.35	193.81	34	276.72	275	62	695.67	687.4
6	198.15	197.22	35	279.22	277.82	63	727.17	726.11
7	199.34	198.09	36	281.18	280.37	64	759.1	757.85
8	201.14	199.29	37	287.15	282.36	65	789.62	789.87
9	204.59	201.1	38	290.91	288.37	66	816.79	820.41
10	211.66	204.55	39	294.71	292.24	67	848.55	847.4
11	216.16	211.64	40	299.47	296.12	68	858.53	868.71
12	218.77	216.25	41	303.38	300.94	69	892.16	888.09
13	221.85	218.95	42	306.43	304.95	70	936.16	921.63
14	226.78	222.07	43	309.2	308.06	71	976.99	966.23
15	230.21	227.04	44	312.31	310.87	72	1004.1	1007.7
16	232.76	230.56	45	319.9	314.01	73	1039.8	1035
17	235.8	233.17	46	327.34	321.66	74	1073.9	1070.7
18	237.94	236.25	47	331.62	329.27	75	1099.8	1105
19	239.85	238.44	48	336.97	333.69	76	1125.9	1130.8
20	241.29	240.38	49	347.93	339.11	77	1156.6	1156.7
21	243.96	241.85	50	354.25	350.19	78	1185.4	1187.3
22	246.66	244.54	51	354.24	356.74	79	1223.6	1216.1
23	250.39	247.28	52	382.52	356.81	80	1271.6	1254.3
24	251.39	251.05	53	402.46	406.21	81	1311	1302.8
25	253.92	252.13	54	422.51	424.59	82	1365.4	1342.8
26	256.47	254.66	55	454.67	443.98	83	1403.8	1397.8
27	259.04	257.26	56	476.64	478.44	84	1427.7	1436.8
28	261.91	259.87	57	515.93	500.3	85		1460.7
29	264.54	262.78						

Tab. 5. YAC = valori osservati Y= previsioni un passo avanti

Osservazione 1 .

Si può dimostrare (siveda Mèlard (1980)) che il modello di crescita lineare è un processo ARIMA(0,2,2) . Nel caso di varianze costanti la dimostrazione è stata data da Harrison(1967).

Analogamente si può mostrare che da un modello polinomiale di grado n , si ottiene un modello ARIMA (0,n,n) con coefficienti dipendenti dal tempo La dimostrazione nel caso di coefficienti costanti è dovuta a Godolphin e Harrison (1975).

Osservazione 2.

Huyberegts (1979) ha dimostrato che in tale modello il filtro di Kalman si può ottenere senza far riferimento al teorema di Bayes e all'assunzione di normalità.

modello lineare di crescita con cambio di pendenza

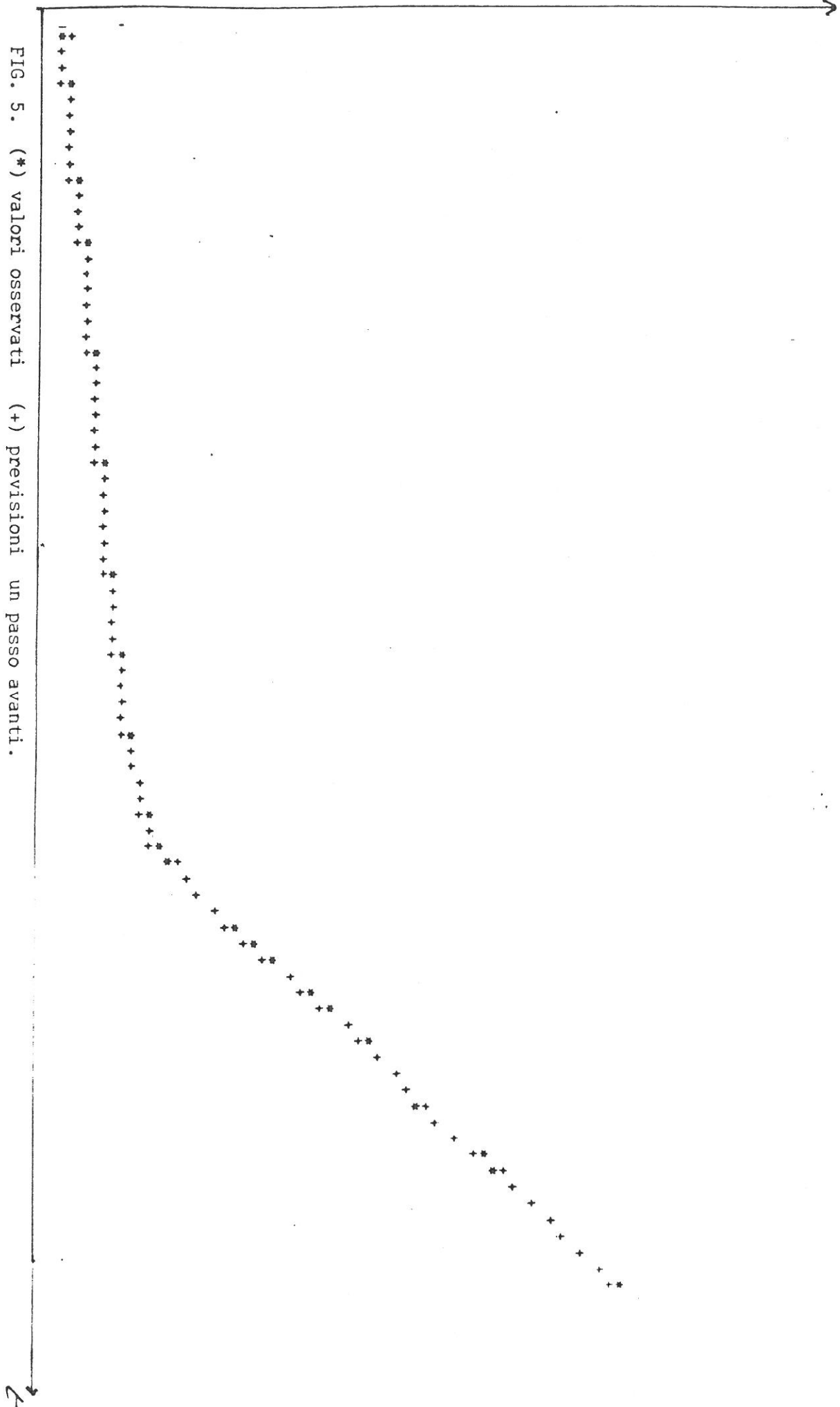


FIG. 5. (*) valori osservati (+) previsioni un passo avanti.

6.2.2. I quattro modelli in parallelo (modello multistato)

I modelli M_1, M_2, M_3, M_4 possono essere considerati in parallelo. Per comprendere tutti e quattro i casi si può supporre che ad ogni istante l'osservazione possa essere generata casualmente da uno dei quattro modelli esposti in precedenza con assegnate probabilità.

Questa situazione può essere trattata come MLD se poniamo:

$$y_t = F_t \theta_t + A_i \Delta y_t + \delta y_t$$

$$\theta_t = G_t \theta_{t-1} + B_i \Delta \theta_t + \delta \theta_t$$

Al variare di i si descrivono i quattro modelli , pertanto

$$\theta_t = (\mu_t , \beta_t)' ; \quad F_t = (1 , 0)$$

$$G_t = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\Delta \theta_t = (\Delta \mu_t , \Delta \beta_t) ; \quad A_1 = A_3 = A_4 = 0 , \quad A_2 = 1$$

$$B_1 = B_2 = 0 ; \quad B_3 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} ; \quad B_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Al tempo $t-1$ assumiamo:

1) la probabilità che il modello $M^{(i)}$ sia quello corretto è $P_{t-1}^{(i)}$.

2) condizionatamente al modello $M^{(i)}$ il sistema è descritto da:

$$(\theta_{t-1} | M_{t-1}^{(i)}, Y^{t-1}) \sim N \left(\begin{pmatrix} m_{t-1}^{(i)} \\ b_{t-1}^{(i)} \end{pmatrix}, C_{t-1}^{(i)} \right)$$

Evidentemente al tempo t può operare ognuno dei quattro modelli.

La descrizione generale del processo è data da una matrice quadrata

P_{ij} , dove p_{ij} indica la probabilità di passare dal modello $M^{(i)}$ al modello $M^{(j)}$.

Tale matrice viene detta matrice di transizione ed è data da:

$t-1 \backslash t$	$M^{(1)}$	$M^{(2)}$	$M^{(3)}$	$M^{(4)}$
$M^{(1)}$	p_{11}	p_{12}	p_{13}	p_{14}
$M^{(2)}$	p_{21}	p_{22}	p_{23}	p_{24}
$M^{(3)}$	p_{31}	p_{32}	p_{33}	p_{34}
$M^{(4)}$	p_{41}	p_{42}	p_{43}	p_{44}

PROBABILITÀ DI
TRANSIZIONE

Siamo ora in grado di calcolare:

$$P(M_{t-1}^{(i)} | M_{t-1}^{(j)} | Y^{t-1}) = P_{t-1}^{(i)} p_{ij}$$

Se applichiamo le equazioni di previsione di un passo utilizzando il filtro di Kalma-Bucy otteniamo:

$$\left(\begin{pmatrix} \mu \\ \beta \end{pmatrix}_t \mid M_{t-1}^{(i)} M_{t-1}^{(j)}, Y^{t-1} \right) \sim N \left(\begin{pmatrix} m_t^{(ij)} \\ b_t^{(ij)} \end{pmatrix}, C_t^{(ij)} \right)$$

dove

$$m_t^{(ij)} = m_{t-1}^{(i)} + b_{t-1}^{(i)} + A_t^{(ij)(1)} (y_t - \mu_{t-1}^{(i)} - b_{t-1}^{(i)})$$

$$b_t^{(ij)} = b_{t-1}^{(i)} + A_t^{(ij)(2)} (y_t - \mu_{t-1}^{(i)} - b_{t-1}^{(i)})$$

e

$$C_t^{(ij)} = R_t^{(ij)} - A_t^{(ij)} B_t^{(ij)} (A_t^{(ij)})^{-1}$$

con

$$R_t^{(ij)} = G C_{t-1}^{(i)} G' + W_t^{(j)}$$

$$B_t^{(ij)} = r_t^{(ij)(11)} + V_t^{(j)}$$

$$A_t^{(ij)} = \begin{pmatrix} r_t^{(ij)(11)} \\ r_t^{(ij)(12)} \end{pmatrix} / B_t^{(ij)}$$

La previsione di un passo dell'osservazione y_t sarà:

$$(y_t \mid M_{t-1}^{(i)} M_{t-1}^{(j)}, Y^{t-1}) \sim N(\hat{y}_t^{(ij)}, B_t^{(ij)})$$

con

$$\hat{y}_t^{(ij)} = m_{t-1}^{(i)} + b_{t-1}^{(i)}$$

Le probabilità finali di transizione possono essere calcolate mediante il teorema di Bayes. Otteniamo, per $1 \leq i, j \leq 4$

$$P_t^{(ij)} = P(M_{t-1}^{(i)} M_t^{(j)} | Y^t) = K p(y_t | M_{t-1}^{(i)} M_t^{(j)} Y^{t-1}) P_{t-1}^{(i)} p_{ij}$$

$$\text{dove } K^{-1} = \sum_{i=1}^4 \sum_{j=1}^4 p(y_t | M_{t-1}^{(i)} M_t^{(j)} Y^{t-1}) P_{t-1}^{(i)} p_{ij}$$

con la condizione che

$$\sum_{i=1}^4 \sum_{j=1}^4 P_t^{(ij)} = 1$$

La verosimiglianza sarà data da:

$$p(y_t | M_{t-1}^{(i)} M_t^{(j)}, Y^{t-1}) \propto (1/B_t^{(ij)})^{1/2} \exp(- (1/2B_t^{(ij)})(y_t - \hat{y}_t^{(ij)})^2).$$

Il procedimento ricorsivo si conclude riducendo le 16 distribuzioni finali a quattro. La procedura di riduzione è la seguente:

$$P_t^{(j)} = \sum_{i=1}^4 P_t^{(ij)} \quad j = 1, 2, 3, 4.$$

In generale $P_t^{(j)}$ indica la probabilità che, dato y^t , l'osservazione y_t provenga dal modello j -esimo. La matrice di transizione finale ci permette di calcolare tale probabilità in maniera ricorsiva.

Avremo inoltre:

$$m_t^{(j)} = (1/P_t^{(j)}) \sum_{i=1}^4 P_t^{(ij)} m_t^{(ij)}$$

$$b_t^{(j)} = (1/P_t^{(j)}) \sum_{i=1}^4 P_t^{(ij)} b_t^{(ij)}$$

Si avrà infine

$$m_t = \sum_{j=1}^4 m_t^{(j)} P_t^{(j)}$$

$$b_t = \sum_{j=1}^4 b_t^{(j)} p_t^{(j)}$$

La previsione dell'osservazione sarà:

$$\hat{y}_{t+1} = m_t + b_t.$$

La matrice di varianze e covarianze del modello J-esimo viene ottenuta nel modo seguente:

$$C_t^{(j)} = (1/p_t^{(j)}) \sum_{i=1}^4 p_t^{(ij)} \left\{ C_t^{(ij)} + \left(\begin{pmatrix} m^{(ij)} \\ b^{(ij)} \end{pmatrix} - \begin{pmatrix} m^{(j)} \\ b^{(j)} \end{pmatrix} \right) \left(\begin{pmatrix} m^{(ij)} \\ b^{(ij)} \end{pmatrix} - \begin{pmatrix} m^{(j)} \\ b^{(j)} \end{pmatrix} \right)' \right\}$$

Dalle cose dette emerge con chiarezza che la metodologia del filtro di Kalman permette di analizzare serie temporali che presentano cambiamenti repentini e che tale metodo è molto flessibile.

Osservazione 1.

In questo modello multi-stato ad ogni istante l'osservazione y_t può provenire da uno dei quattro possibili modelli. In tal caso le distribuzioni finali dei parametri risultano misture di distribuzioni normali e sono poco maneggevoli, sorge allora il problema di approssimare tali distribuzioni. Per tali metodologie si veda : Gathercole-Smith (1981), Gathercole (1982) , Smith-Makov(1977).

Osservazione 2.

Le relazioni tra tali modelli e la teoria delle catastrofi si possono trovare nei seguenti lavori: Harrison-Smith(1981), Smith - Harrison-Zeeman (1981).

Osservazione 3.

Dalle cose dette è facile vedere che la metodologia adottata permette di analizzare serie temporali che presentano cambiamenti improvvisi. L'approccio del filtro di Kalman permette di analizzare il punto di cambiamento di una serie. Si veda: Smith (1980), Smith-Makov (1980).

Osservazione 4.

Per un'applicazione del modello multi stato si veda: Smith- Wess (1983).

Vediamo ora alcuni esempi in cui abbiamo utilizzato il processo multistato.

In tale modello l'osservazione può essere generata casualmente da uno dei quattro modelli esposti. Gli esempi riguardano le importazioni e le esportazioni di beni e servizi dal 1960 al 1982 (Fonte ISCO, Quadri di contabilità nazionale: conto economico delle risorse e degli impieghi).

Esempio A) : Importazioni di beni e servizi.

Assunzioni iniziali:

La probabilità di transizione da un modello ad un altro è:

0,80	0,05	0,05	0,10
0,80	0,05	0,05	0,10
0,080	0,05	0,05	0,10
0,80	0,05	0,05	0,10

La varianza delle osservazioni $V = 20$ e $V^* = 2000$. Le varianze del sistema sono $w_1 = 3$, $w_2 = 1$, $w_1^* = 1000$, $w_2^* = 1000$.

Al tempo $t=0$ assumiamo che il modello $M^{(1)}$ sia il modello corretto

con probabilità 1 e che

$$m_0^{(1)} = 3,50, \quad b_0^{(1)} = 0 \quad \text{e} \quad C_0 = \begin{pmatrix} 4 & 1 \\ 1 & 1 \end{pmatrix}$$

I risultati sono dati nella tabella 6 nel grafico 6

YACT	Y	YACT	Y
*****	*****	*****	*****
3.611	3.5	11.897	11.462
4.114	3.5538	13.142	12.549
4.738	3.908	13.425	13.796
5.777	4.5492	12.134	14.398
5.396	5.6109	14.005	13.52
5.502	5.8139	13.893	14.323
6.287	5.9146	15.122	14.488
7.141	6.4624	17.206	15.368
7.54	7.2813	18.632	17.22
8.999	7.8841	17.645	19.018
10.431	9.1532	18.017	18.951
10.679	10.698		18.992

Tab. 6. YACT= valori osservati ; Y= previsioni un passo avanti

Modello multi-stato.

Importazione di beni e servizi (1960- 1982)

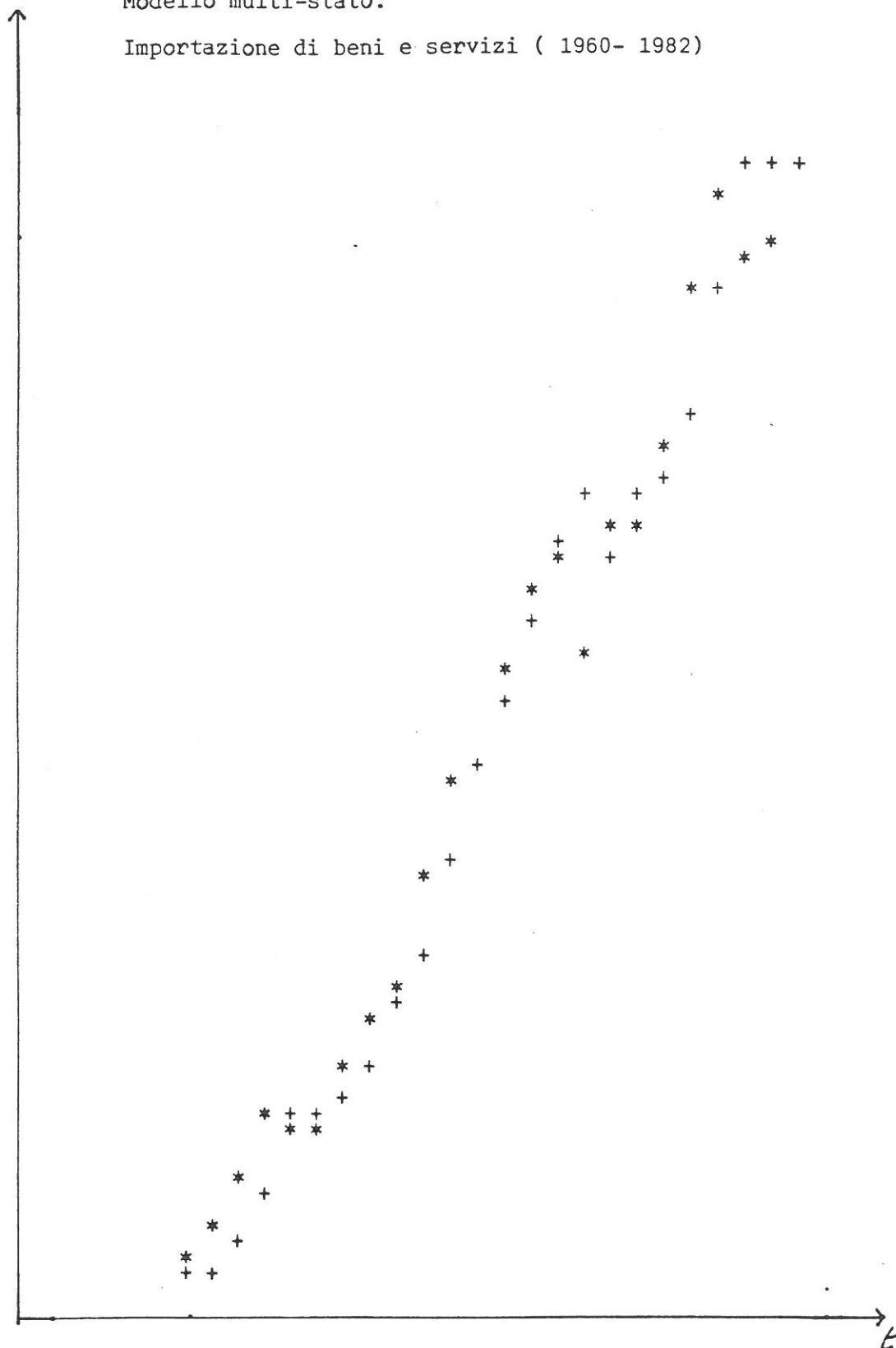


FIG. 6 . (*) valori osservati ; (+) previsioni un passo avanti

Esempio B) ESportazione di beni e servizi.

In tale esempio abbiamo mantenuto le stesse assunzioni iniziali
 tranne $m_0^{(1)} = 3,20$. I risultati sono riportati nella tabella
 e nel grafico sottostante.

YACT=VALORI STORICI, Y=PREVISIONI UN PASSO AVANTI

YACT	Y	YACT	Y
*****	*****	*****	*****
3.207	3.2	11.931	11.423
3.722	3.2034	12.382	12.508
4.14	3.5231	13.607	13.18
4.441	4.004	14.108	14.225
4.993	4.4594	15.969	14.91
6.008	5.0529	17.032	16.43
6.68	6.0331	18.75	17.747
7.259	6.9246	20.452	19.434
8.393	7.6775	19.563	21.26
9.423	8.7533	20.58	21.2
9.985	9.8886	20.829	21.663
10.7	10.686		21.837

Tab. 7

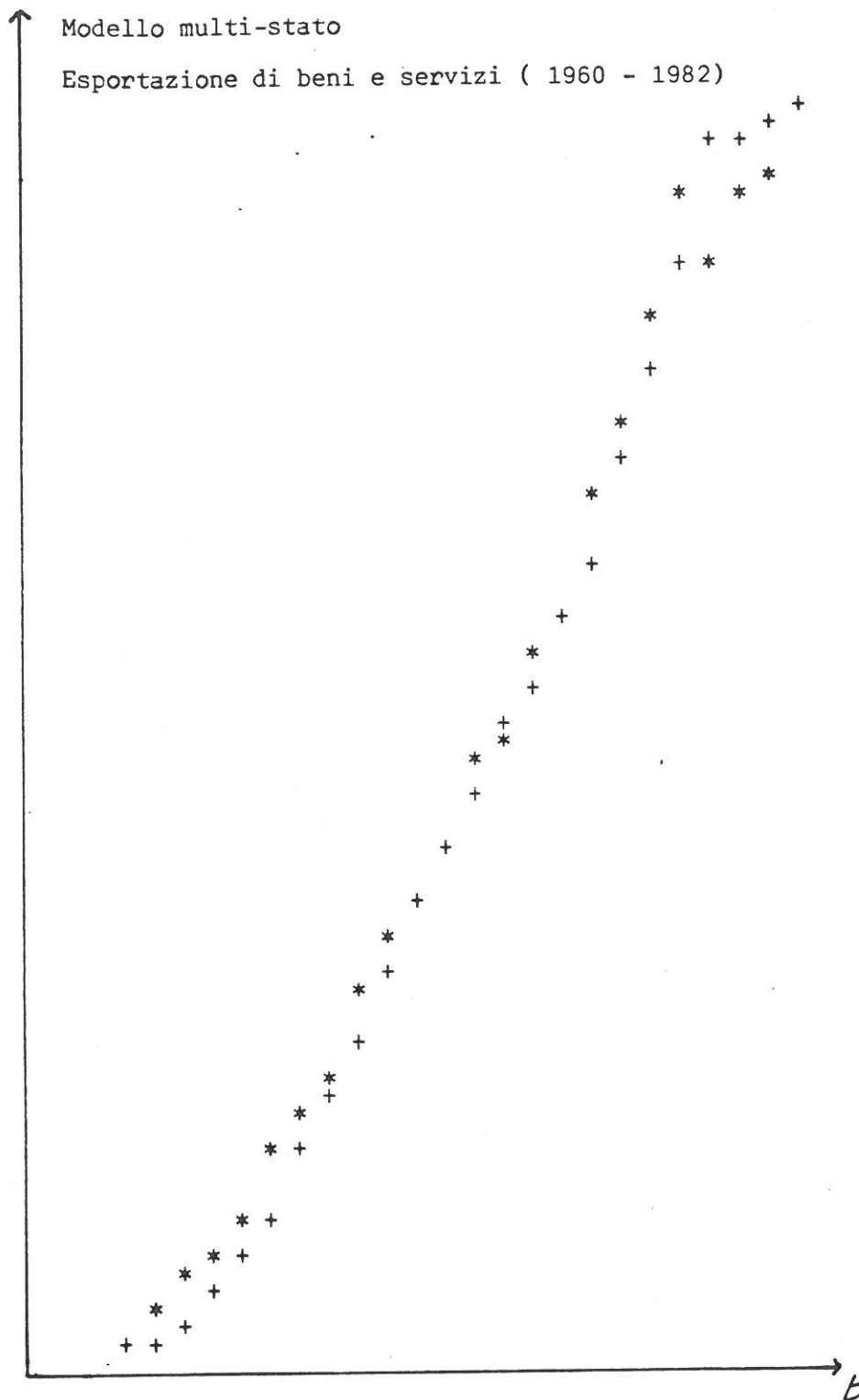


FIG. 7 . (*) valori osservati (+) previsioni un passo avanti.

Facciamo osservare che gli esempi riguardanti le importazioni e le esportazioni avremmo potuto studiarli anche con un semplice modello di crescita lineare, i risultati si sarebbero discostati di poco. Per mettere invece in luce il ruolo del modello multi-stato abbiamo costruito una nuova variabile

$$Y = (\text{Esportazioni} - \text{Importazioni}) / \text{PIL}$$

Le assunzioni iniziali sono state così modificate:

La matrice di transizione è:

0.02	0.20	0.50	0.28
0.02	0.20	0.50	0.28
0.02	0.20	0.50	0.28
0.02	0.20	0.50	0.28

Le probabilità che il modello $M^{(i)}$ sia il modello corretto al tempo $t=0$ sono:

$$P_0^{(1)} = 0.02 \quad P_0^{(2)} = 0.20 \quad P_0^{(3)} = 0.50 \quad P_0^{(4)} = 0.28$$

Le varianze delle osservazioni e le varianze del sistema sono rimaste immutate. Inoltre $m_0^{(1)} = 0,01$, $b_0^{(1)} = -0,001$

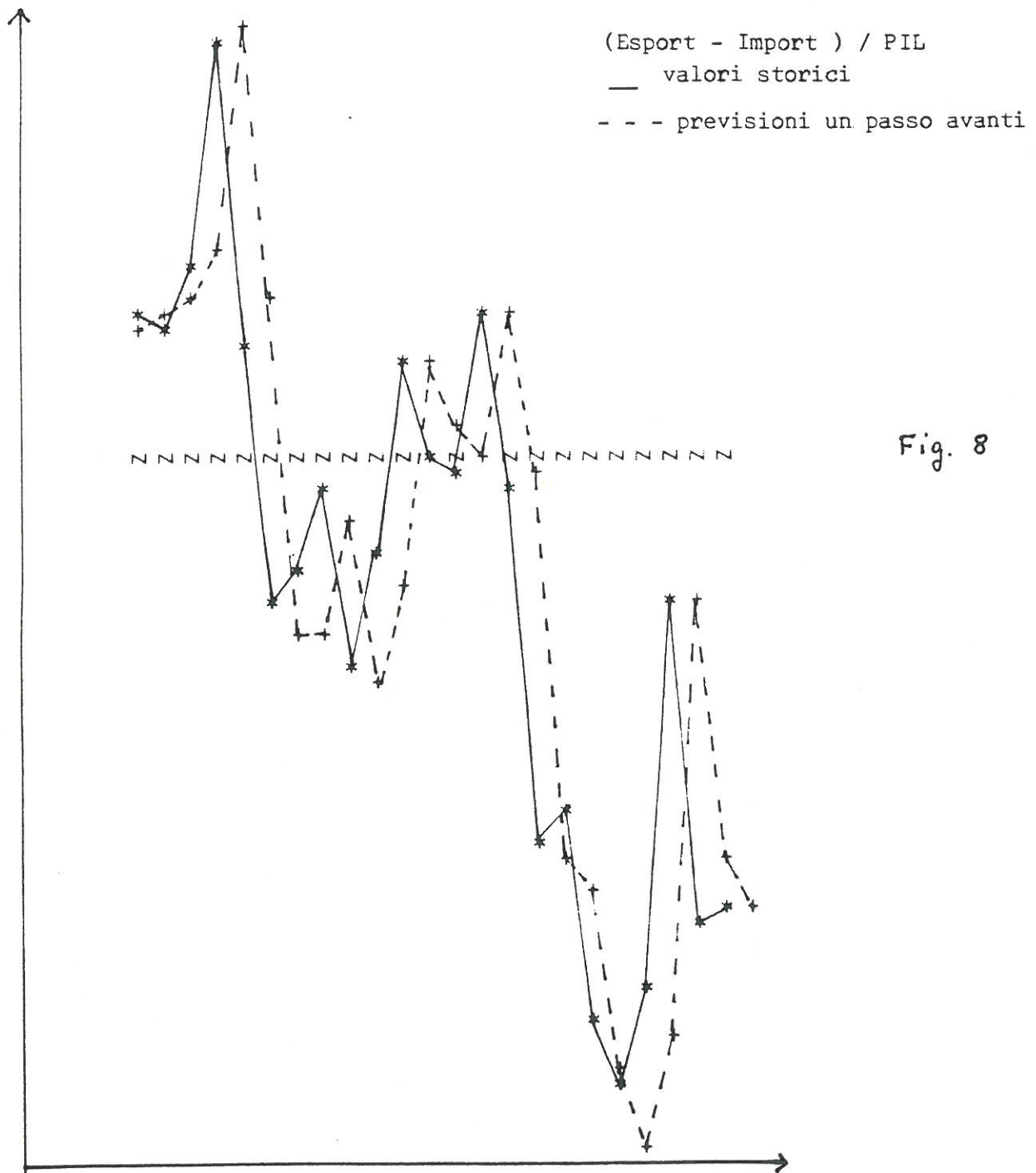
$$C_0 = \begin{pmatrix} 0,001 & 0 \\ 0 & 0,001 \end{pmatrix}$$

I risultati sono riportati nella tabella e nel grafico sottostante.

YACT=VALORI STORICI, Y=PREVISIONI UN PASSO AVANTI

YACT	Y	YACT	Y
*****	*****	*****	*****
.011193	.01	-5.1544E-4	.0019911
.010037	.010223	.010765	2.0516E-4
.014417	.011446	-.0024753	.011328
.030499	.015649	-.027861	-9.0552E-4
.0089496	.032107	-.026183	-.02946
-.010881	.011851	-.041068	-.031313
-.007974	-.012573	-.046224	-.045017
-.0022339	-.01266	-.039423	-.050803
-.015156	-.0051323	-.010882	-.04281
-.0071008	-.016297	-.034256	-.010767
.0070925	-.0092007	-.032935	-.030023
-3.2856E-4	.0072621		-.032824

Tab. 8



6.3. Il modello stagionale.

Sia T la lunghezza del ciclo (è ovvio che per dati mensili la lunghezza del ciclo sia $T=12$)

Sia $S_{i,t}$ il fattore stagionale nel mese i .esimo dell'anno t .esimo.

Naturalmente $S_{i,t}$ deve soddisfare il seguente vincolo

$$\sum_{i=1}^T S_{i,t} = 0 \quad \text{per ogni } t$$

In diciamo con $S'_t = (S_{1,t}, S_{2,t}, \dots, S_{T,t})$ i dati dell'anno t .esimo
Se l'evoluzione degli effetti stagionali è data dalla seguente equazione

$$S'_t = S'_{t-1} + \delta S'_t \quad \delta S_t \sim N(0, W)$$

allora

$$\sum_{i=1}^T S_{i,t} = \sum_{i=1}^T S_{i,t} = 0$$

Pertanto

$$\text{Var}\left(\sum_{i=1}^T \delta S_{i,t}\right) = 0$$

ossia

$$\underline{1}' W \underline{1} = 0$$

dove $\underline{1}'$ è un vettore $(1 \times T)$ formato da elementi unitari.

Evidentemente se $W = [w_{ij}] \quad i, j = 1, 2, \dots, T$ si ha

$$\sum_{i=1}^T \sum_{j=1}^T w_{ij} = 0.$$

6.4. Il modello stagionale additivo.

La forma additiva del modello stagionale è descritta da una equazione delle osservazioni e da una equazione del sistema. Le equazioni per il mese i .esimo dell'anno t .esimo sono:

$$\begin{aligned} Y_{i,t} &= F_i S_t + v_t & v_t &\sim N(0, V_t) \\ S_t &= S_{t-1} + w_t & w_t &\sim N(0, W_t) \end{aligned}$$

dove $F_i = (0, \dots, 0, 1, 0, 0, \dots, 0)$ è un vettore di dimensioni $(1 \times T)$ composto da zero eccetto nel mese i .esimo che assume il valore 1.

La restrizione

$$\sum_{i=1}^T \sum_{j=1}^T w_{ij} = 0$$

impone che

$$w_{ij,t} = -(1/T-1)w_{ii,t}$$

dove

$$W_t = [w_{ij}]_t$$

In tale modello è necessario porre molta attenzione nella scelta delle distribuzioni iniziali.

Per semplicità la distribuzione iniziale viene data da:

$$S_0 | Y_0 \sim N(m_0, C_0)$$

con $\sum_{i=1}^T m_{i,0} = 0$, $c_{ii} = \sigma^2$ e $c_{ij} = -(1/T-1)\sigma^2$ ($i \neq j$).

$C_0 = [c_{ij}]$, σ^2 è scelta opportunamente per soddisfare la condizione $\underline{1}'W\underline{1} = 0$.

6.5. Il modello stagionale di crescita lineare.

Tale modello è una combinazione lineare tra il modello di crescita lineare e quello stagionale additivo. Spesso il modello si presenta in forma moltiplicativa, è ovvio che una semplice trasformazione logaritmica dei dati reali porta ad un modello additivo.

Sia

T = numero di dati per anno

t = indice del tempo ($t = 1, 2, \dots$)

$m(t)$ = mese associato al tempo t

Y_t = valore dell'osservazione al tempo t .

Supponiamo che il processo sia moltiplicativo con i seguenti parametri al tempo t

M_t = livello

B_t = crescita

$q_{i,t}$ = fattore stagionale per il mese i ($i = 1, \dots, T$)

Il modello è:

$$Y_t = M_t q_{m(t),t} \varepsilon_t$$

$$M_t = M_{t-1} B_t \delta M_t$$

$$B_t = B_{t-1} \delta B_t$$

$$q_{i,t} = q_{i,t-1} \delta q_{i,t} \quad i = 1, \dots, T$$

dove δM_t , δB_t , $\delta q_{i,t}$ sono disturbi lognormali nel livello, nella crescita e nella stagionalità ed ε_t è l'errore di osservazione anch'esso lognormale. Il modello presentato è non lineare nei $(T+2)$ parametri ma può essere facilmente trasformato in un modello lineare prendendo i logaritmi.

Ponendo $y_t = \log Y_t$ avremo:

$$y_t = u_t + S_{m(t),t} + v_t$$

$$u_t = u_{t-1} + \beta_t + \delta u_t$$

$$\beta_t = \beta_{t-1} + \delta \beta_t$$

$$S_{i,t} = S_{i,t-1} + \delta S_{i,t} \quad (i = 1, 2, \dots, T)$$

con il vincolo

$$\sum_{t=1}^T S_{i,t} = \sum_{t=1}^T \delta S_{i,t} = 0$$

In generale vi sono $T+2$ paramentri non noti ma dato il vincolo sui fattori stagionali vi sono solamente $T+1$ parametri indipendenti. Tale modello può essere scritto nella forma generale del MLD presentata nelle equazioni (3.1) e (3.2) come segue:

$$y_t = F_t \theta_t + v_t \quad v_t \sim N(0, V_t)$$

$$\theta_t = G_t \theta_{t-1} + w_t \quad w_t \sim N(0, W_t)$$

con $m=1$, $n = T+2$, $\theta_t = (u_t, \beta_t, S_{1,t}, S_{2,t}, \dots, S_{T,t})$ e

$F_t = (1, 0, | 0, \dots, 0, 1, 0, \dots, 0)$. La matrice G può essere partizionata nel modo seguente:

$$G = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix}$$

con

$$G_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$G_2 = I \text{ (matrice identità)}$$

Osserviamo inoltre che la matrice W (matrice di varianze e covarianze degli errori del sistema) è data da:

$$W_t = \begin{pmatrix} W_1 & 0 \\ 0 & W_2 \end{pmatrix}$$

dove

$$W_1 = \begin{pmatrix} w_1 + w_2 & w_2 \\ w_2 & w_2 \end{pmatrix}$$

e

$$W_2 = w_3 \begin{pmatrix} 1 & -(1/T-1), \dots, & -(1/T-1) \\ -(1/T-1) & 1 & , \dots, & -(1/T-1) \\ . & . & & . \\ : & & & : \\ -(1/T-1) & , \dots, & & 1 \end{pmatrix}$$

Evidentemente w_1, w_2, w_3 saranno scelti in modo opportuno.

Osservazione 1.

E' appena il caso di notare che anche in questo modello si può pensare che ad ogni istante il processo possa essere uno dei tanti possibili (es. cambio di livello, cambio di crescita, outlier, nessun cambiamento).

Si può assumere pertanto , come abbiamo fatto in precedenza al punto 6.2.2 che questi quattro modelli possano accadere casualmente con determinate probabilità ed indipendentemente dallo stato passato.

Per ulteriori approfondimenti si veda Johnston- Harrison(1980), Harrison -Stevens(1975).

Vediamo ora due esempi.

Come primo esempio per il modello stagionale dicrescita lineare abbiamo considerato le importazioni di merci (quantità) per gli anni 1971- 1972- 1973- 1974 , dati mensili.) Fonte: nostra elaborazione su dati Istat.

Assunzioni iniziali.

Varianza delle osservazioni : $V = 5$

Varianze del sistema : $w_1 = 4$, $w_2 = 1$, $w_3 = 22$.

Coefficienti di stagionalità: 0,952 ; 0,965; 1,018, 0,962; 1,039;1,028;
0,835;1,029; 1,016; 0,974;1,135.

$$m_0 = 755 \quad ; \quad b_0 = 0.$$

C_0 è una matrice diagonale che ha sulla diagonale principale il valore 0,01.

I risultati sono riportati nella tabella 9 e nella figura 9.

IMPQUA=VALORI STORICI,YY=PREVISIONI UN PASSO AVANTI

IMPQUA	YY	IMPQUA	YY	IMPQUA	YY	IMPQUA	YY
*****	*****	*****	*****	*****	*****	*****	*****
754	755.95	783.6	823.24	959	902.57	1007.7	1064
778.5	755.77	900.6	909.58	755.9	1040.5	1119.7	843.07
883.7	759.44	831.6	752.82	877.4	874.31	1100.3	1050.4
718.9	788.03	912	934.28	1096.7	947.43	970.1	1289.4
879.3	777.88	832.2	834.55	1131.1	910.4	935.2	1217.1
777	812.23	864.2	805.31	969.5	1016.5	1242.8	953.24
742.4	811.57	752.6	763.12	856.5	897.7	806.9	922.47
673.9	797.39	841.3	827.38	1115.1	977.3	991.8	1133.9
738.2	762.04	894.8	887.43	974	1082.4	1047.8	938.74
791.1	751.78	886.1	873.96	900.8	1045.3	856.4	891.55
771.5	760.24	1058.6	983.07	1224.2	1170.2	1049.9	1192.9
875.5	761.99	892.1	965.53	956.6	1022.1		869.73
827.9	790.48						

TAB. 9

Modello stagionale di crescita lineare
 Importazioni di merci (1971-1974) dati mensili.

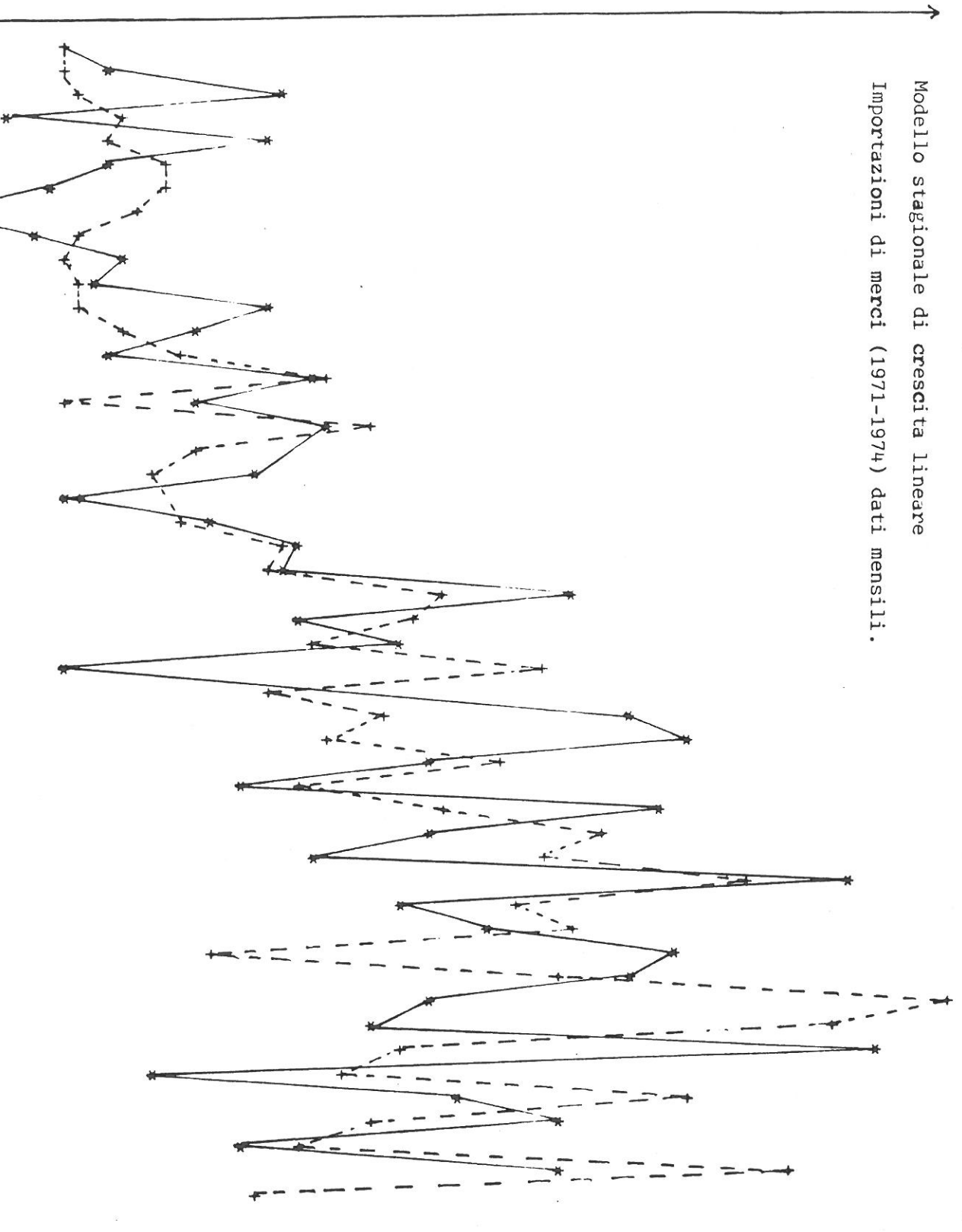


FIG.9. (*) valori storici. (+) previsioni un passo avanti.

Nel secondo esempio abbiamo considerato l'indice generale della produzione industriale.

Gli indici sono depurati dalle variazioni di calendario.

Gli anni considerati sono: 1972, 1973, 1974. Base 1970 = 100.

Fonte Istat.

Assunzioni iniziali.

Coefficienti di stagionalità: 0,952, 0,965, 1,018, 0,962, 1,039, 1,028, 0,835, 1,029, 1,016, 0,974.

$$m_0 = 100, \quad b_0 = 0.$$

Varianza delle osservazioni : $V = 5$

Varianze del sistema: $w_1 = 100$, $w_2 = 80$, $w_3 = 200$.

C_0 è una matrice identica a quella precedente.

I risultati sono riportati nella tabella 10 e nella figura 10.

IPI72=VALORI STORICI,YY=PREVISIONI UN PASSO AVANTI

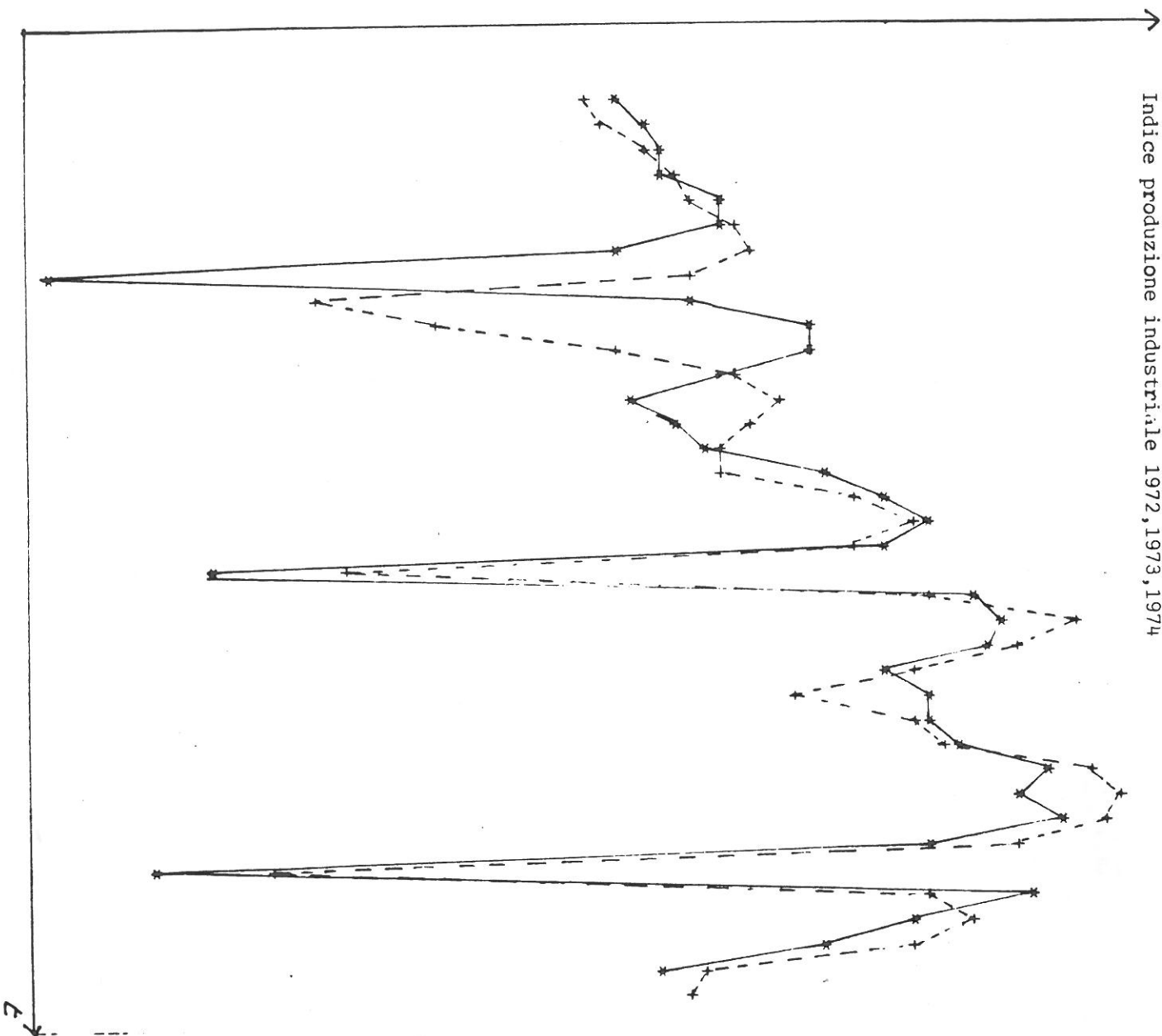
IPI72	YY	IPI72	YY	IPI72	YY	IPI72	YY
*****	*****	*****	*****	*****	*****	*****	*****
102.6	101	114.8	102.41	77.3	85.935	128.6	134.97
104.6	101.87	109.8	110.07	125.3	122.3	130.9	134.02
105.2	104.15	103.9	112.94	126.9	131.67	122.4	128.77
105.7	106.03	106.7	110.95	126.8	128.38	72.7	81.103
109.4	107.19	107.9	109.41	119.4	121.92	129.7	122.23
109.6	109.94	116.4	109.34	122.9	113.91	121.9	125.71
102.2	111.37	120	117.94	122.6	121.22	116	121.55
66.5	107.04	122.9	121.64	124.8	123.35	105.8	108.5
107	83.49	120.2	117.75	129.8	132.87		107.41
115.1	91.334						

Tab. 10

Modello stagionale di crescita lineare
Indice produzione industriale 1972, 1973, 1974

— valori storici
--- previsioni un
passo avanti

FIG.10



6.6. Il modello lineare generale statico.

Si consideri il modello lineare

$$Y = X \beta + \varepsilon$$

con Y vettore ($m \times 1$) aleatorio di variabili osservabili

β vettore ($n \times 1$) di parametri aleatori

X matrice ($m \times n$) non stocastica

ε vettore ($m \times 1$) di errori aleatori non osservabili.

Tale modello si può porre sotto forma di MLD ponendo:

$$F_t = X, \quad \theta_t = \theta_{t-1} = \dots = \beta, \quad G_t = I$$

Assumiamo ora che ε si distribuisca normalmente con

$$E(\varepsilon) = 0 \quad \text{e} \quad E(\varepsilon \varepsilon') = \sigma^2 I_m$$

Quindi

$$Y | X, \beta, \sigma^2 \sim N(X \beta, \sigma^2 I_n)$$

Assumiamo inoltre che σ^2 sia nota e che la distribuzione iniziale per β coniugata al modello sia una normale multivariata con media μ e matrice di varianze e covarianze C_β

E' immediato verificare che la distribuzione finale sarà ancora una normale con valore atteso

$$E(\beta | Y) = ((X'X) \sigma^{-2} + C_\beta^{-1})^{-1} (\sigma^{-2} X'Y + C_\beta^{-1} \mu)$$

e matrice di varianze e covarianze

$$((X'X) \sigma^{-2} + C_\beta^{-1})^{-1}$$

Utilizzando il lemma matriciale di inversione ricordato a pag. 10, semplici calcoli ci portano ai risultati:

$$E(\beta | Y) = \mu + C_\beta X' (\sigma^2 I_n + X C_\beta X')^{-1} (Y - X \mu)$$

$$\text{Cov}(\beta | Y) = C_\beta - C_\beta X' (\sigma^2 I_n + X C_\beta X')^{-1} X C_\beta$$

Tali espressioni coincidono con quelle del filtro di Kalman-Bucy.

6.7. Il modello di regressione dinamica.

Un modello di regressione con parametri che variano nel tempo, con $y(t)$ variabile dipendente e $x_1(t), \dots, x_n(t)$ variabili indipendenti può essere scritto:

$$y(t) = \sum_{i=1}^m a_i(t) x_i(t) + v(t)$$

Se si assume che i coefficienti di regressione variano nel tempo secondo un processo markoviano

$$a_i(t) = \alpha_i a_i(t-1) + \zeta_i(t) \quad i = 1, 2, \dots, m$$

allora è agevole esprimere tale modello sotto forma di modello lineare dinamico.

L'espressione generale, come più volte ricordato è la seguente

$$\begin{aligned} y_t &= F_t \theta_t + v_t \\ \theta_t &= G_t \theta_{t-1} + w_t \end{aligned}$$

Nel nostro caso, basta porre:

$$\theta_t' = (a_1(t), \dots, a_m(t))$$

$$F_t = (x_1(t), \dots, x_m(t))$$

$$G_t = \begin{pmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_m \end{pmatrix}$$

$$w_t = \begin{bmatrix} \zeta_1(t) \\ \zeta_2(t) \\ \vdots \\ \zeta_m(t) \end{bmatrix}$$

Una ulteriore generalizzazione può essere la seguente. Se assumiamo che le variabili indipendenti $x_i(t)$ possano essere osservate come

$$b_i(t) = x_i(t) + e_i(t) \quad i = 1, 2, \dots, m$$

e supponiamo inoltre che le $x_i(t)$ soddisfino

$$x_i(t) = \phi_i x_i(t-1) + \gamma_i(t)$$

allora il modello può essere scritto sotto forma di modello lineare se i coefficienti sono costanti mentre nel caso di coefficienti variabili il modello è non lineare (si veda Mehra (1976), (1979)).

Per meglio esemplificare il modello di regressione con coefficienti variabili consideriamo un caso molto semplice.

Sia

$$\begin{aligned} y_t &= x'_t \theta_t + \varepsilon_t \\ \theta_t &= \theta_{t-1} + w_t \end{aligned}$$

Assumiamo che y_t sia uno scalare e riparametrizziamo la matrice di varianze e covarianze W_t .

$$W_t = \text{Cov}(w_t) = \sigma^2 A$$

dove $\sigma^2 = \text{Var}(\varepsilon_t)$

Allora si deduce che $\text{Var}(w_t) / \text{Var}(\varepsilon_t) = A$.

Possiamo ora applicare le equazioni del filtro di Kalman per stimare θ_t . La previsione di un passo di θ_t , note le osservazioni fino al tempo $t-1$ sarà:

$$E(\theta_t | y^{t-1}) = m_{t-1}$$

e la matrice di varianze e covarianze

$$\text{Cov}(\theta_t | y^{t-1}) = C_{t-1} + A = R_t$$

La stima di θ_t note le osservazioni fino al tempo t sarà :

$$E(\theta_t | y^t) = m_t$$

con

$$\begin{aligned} m_t &= m_{t-1} + (x'_t R_t x_t + 1)^{-1} R_t x_t (y_t - x'_t m_{t-1}) \\ &= m_{t-1} + (1 + x'_t (C_{t-1} + A) x_t)^{-1} (C_{t-1} + A) x_t (y_t - x'_t m_{t-1}) \end{aligned}$$

e

$$C_t = (C_{t-1} + A) - (1 + x_t'(C_{t-1} + A)x_t)^{-1}(C_{t-1} + A)x_t x_t'(C_{t-1} + A)$$

Se si pongono i coefficienti costanti ($A = 0$) si ha:

$$m_t = m_{t-1} + (1 + x_t' C_{t-1} x_t)^{-1} C_{t-1} x_t (y_t - x_t' m_{t-1})$$

$$C_t = C_{t-1} - (1 + x_t' C_{t-1} x_t)^{-1} C_{t-1} x_t x_t' C_{t-1}$$

Queste equazioni sono le stesse espressioni ottenute da Plackett(1950).

Consideriamo il caso di un solo coefficiente

$$y_t = b_t x_t + \varepsilon_t$$

$$b_t = b_{t-1} + w_t$$

con $\text{Var}(w_t) = \sigma^2 w$ dove w è il rapporto $\text{Var}(w_t)/\text{Var}(\varepsilon_t)$.

Le equazioni delle stime diventano:

$$m_t = m_{t-1} + C_t x_t (y_t - x_t' m_{t-1})$$

con

$$\begin{aligned} C_t &= (C_{t-1} + w) - (x_t^2 (C_{t-1} + w)^2) / (1 + x_t^2 (C_{t-1} + w)) = \\ &= (C_{t-1} + w) / (1 + x_t^2 (C_{t-1} + w)). \end{aligned}$$

Se i coefficienti sono costanti $w=0$ si ha :

$$C_t = C_{t-1} / (1 + x_t^2 C_{t-1})$$

$$m_t = m_{t-1} + C_t x_t (y_t - x_t' m_{t-1})$$

E' agevole verificare che

$$C_t = (1 + C_1 \sum_{j=2}^{t+1} x_j^2)^{-1} C_1$$

Ponendo $C_1 = 1/x^2$ si ottiene

$$c_t = 1 / \left(\sum_{j=1}^{t+1} x_j^2 \right)$$

e

$$m_t = m_{t-1} + \left(x_t / \left(\sum_{j=1}^{t+1} x_j^2 \right) \right) (y_t - x_t m_{t-1}).$$

Per poter calcolare le stime e le previsioni è necessario, come si è visto, specificare la matrice A. Molte sono le impostazioni utilizzate in letteratura (si veda Harrison - Stevens (1976) e Mehra(1979) per una rassegna).

6.8. Il modello ARMA.

Un processo ARMA(p,q) univariato ammette, come è noto, la seguente rappresentazione:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

$$t = 1, 2, \dots, n; \varepsilon_t \sim N(0, \sigma^2).$$

Tale modello può essere trasformato in modello lineare dinamico (si veda Akaike (1974)) definendo un vettore α_t di dimensioni $(k \times 1)$ che soddisfa la seguente equazione del sistema

$$\alpha_t = H \alpha_{t-1} + R w_t$$

$t = 1, 2, \dots, n$ e $k = \max(p, q+1)$ con

$$H = \begin{bmatrix} \phi_1 & : & & & \\ \phi_2 & : & & & \\ \vdots & : & & & \\ \vdots & : & I_{k-1} & & \\ \phi_k & : & \dots & \dots & \dots \\ & & 0'_{k-1} & & \end{bmatrix}$$

$$R = \begin{bmatrix} 1 \\ \theta_1 \\ \vdots \\ \vdots \\ \theta_{k-1} \end{bmatrix}$$

Facciamo osservare che $\phi_j = 0$ per $j > p$ e $\theta_j = 0$ per $j > q$.

I_{k-1} è una matrice identità di dimensioni $(k \times 1)$ $(k \times 1)$.

L'equazione delle osservazioni è:

$$y_t = (1, 0, \dots, 0) \alpha_t$$

7. UN METODO ALTERNATIVO PER DETERMINARE IL FILTRO DI KALMAN-BUCY.

In questo numero presentiamo un metodo alternativo per determinare il filtro di kalman-Bucy. Tale procedimento si basa su un noto risultato di analisi statistica multivariata e su alcune proprietà della distribuzione normale.

TEOREMA(correlazione normale, si veda Anderson (1958) pag. 28-29)

Siano X_1 e X_2 due vettori aventi distribuzione normale con

$$m_1 = E(X_1) \quad m_2 = E(X_2)$$

$$D_{11} = \text{Cov}(X_1 X_1) \quad D_{12} = \text{Cov}(X_1 X_2) \quad D_{22} = \text{Cov}(X_2 X_2)$$

Allora la distribuzione di X_1 dato X_2 è ancora normale con:

$$E(X_1 | X_2) = m_1 + D_{12} D_{22}^{-1} (X_2 - m_2) \quad (7.1)$$

$$\text{Cov}(X_1 X_1 | X_2) = D_{11} - D_{12} D_{22}^{-1} D_{21} \quad (7.2)$$

Corollario.

Se $X_1 | X_2$ si distribuisce normalmente con valore atteso dato dalla (7.1) e matrice di varianze e covarianze data dalla (7.2), quando $X_2 \sim N(m_2, D_{22})$ allora

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N \left[\begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix} \right] \quad (7.3)$$

La (7.1) è, come è noto, la funzione di regressione e $(D_{12} D_{22}^{-1})$ è il coefficiente di regressione dei minimi quadrati.

Sia

$$X_1 \Leftrightarrow e_t$$

$$X_2 \Leftrightarrow \theta_t$$

Dalla (4.7) si deduce che

$$(\theta_t | y^{t-1}) \sim N(G_t m_{t-1}, R_t)$$

E' ovvio allora stabilire le relazioni:

$$m_2 \Leftrightarrow G_t m_{t-1}$$

$$D_{22} \Leftrightarrow R_t$$

Se nella (7.1) sostituiamo X_1 , X_2 , m_2 , D_{22} rispettivamente con e_t

θ_t , $G_t m_{t-1}$, R_t allora, ricordando il risultato dell'equazione (4.17)

$$(e_t | y^{t-1}) \sim N(F_t(\theta_t - G_t m_{t-1}), V_t)$$

si ha :

$$m_1 + D_{12} D_{22}^{-1} (X_2 - m_2) = m_1 + D_{12} R_t^{-1} (\theta_t - G_t m_{t-1})$$

da cui

$$m_1 + D_{12} R_t^{-1} (\theta_t - G_t m_{t-1}) \Leftrightarrow F_t(\theta_t - G_t m_{t-1})$$

Pertanto

$$m_1 \Leftrightarrow 0$$

$$D_{12} \Leftrightarrow F_t R_t$$

Operariamo in maniera analoga nell'espressione (7.2)

$$D_{11} - D_{12} D_{22}^{-1} D_{21} = D_{11} - F_t' R F_t \Leftrightarrow V_t$$

da cui

$$D_{11} \Leftrightarrow V_t + F_t' R F_t$$

Il corollario al teorema della correlazione normale ci permette di scrivere:

$$\begin{pmatrix} \theta_t \\ e_t \end{pmatrix} \Big| y^{t-1} \sim N \left[\begin{pmatrix} G_t m_{t-1} \\ 0 \end{pmatrix}, \begin{pmatrix} R_t & R_t F_t' \\ F_t R_t & V_t + F_t' R F_t \end{pmatrix} \right]$$

Applicando ora il teorema si ottiene:

$$(\theta_t | e_t, y^{t-1}) \sim N(m_t, C_t)$$

dove

$$m_t = G_t m_{t-1} + R_t F_t' (V_t + F_t' R F_t)^{-1} e_t \quad (7.4)$$

$$C_t = R_t - R_t F_t' (V_t + F_t' R F_t)^{-1} F_t R_t \quad (7.5)$$

Come si può vedere la (7.4) e la (7.5) coincidono con le espressioni (4.11) e (4.12)

Tale metodo ci permette di cogliere meglio alcuni aspetti del filtro di Kalman-Bucy. Si può osservare infatti che $E(\theta_t | e_t, y^{t-1})$ è la funzione di regressione di θ_t su e_t ed è data dalla somma di due quantità, la media della distribuzione iniziale ed un multiplo dell'errore di previsione. Il multiplo dell'errore di previsione (matrice guadagno) è il coefficiente di regressione dei minimi quadrati di θ_t su e_t . Il filtro

pertanto può essere visto come l'evoluzione di una serie di funzioni di regressione di θ_t su e_t al tempo $0, 1, \dots, t-1, t$, ognuna con una diversa intercetta ed un diverso coefficiente di regressione.

RIFERIMENTI BIBLIOGRAFICI

Anderson-Moore(1979) , Kailath (1981) , Meinhold-Singpurwalla(1983)

8. UNA GENERALIZZAZIONE DEL MODELLO LINEARE DINAMICO.

8.1. Definizione del modello ed assunzioni di base.

Nell'usuale teoria del filtro di Kalman-Bucy la matrice F_t è assunta dipendente dal tempo ma indipendente dalle osservazioni y_t ($t = 1, 2, \dots$). Se l'assunzione di indipendenza dalle osservazioni viene rimossa è necessario porre molta attenzione all'utilizzo dei risultati forniti nei paragrafi precedenti.

Un esempio classico di dipendenza della matrice F_t dalle osservazioni si trova nel modello autoregressivo.

E' noto che il modello autoregressivo è dato dalla seguente espressione:

$$Y_t = \sum_{i=1}^p \phi Y_{t-i} + \varepsilon_t$$

Se si vuole porre il modello autoregressivo sotto forma di modello lineare dinamico si ha:

$$\begin{aligned} Y_t &= (Y_{t-1}, \dots, Y_{t-p}) \phi_t + \delta y_t \\ \phi_t &= \phi_{t-1} + \delta \phi_t \end{aligned}$$

Come si vede F_t dipende dalle osservazioni passate.

Altra assunzione di base che viene avanzata nell'usuale modello lineare dinamico presentato nel paragrafo 3 è quella che impone v_t e w_t stocasticamente indipendenti.

Se rimuoviamo queste due assunzioni di base è evidentemente necessario introdurre un modello che generalizzi quello da noi presentato in precedenza. Tale generalizzazione è stata introdotta in letteratura da Liptser-Shiryayev (1972).

Sia dato un processo stocastico parzialmente osservabile $(\theta, Y) = (\theta_t, Y_t)$,
 $t = 0, 1, 2, \dots$ definito dalle seguenti equazioni:

$$\theta_{t+1} = a_0(t, y) + a_1(t, y) \theta_t + b(t, y) w_1(t+1) \quad (8.1)$$

$$Y_{t+1} = A_0(t, y) + A_1(t, y) \theta_t + B(t, y) w_2(t+1) \quad (8.2)$$

dove w_1 e w_2 sono vettori aleatori mutuamente dipendenti e distribuiti normalmente; mentre $a_i(t, y)$ ed $A_i(t, y)$ e le matrici $b(t, y)$ e $B(t, y)$ dipendono in ogni momento solamente dai valori $(Y_0, Y_1, \dots, Y_t)$.
 La necessità di considerare tali coefficienti dipendenti dalla " storia passata " è di estrema importanza in molti problemi di controllo e di teoria dell'informazione .

Anche in questo modello il problema che si pone è quello di stimare lo stato non osservabile θ_t basandosi su $(Y_0, Y_1, \dots, Y_t) = Y^t$.
 Abbiamo visto nei paragrafi precedenti che tale problema è detto " problema di filtraggio " e che lo stimatore bayesiano di θ_t , con una funzione di danno quadratica, è dato da:

$$E(\theta_t | Y^t) = m_t$$

Abbiamo dimostrato nel paragrafo 4 che tale problema è un semplice problema di stima bayesiana di un parametro aleatorio mediante le osservazioni Y^t . L'unica difficoltà che abbiamo incontrato è dovuta al fatto che lo schema considerato è più complesso visto che abbiamo dovuto considerare il processo aleatorio θ_t piuttosto che la variabile casuale .
 Abbiamo inoltre constatato che le formule che definiscono il filtro m_t sono ricorsive e facili da un punto di vista computazionale.

Anche nel modello definito dalle espressioni (8.1) , (8.2) che generalizza il modello dato in (3.1) e (3.2) l'obiettivo che ci poniamo è quello di ottenere espressioni che siano ricorsive.

Nel caso discreto $t = 0, 1, 2, \dots$ la forma più semplice di relazione ricorrente che si possa immaginare è la seguente

$$m_{t+1} = m_t + a(t, m_t) + b(t, m_t)(y_{t+1} - y_t)$$

E' abbastanza evidente che senza speciali assunzioni sulla struttura del processo (θ, Y) è difficile pensare che m_{t+1} soddisfaccia ad equazioni ricorsive del tipo di quella enunciata. Per ottenere equazioni ricorsive sul valore atteso e sulla matrice di varianze e covarianze della distribuzione finale, l'assunzione di base è la seguente:

$$P(\theta_0 \leq x \mid Y_0) \sim N(m_0, C_0)$$

Tale assunzione permette di dimostrare che la successione (θ_t, Y_t) che soddisfa le equazioni (8.1) e (8.2) è condizionatamente gaussiana e che pertanto vale anche la seguente:

$$P(\theta_t \leq x \mid Y^t) \sim N(m_t, C_t)$$

Vediamo ora come derivare le espressioni di m_t e C_t .

8.2. EQUAZIONI RICORSIVE PER IL FILTRO.

Le equazioni (8.1) e (8.2) che descrivono l'evoluzione del processo possono essere riscritte, in maniera equivalente, nel modo seguente:

$$\theta_{t+1} = a_0(t, y) + a_1(t, y) \theta_t + b_1(t, y) \epsilon_1(t+1) + b_2(t, y) \epsilon_2(t+1) \quad (8.4)$$

$$Y_{t+1} = A_0(t, y) + A_1(t, y) \theta_t + B_1(t, y) \epsilon_1(t+1) + B_2(t, y) \epsilon_2(t+1) \quad (8.5)$$

dove

$$\epsilon_1(t) = (\epsilon_{1,1}(t), \dots, \epsilon_{1,K}(t))$$

$$\epsilon_2(t) = (\epsilon_{2,1}(t), \dots, \epsilon_{2,K}(t))$$

sono variabili aleatorie indipendenti e distribuite normalmente con

$$E(\epsilon_{i,j}(t)) = 0$$

$$E(\epsilon_{i_1,j_1}(t) \epsilon_{i_2,j_2}(s)) = \delta(i_1 i_2) \delta(j_1 j_2) \delta(t,s)$$

dove $\delta(x,y)$ è il simbolo di Kroneker:

$$\delta(x,y) = \begin{cases} 1 & \text{se } x=y \\ 0 & \text{se } x \neq y \end{cases}$$

I vettori

$$a_0(t,y) = (a_{0,1}(t,y), \dots, a_{0,k}(t,y))$$

$$A_0(t,y) = (A_{0,1}(t,y), \dots, A_{0,k}(t,y))$$

e le matrici

$$b_1(t,y) = ||b_{ij}^{(1)}(t,y)|| \quad ; \quad b_2(t,y) = ||b_{ij}^{(2)}(t,y)||$$

$$B_1(t,y) = ||B_{ij}^{(1)}(t,y)|| \quad ; \quad B_2(t,y) = ||B_{ij}^{(2)}(t,y)||$$

$$a_1(t,y) = ||a_{ij}^{(1)}(t,y)|| \quad ; \quad A_1(t,y) = ||A_{ij}^{(1)}(t,y)||$$

hanno dimensione rispettivamente

$$(k \times k) ; (k \times 1) ; (1 \times k) ; (1 \times 1) ; (k \times k) ; (1 \times k)$$

Per ogni t gli elementi dei vettori e delle matrici dipendono solamente dai valori delle osservazioni fino a t .

Il sistema dato dalle equazioni (8.4) e (8.5) può essere risolto imponendo la seguente condizione:

$$(\theta_0, Y_0) \perp\!\!\!\perp \varepsilon_1(t), \varepsilon_2(t) \quad \text{per } t=0,1,2,\dots$$

Per quanto riguarda i coefficienti e la distribuzione iniziale (θ_0, Y_0) avanziamo le seguenti assunzioni:

- 1) se $g(t,y)$ è una delle funzioni di $a_{0,i}$, $A_{0,i}$, $b_{ij}^{(1)}$, $b_{ij}^{(2)}$, $B_{ij}^{(1)}$, $B_{ij}^{(2)}$ allora

$$E | g(t,y) |^2 < \infty$$

- 2) con probabilità 1

$$|a_{ij}^{(1)}(t,y)| \leq c, \quad |A_{ij}^{(1)}(t,y)| \leq c$$

$$3) E(||\theta_0||^2 + ||Y_0||^2) < \infty \quad ||x|| = \sum_i x_i^2$$

- 4) La distribuzione condizionata

$$P(\theta_0 \leq x \mid Y_0)$$

è, con probabilità 1, normale di parametri m_0 e C_0 .

Tali assunzioni permettono di dimostrare (si veda Liptser e Shirayev (1972)) che la distribuzione

$$P(\theta_{t+1} \leq x \mid Y^t)$$

è normale con valore atteso m_{t+1} e matrice di varianze e covarianze C_{t+1} .

Le espressioni dei parametri m_{t+1} e C_{t+1} si possono agevolmente determinare con il seguente procedimento. Per semplicità di notazione elimineremo nella dimostrazione il simbolo (t,y) .

Assumiamo che la distribuzione

$$P(\theta_t \leq \alpha \mid Y^t)$$

sia normale con valore atteso m_t e matrice di varianze e covarianze C_t .

Dalle espressioni (8.4) e (8.5) si deduce che la distribuzione condizionata

$$P(\theta_{t+1} \leq \alpha, Y_{t+1} \leq x \mid Y^t, \theta_t = \alpha)$$

è normale con valore atteso

$$A_0 + A_1 \alpha = \begin{pmatrix} a_0 + a_1 \alpha \\ A_0 + A_1 \alpha \end{pmatrix} \quad (8.6)$$

$$B = \begin{pmatrix} b o b & b o B \\ (b o B)' & B o B \end{pmatrix} \quad (8.7)$$

dove

$$\begin{aligned} b o b &= b_1 b_1' + b_2 b_2' \\ b o B &= b_1 B_1' + b_2 B_2' \\ B o B &= B_1 B_1' + B_2 B_2' \end{aligned}$$

Dalla (8.6) si ha

$$E(\theta_{t+1} \mid Y^t) = a_0 + a_1 m_t \quad (8.8)$$

$$E(Y_{t+1} \mid Y^t) = A_0 + A_1 m_t \quad (8.9)$$

Per determinare la matrice di varianze e covarianze sottraiamo dalla (8.4)

e (8.5) le espressioni (8.8) e (8.9) rispettivamente.

Otteniamo:

$$\theta_{t+1} - E(\theta_{t+1} \mid Y^t) = a_1 (\theta_t - m_t) + b_1 \epsilon_1 + b_2 \epsilon_2 \quad (8.10)$$

$$Y_{t+1} - E(Y_{t+1} \mid Y^t) = A_1 (\theta_t - m_t) + B_1 \epsilon_1 + B_2 \epsilon_2 \quad (8.11)$$

Dalla (8.10) e (8.11) si ottiene

$$D_{11} = \text{Cov}(\theta_{t+1} | Y^t) = a_1 C_t a_1' + b \circ b$$

$$D_{12} = \text{Cov}(\theta_{t+1}, Y_{t+1} | Y^t) = a_1 C_t A_t' + b \circ B$$

$$D_{22} = \text{Cov}(Y_{t+1} | Y^t) = A_1 C_t A_1' + B \circ B$$

Dal teorema della correlazione normale (si veda il paragrafo 7)
è agevole verificare che:

$$m_{t+1} = E(\theta_{t+1} | Y_{t+1}, Y^t) = E(\theta_{t+1} | Y^t) + D_{12} D_{22}^{-1} (y_{t+1} - E(Y_{t+1} | Y^t)) \quad (8.12)$$

$$C_{t+1} = \text{Cov}(\theta_{t+1} | Y_{t+1}, Y^t) = D_{11} - D_{12} D_{22}^{-1} D_{21} \quad (8.13)$$

Sostituendo nella (8.12) e (8.13) le espressioni di

$$E(\theta_{t+1} | Y^t), E(Y_{t+1} | Y^t), D_{11}, D_{12}, D_{22}$$

otteniamo

$$m_{t+1} = (a_0 + a_1 m_t) + (b \circ B + a_1 C_t A_t') (B \circ B + A_1 C_t A_1')^{-1} (y_{t+1} - A_0 + A_1 m_t) \quad (8.14)$$

$$C_{t+1} = (b \circ b + a_1 C_t a_1') - (b \circ B + a_1 C_t A_t') (B \circ B + A_1 C_t A_1')^{-1} (b \circ B + a_1 C_t A_t') \quad (8.15)$$

Le equazioni (8.14) e (8.15) forniscono le espressioni ricorsive per il filtro non lineare ottenute per successioni di variabili aleatorie condizionatamente normali.

Vediamo ora di fare alcune rapide osservazioni:

OSSERVAZIONE 1.

Se poniamo :

$$a_0(t,y) = a_0(t) + a_2(t)y_t \quad ; \quad A_0(t,y) = A_0(t) + A_2(t)y_t$$

$$a_1(t,y) = a_1(t) \quad ; \quad A_1(t,y) = A_1(t)$$

$$b_i(t,y) = b_i(t) \quad i=1,2 \quad ; \quad B_i(t,y) = B_i(t) \quad i=1,2$$

dove tutte le funzioni dipendono solamente da t e non da y e se il vettore (θ_0, Y_0) è distribuito normalmente, allora il processo $(\theta_t, Y_t) \quad t=0,1,2,\dots$ è anch'esso distribuito normalmente. In questa circostanza la matrice di varianze e covarianze C_t non è aleatoria.

OSSERVAZIONE 2.

Sia (θ_t, Y_t) una successione parzialmente osservabile che soddisfi per $t \geq 1$ il seguente sistema di equazioni:

$$\theta_{t+1} = a_0(t,y) + a_1(t,y)\theta_t + b_1(t,y)\epsilon_1(t+1) + b_2(t,y)\epsilon_2(t+1) \quad (8.16)$$

$$Y_{t+1} = A_0^*(t-1,y) + A_1^*(t-1,y)\theta_t + B_1^*(t-1,y)\epsilon_1(t) + B_2^*(t-1,y)\epsilon_2(t) \quad (8.17)$$

Assumiamo che

$$P(\theta \leq \mid Y_1) \sim N(m_1, C_1).$$

Sebbene il sistema di equazioni (8.16) e (8.17) non possa essere formalmente ricondotto allo schema (8.4) e (8.5) ciò nonostante per poter determinare m_t e C_t è possibile utilizzare le equazioni ricorsive date nelle espressioni (8.14) e (8.15).

Dalla (8.17) si ha

$$Y_{t+1} = A_0^*(t,y) + A_1^*(t,y)(a_0(t,y) + a_1(t,y) \theta_t + b_1(t,y) \epsilon_1(t+1) + b_2(t,y) \epsilon_2(t+1)) \\ + B_1^*(t,y) \epsilon_1(t+1) + B_2^*(t,y) \epsilon_2(t+1)$$

Ponendo nella (8.5)

$$A_0 = A_0^* + A_1^* a_0$$

$$A_1 = A_1^* a_1$$

$$B_1 = A_1^* b_1 + B_1^*$$

$$B_2 = A_1^* b_2 + B_2^*$$

è immediato verificare che (θ_t, Y_t) soddisfa le equazioni (8.4) e (8.5).

Pertanto m_{t+1} e C_{t+1} possono essere determinate mediante le equazioni (8.14) e (8.15).

OSSERVAZIONE 3. (IL FILTRO DI KALMAN-BUCY)

Supponiamo che la successione gaussiana (θ_t, Y_t) $t=0,1,2,\dots$ soddisfaccia il seguente sistema di equazioni:

$$\theta_{t+1} = a_0(t) + a_1(t) \theta_t + b_1(t) \epsilon_1(t+1) + b_2(t) \epsilon_2(t+1) \quad (8.18)$$

$$Y_t = A_0(t) + A_1(t) \theta_t + B_1(t) \epsilon_1(t) + B_2(t) \epsilon_2(t) \quad (8.19)$$

In tale sistema, come si può notare, i coefficienti dipendono solamente dal tempo t e non dalle osservazioni y_t .

Se nella (8.5) effettuiamo le seguenti sostituzioni

$$A_0 = A_0^* + A_1^* a_0(t) \quad ; \quad A_1 = A_1^* a_1(t)$$

$$B_1 = A_1 b_1 + B_1 \quad ; \quad B_2 = A_1 b_2 + B_2$$

otteniamo

$$Y_{t+1} = A_0 + A_1 (a_0 + a_1 \theta_t + b_1 \epsilon_1 + b_2 \epsilon_2) + B_1 \epsilon_1 + B_2 \epsilon_2$$

Il sistema (8.18), (8.19) può ricondursi pertanto al sistema (8.4), (8.5),
pertanto possiamo applicare le equazioni ricorsive (8.14) e (8.15) per
determinare l'evoluzione di m_t e C_t :

$$m_{t+1} = (a_0 + a_1 m_t) + P(t)(Q(t))^{-1} (y_{t+1} - A_0(t+1) - A_1(t+1)a_0(t) - A_1(t+1)a_1(t)m_t) \quad (8.20)$$

con

$$P(t) = b_1(t)(A_1(t+1)b_1(t) + B_1(t+1))' b_2(t)(A_1(t+1)b_2(t) + B_2(t+1))' + a_1(t)C_t(a_1(t)'(A_1(t+1) \quad (8.21)$$

$$Q(t) = (A_1(t+1)b_1(t) + B_1(t+1))(A_1(t+1)b_1(t) + B_1(t+1))' + \\ (A_1(t+1)b_2(t) + B_2(t+1))(A_1(t+1)b_2(t) + B_2(t+1))' + \\ (A_1(t+1)a_1(t) C_t a_1(t)' A_1(t+1)') \quad (8.22)$$

E' agevole verificare che

$$C_{t+1} = (a_1(t) C_t a_1(t)' + b_1 b_1) - P(t)(Q(t))^{-1} P(t)' \quad (8.23)$$

Le equazioni (8.20) e (8.23) sono le equazioni del filtro di Kalman-Bucy
dove $P(t)(Q(t))^{-1}$ è la matrice guadagno.

Il teorema della correlazione normale ci permette altresì di determinare le espressioni delle condizioni iniziali:

$$m_0 = E(\theta_0 | Y_0) = E(\theta_0) + \text{Cov}(\theta_0) A_1'(0) (A_1(0) \text{Cov}(\theta_0) A_1'(0) + B \circ B)^{-1} (y_0 - A_0(0) - A_1(0) E(\theta_0)) \quad (8.24)$$

$$C_0 = \text{Cov}(\theta_0 | Y_0) = \text{Cov}(\theta_0) - \text{Cov}(\theta_0) A_1(0)' (A_1(0) \text{Cov}(\theta_0) A_1(0) + B \circ B)^{-1} A_1(0) \text{Cov}(\theta_0) \quad (8.25)$$

8.3 ALCUNI SEMPLICI ESEMPI.

8.3.1. Stima dei parametri.

Sia $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ un vettore gaussiano con valore atteso m e matrice di varianze e covarianze C . Vogliamo stimare θ mediante il processo stocastico $\{Y_t\} \quad t = 0, 1, 2, \dots$

Tale processo soddisfa la seguente equazione ricorsiva

$$Y_{t+1} = A_0(t, y) + A_1(t, y) \theta + B_1(t, y) \varepsilon(t+1)$$

con $Y_0 = 0$.

Assumendo valide le assunzioni fatte sui coefficienti nel paragrafo precedente otteniamo

$$E(\theta | y^{t+1}) = m_{t+1} = m_t + C_t A_1' (B_1 B_1' + A_1 C_t A_1')^{-1} (y_{t+1} - A_0 - A_1 m_t)$$

$$\text{Cov}(\theta | y^{t+1}) = C_{t+1} = C_t - C_t A_1' (B_1 B_1' + A_1 C_t A_1')^{-1} A_1 C_t$$

Semplificando la (8.26) con $A_0 = 0$ e $B_1 = 1$ otteniamo

$$Y_{t+1} = A_1(t, y) \theta + \varepsilon(t+1)$$

In tal caso si ha:

$$m_{t+1} = m_t + C_t A_1' (1 + A_1 C_t A_1')^{-1} (y_{t+1} - A_1 m_t)$$

Se si pone $A_1 = 1$ si deduce

$$Y_{t+1} = \theta + \varepsilon(t+1) \quad t = 0, 1, 2, \dots$$

e la stima di θ è (allorché $\varepsilon \sim N(0, 1)$)

$$m_{t+1} = m + (1 / (1+tC)) (C \sum_{i=1}^t y_i)$$

8.3.2. Il modello lineare dinamico nella formulazione presentata nel paragrafo 3.

Sia

$$Y_t = F_t \theta_t + v_t \quad v_t \sim N(0, V_t)$$

$$\theta_t = G_t \theta_{t-1} + w_t \quad w_t \sim N(0, W_t)$$

Se confrontiamo tale modello con quello proposto all'osservazione n.3

occorre porre:

$$a_0(t) = 0 ; A_0(t) = 0 ; a_1(t) = G_{t+1} ; A_1(t) = F_t ;$$

$$b_1(t) = 0 ; b_2(t) = (W_{t+1})^{1/2} ; B_1(t) = (V_t)^{1/2} ; B_2(t) = 0$$

Applicando le equazioni ricorsive (8.20), (8.21), (8.22), (8.23) si ottiene

$$m_t = G_t m_{t-1} + \underbrace{P(t)(Q(t))^{-1}}_{K_t} (y_t - F_t G_t m_{t-1})$$

con

$$P(t) = W_t F_t' + G_t C_{t-1} G_t' F_t' = (W_t + G_t C_{t-1} G_t') F_t'$$

$$Q(t) = (V_t + F_t (W_t + G_t C_{t-1} G_t') F_t')$$

E' agevole verificare che tali espressioni coincidono con quelle date nel paragrafo 3 e precisamente le espressioni (4.11) e (4.12).

RIFERIMENTI BIBLIOGRAFICI

Liptser- Shirayayev (1972), (1977). Kallianpur-Streibl (1968) (1969)
Kallianpur (1980)

9. IL PROBLEMA DELL'INTERPOLAZIONE (SMOOTHING)

Il problema che ci poniamo è quello di stimare θ_s note le osservazioni fino al tempo t , y^t , con $t > s$.

Indichiamo con

$$m(s,t) = E(\theta_s | y^t) \quad (9.1)$$

$$C(s,t) = \text{Cov}(\theta_s | y^t) \quad (9.2)$$

il valore atteso e la matrice di varianze e covarianze della distribuzione $(\theta_s | y^t)$. Evidentemente $m(s,t)$ è lo stimatore bayesiano di θ_s note le osservazioni fino al tempo t . Tale stimatore può essere definito da due differenti tipi di equazioni:

- a) equazioni in avanti (su t fissato s)
- b) equazioni all'indietro (su s fissato t)

9.1. Il modello lineare dinamico semplice.

Il filtro di Kalman -Bucy porta alla stima di θ_t in ogni istante t . L'unico stimatore che utilizza tutte le osservazioni è m_t , lo stimatore dello stato dell'ultimo periodo. Le equazioni di smoothing iniziano da m_t e C_t . Consideriamo le equazioni all'indietro.

Se $m(s,t)$ indica lo stimatore θ_s note le osservazioni fino al tempo t e $C(s,t)$ indica la matrice di varianze e covarianze al tempo s , le equazioni di smoothing sono le seguenti:

$$m(s,t) = m(s,s) + C^*(s,s) (m(s+1,t) - G_{s+1} m(s,s)) \quad (9.3)$$

$$C(s,t) = C(s,s) + C^*(s,s) (C(s+1,t) - C(s+1,s)) C^{*'}(s,s) \quad (9.4)$$

$$\text{con } c^*(s,s) = C(s,s)G'_{s+1}(C(s+1,s))^{-1} \quad s=t-1, t-2, \dots, 1 \quad (9.5)$$

Per illustrare la procedura vediamo un semplice esempio.

Consideriamo il modello:

$$y_t = \theta_t + v_t$$

$$\theta_t = \theta_{t-1} + w_t$$

dove y_t e θ_t sono scalari, $v_t \sim N(0, a)$, $w_t \sim N(0, b)$.

La previsione di un passo di θ_t è data dall'equazione (4.7). Nel nostro esempio si ha

$$E(\theta_t | y^{t-1}) = m_{t-1}$$

$$\text{Cov}(\theta_t | y^{t-1}) = C_{t-1} + b.$$

L'equazione (4.11) fornisce la stima di m_t

$$m_t = m_{t-1} + (C_{t-1} + b)(C_{t-1} + a + b)^{-1}(y_t - m_{t-1})$$

e

$$C_t = (C_{t-1} + b) - (C_{t-1} + b)^2(C_{t-1} + a + b)^{-1}.$$

Pertanto gli stimatori di smoothing sono: (posto $t=10$, $s=3$)

$$m(3,10) = m(3,3) + C(3,3)(m(4,10) - m(3,3)) / C(4,3)$$

$$C(3,10) = C(3,3) + (C(3,3)/C(4,3))^2 (C(4,10) - C(4,3)).$$

9.2. Il modello lineare dinamico generalizzato

Assumiamo che $(\theta_s | y^s)$ si distribuisca normalmente allora si può dimostrare che la distribuzione condizionata $((\theta_s, Y_t) | Y^{t-1})$ è anch'essa normale. Pertanto posto $t = s+1$ sarà distribuita normalmente anche $(\theta_s, Y_{s+1}) | Y^s$. Non è difficile provare allora che $(\theta_s | Y^{s+1})$ è normale. Procedendo per induzione si ha che per ogni $t > s$

$$(\theta_s | y^t) \sim N(m(s,t), C(s,t))$$

Il problema ora è quello di determinare le equazioni ricorsive in avanti per $m(s,t)$ e $C(s,t)$.

Essendo $(\theta_s, Y_t) | Y^{t-1}$ distribuita normalmente per il teorema della correlazione normale si ha:

$$E(\theta_s | Y^{t-1}, Y_t) = E(\theta_s | Y^{t-1}) + D_{12} D_{22}^{-1} (y_t - E(y_t | Y^{t-1})). \quad (9.6)$$

Determiniamo le espressioni D_{12} e D_{22} .

$$1) D_{22} = \text{Cov}(y_t y_t | Y^{t-1}) = A_1(t-1, y) C_{t-1} A_1'(t-1, y) + B \sigma B'(t-1, y) \quad (9.7)$$

2) Prima di determinare D_{12} è necessario calcolare $E(\theta_{t-1} | Y^{t-1})$

$$m_{t-1} = E(\theta_{t-1} | Y^{t-1}) = E\{E(\theta_{t-1} | Y^{t-1}, \theta_s) | Y^{t-1}\} \quad (9.8)$$

Essendo

$$E(\theta_t | Y^t, \theta_s = \alpha) = \varphi_s^t \alpha + \psi_s^t \quad (9.9)$$

con

$$\varphi_s^s = I$$

$$\varphi_s^t = \prod_{u=s}^{t-1} (a_1 - (b_0 B + a_1 C(u,s) A'_1) (B_0 B + A_1 C(u,s) A'_1)^{-1} A_1)$$

e

$$\psi_s^t = \sum_{u=s}^{t-1} \varphi_u^{t-1} (a_0 + (b_0 B + a_1 C(u,s) A'_1) (B_0 B + A_1 C(u,s) A'_1)^{-1} (y_{u+1} - A_0))$$

(per semplicità abbiamo eliminato il simbolo (u,y) ai coefficienti)

Per la (9.9) , la (9.8) diventa:

$$\begin{aligned} m_{t-1} &= E(\theta_{t-1} | y^{t-1}) = E(\varphi_s^{t-1} \theta_s + \psi_s^{t-1} | y^{t-1}) \\ &= \varphi_s^{t-1} E(\theta_s | y^{t-1}) + \psi_s^{t-1} \\ &= \varphi_s^{t-1} m(s, t-1) + \psi_s^{t-1} \end{aligned} \quad (9.10)$$

Inoltre

$$\begin{aligned} E((\theta_{t-1} - m_{t-1}) | y^{t-1}, \theta_s) &= \\ &= \varphi_s^{t-1} \theta_s + \psi_s^{t-1} - (\varphi_s^{t-1} m(s, t-1) + \psi_s^{t-1}) \\ &= \varphi_s^{t-1} (\theta_s - m(s, t-1)) \end{aligned} \quad (9.11)$$

Dalle equazioni che reggono il modello si deduce altresì che:

$$E(y_t | y^{t-1}) = A_0 + A_1 m_{t-1} \quad (9.12)$$

e che

$$\begin{aligned} E((y_t - E(y_t | y^{t-1})) | y^{t-1}, \theta_s) &= \\ &= E([A_1(\theta_{t-1} - m_{t-1}) + B_1 \epsilon_1 + B_2 \epsilon_2] | y^{t-1}, \theta_s) = \\ &= E([A_1(\theta_{t-1} - m_{t-1})] | y^{t-1}, \theta_s) = \end{aligned}$$

$$\begin{aligned}
&= A_1' E((\theta_{t-1} - m_{t-1})' | y^{t-1}, \theta_s) \\
&= (\theta_s - m(s, t-1))' (\varphi_s^{t-1})' A_1' \quad (9.13)
\end{aligned}$$

Siamo ora in grado di calcolare D_{12}

$$\begin{aligned}
D_{12} &= \text{Cov} (\theta_s, y_t | y^{t-1}) = E\{ (\theta_s - m(s, t-1))(y_t - E(y_t | y^{t-1}))' | y^{t-1} \} = \\
&= E\{ (\theta_s - m(s, t-1))(\theta_s - m(s, t-1))' (\varphi_s^{t-1})' A_1' | y^t \} = \\
&= C(s, t-1) (\varphi_s^{t-1})' A_1'(t-1, y) \quad (9.14)
\end{aligned}$$

Sostituendo ora la (9.7) e la (9.14) nella (9.6) otteniamo

$$\begin{aligned}
m(s, t+1) &= m(s, t) + C(s, t) (\varphi_s^{t-1})' A_1'(t, y) (B \otimes B + A_1(t, y) C_t A_1'(t, y))^{-1} \\
&\quad \times (y_{t+1} - A_0(t, y) + A_1(t, y) m_t) \quad (9.15)
\end{aligned}$$

Analogamente per

$$C(s, t) = D_{11} - D_{12} D_{22}^{-1} D_{12}' \quad (9.16)$$

Essendo $D_{11} = C(s, t-1)$ ed utilizzando la (9.7) e la (9.14) si ha

$$\begin{aligned}
C(s, t+1) &= C(s, t) - C(s, t) (\varphi_s^t)' A_1'(t, y) (B \otimes B + A_1(t, y) C_t A_1'(t, y))^{-1} \\
&\quad \times (A_1(t, y) \varphi_s^t C(s, t)) \quad (9.17)
\end{aligned}$$

Nella (9.15) e (9.17) $m_t = m(t, t)$ e $C_t = C(t, t)$.

Vediamo ora come dedurre le equazioni all'indietro per $m(s, t)$ e $C(s, t)$

Abbiamo ricordato che $(\theta_s, \theta_t | y^t)$ si distribuisce normalmente allora per il teorema della correlazione normale

$$E(\theta_s | y^t, \theta_t = \beta) = m^*(s, t) = E(\theta_s | y^t) + D_{12} D_{22}^{-1} (\beta - E(\theta_t | y^t)) \quad (9.18)$$

e

$$\text{Cov}(\theta_s | y^t, \theta_t = \beta) = C^*(s, t) = D_{11} - D_{12} D_{22}^{-1} D_{12}'$$

dove

$$D_{11} = \text{Cov}(\theta_s, \theta_s | y^t) = C(s, t) \quad (9.19)$$

$$D_{12} = \text{Cov}(\theta_s, \theta_t | y^t)$$

$$D_{22} = \text{Cov}(\theta_t, \theta_t | y^t)$$

Dalla (9.11) ricaviamo che

$$\begin{aligned} E((\theta_t - m_t)' | y^t, \theta_s) &= \theta_s' (\varphi_s^t)' + (\psi_s^t)' - (m'(s, t) (\varphi_s^t)' + (\psi_s^t)') \\ &= (\theta_s - m(s, t))' (\varphi_s^t)' \end{aligned} \quad (9.20)$$

da cui

$$\begin{aligned} D_{12} &= E((\theta_s - m(s, t))(\theta_t - m_t)' | y^t) = \\ &= E\{(\theta_s - m(s, t))\} E\{(\theta_t - m_t)' | y_t \theta_s | y^t\} \\ &= D_{11} (\varphi_s^t)' = C(s, t) (\varphi_s^t)' \end{aligned} \quad (9.21)$$

Sostituendo nella (9.18) otteniamo

$$m^*(s,t) = m(s,t) + C(s,t) \begin{pmatrix} \varphi_s^t \end{pmatrix}' C_t^{-1} (\beta - m_t) \quad (9.22)$$

e

$$C^*(s,t) = C(s,t) - C(s,t) \begin{pmatrix} \varphi_s^t \end{pmatrix}' C_t^{-1} \begin{pmatrix} \varphi_s^t \end{pmatrix} C(s,t) \quad (9.23)$$

con $m^*(s,s) = \beta$ e $C^*(s,s) = 0$.

Se sono soddisfatte le condizioni sui coefficienti del modello allora

si dimostra che:

$$m^*(s,t) = m(s,s+1) + C(s,s+1) \begin{pmatrix} \varphi_s^{s+1} \end{pmatrix}' C_{s+1}^{-1} (m^*(s+1,t) - m_{s+1}) \quad (9.24)$$

$$C^*(s,t) = C^*(s,s+1) + C(s,s+1) \begin{pmatrix} \varphi_s^{s+1} \end{pmatrix}' C_{s+1}^{-1} C^*(s+1,t) C_{s+1}^{-1} \begin{pmatrix} \varphi_s^{s+1} \end{pmatrix} C(s,s+1) \quad (9.25)$$

con $m^*(t,t) = \beta$ e $C^*(t,t) = 0$.

In conclusione possiamo dare le equazioni all'indietro del valore atteso e della matrice di varianze e covarianze della distribuzione

$$\begin{pmatrix} \theta_s | y^t \end{pmatrix} \quad \text{per } s < t$$

$$m(s,t) = m(s,s+1) + C(s,s+1) \begin{pmatrix} \varphi_s^{s+1} \end{pmatrix}' C_{s+1}^{-1} (m(s+1,t) - m_{s+1}) \quad (9.26)$$

e

$$C(s,t) = C^*(s,s+1) + C(s,s+1) \begin{pmatrix} \varphi_s^{s+1} \end{pmatrix}' C_{s+1}^{-1} C(s+1,t) C_{s+1}^{-1} \begin{pmatrix} \varphi_s^{s+1} \end{pmatrix} C(s,s+1) \quad (9.27)$$

con $m(t,t) = m_t$, $C(t,t) = C_t$,

Riferimenti bibliografici

Liptser-Shiryayev (1972)(1977), Kallianpur (1982).

10. IL PROBLEMA DELLA PREVISIONE.

Nella pratica sovente è desiderabile effettuare previsioni non di un solo periodo ma di più periodi. In tal caso è necessario determinare la distribuzione predittiva di una futura osservazione y_{t+k} con $k \geq 1$.

10.1. IL modello lineare dinamico semplice.

Nel paragrafo 4 abbiamo dato la previsione di un passo di θ_t e di y_t . La previsione di θ_t note le osservazioni fino al tempo $t-1$ è data dal valore atteso della distribuzione $(\theta_t | y^{t-1})$.

Dalla (4.7) si deduce che:

$$E(\theta_t | y^{t-1}) = G_t m_{t-1}$$

$$\text{Cov}(\theta_t | y^{t-1}) = G_t C_{t-1} G_t' + W_t = R_t$$

La previsione di y_t note le osservazioni fino al tempo $t-1$ è data:

$$E(y_t | y^{t-1}) = F_t E(\theta_t | y^{t-1}) = F_t G_t m_{t-1}$$

$$\begin{aligned} \text{Cov}(y_t | y^{t-1}) &= F_t \text{Cov}(\theta_t | y^{t-1}) F_t' + V_t = \\ &= F_t R_t F_t' + V_t. \end{aligned}$$

Dalle equazioni del modello (3.1) e (3.2) possiamo dedurre l'osservazione y_{t+k} come segue:

$$y_{t+k} = F_{t+k} \theta_{t+k} + v_{t+k} \quad (10.1)$$

$$\theta_{t+k} = G \theta_{t+k-1} + w_{t+k} \quad (10.2)$$

In tale modello a differenza del modello introdotto nelle (3.1) e (3.2) abbiamo assunto G indipendente dal tempo.

Assumiamo inoltre note le varianze dei vettori aleatori v_{t+k} e w_{t+k} .

Per effettuare la previsione del valore y_{t+k} è necessario dapprima inferire su θ_{t+k} e su F_{t+k} .

Dobbiamo pertanto determinare :

$$m_{k,t} = E(\theta_{t+k} | y^t) \quad (10.3)$$

$$C_{k,t} = \text{Cov}(\theta_{t+k} | y^t) \quad (10.4)$$

dove

$$m_{0,t} = E(\theta_t | y^t)$$

$$C_{0,t} = \text{Cov}(\theta_t | y^t)$$

Le espressioni di $m_{0,t}$ e $C_{0,t}$ si determinano applicando le espressioni (4.11) e (4.12) che fornisco nel filtro di KALMAN- BUCY.

Allora dalla (11.2) si ha:

$$m_{k,t} = G m_{k-1,t} \quad (10.5)$$

$$C_{k,t} = G C_{k-1,t} G' + W_{t+k}$$

Il valore atteso e la matrice di varianze e covarianze di $(\theta_{t+1} | y^t)$ può calcolarsi ricorsivamente per $k = 1, 2, \dots$.

Assumiamo inoltre che F_{t+k} sia nota. Avremo:

$$\hat{y}_{t+k} = F_{t+k} m_{k,t} \quad (10.6)$$

$$\begin{aligned} \text{Cov}(y_{t+k} | y^t) &= F_{t+k} \text{Cov}(\theta_{t+k} | y^t) F_{t+k}' + v_{t+k} \\ &= F_{t+k} C_{k,t} F_{t+k}' + v_{t+k} \end{aligned}$$

Facciamo osservare infine che il caso in cui F_{t+k} non è nota la previsione è più complicata ed inoltre l'analisi di Harrison e Stevens è errata(si veda Blight(1976)).

Diamo ora le previsioni di K passi in avanti di tre esempi presentati nel paragrafo 6.

PREVISIONI DA 1 A 12 PASSI AVANTI - INDICE GEN. PREZZI

MSP

 564.45
 569.4
 574.35
 579.3
 584.25
 589.19
 594.14
 599.09
 604.04
 608.99
 613.94
 618.89

PREVISIONI DA 1 A 10 PASSI AVANTI - IMPORT

MSP

 19.02
 19.598
 20.176
 20.754
 21.332
 21.91
 22.488
 23.067
 23.645
 24.223

PREVISIONI DA 1 A 10 PASSI AVANTI - EXPORT

MSP

 21.862
 22.519
 23.177
 23.834
 24.492
 25.149
 25.807
 26.465
 27.122
 27.78

10.2. Il modello lineare dinamico generalizzato.

Dato un processo parzialmente osservabile (θ_t, y_t) retto dalle equazioni (8.4) e (8.5) il problema della previsione consiste nella stima dei vettori θ_t e y_t note le osservazioni fino ad s , $(y_0, y_1, \dots, y_s) = y^s$. Evidentemente $s < t$. Le previsioni di θ_t e y_t si possono dedurre solo in casi particolari poichè la distribuzione condizionata

$$((\theta_t, y_t) | y^s)$$

non si distribuisce normalmente.

E' possibile dimostrare che se sono soddisfatte le assunzioni avanzate al numero 8.2 allora per ogni t esiste un vettore gaussiano $\bar{\epsilon}(t) = (\bar{\epsilon}_1(t), \dots, \bar{\epsilon}_k(t))$ con $E(\bar{\epsilon}(t)) = 0$ e $E(\bar{\epsilon}(t) \bar{\epsilon}'(s)) = \delta(t-s) I$ tale che valga la seguente rappresentazione dei vettori y_{t+1} :

$$y_{t+1} = A_0(t, y) + A_1(t, y) m_t + (B \circ B(t, y) + A_1(t, y) C_t A_1'(t, y))^{1/2} \bar{\epsilon}(t+1). \quad (10.7)$$

Se sostituiamo nella espressione del valore atteso della distribuzione finale $(\theta_t | y^t)$ la (10.7) otteniamo:

$$m_{t+1} = a_0(t, y) + a_1(t, y) m_t + (B \circ B(t, y) + A_1(t, y) C_t A_1'(t, y))^{1/2} \bar{\epsilon}(t+1) \quad (10.8)$$

Prendiamo ora il valore atteso delle espressioni (10.7) e (10.8), essendo $E(\bar{\epsilon}(t+1)) = 0$, avremo:

$$E(E(y_{t+1} | y^t) | y^s) = A_0(t, y) + A_1(t, y) E(E(\theta_t | y^t) | y^s) \quad (10.9)$$

$$E(E(\theta_{t+1} | y^t) | y^s) = a_0(t, y) + a_1(t, y) E(E(\theta_t | y^t) | y^s) \quad (10.10)$$

La determinazione simultanea dei valori attesi $E(Y_{t+1} | Y^S)$ e $E(m_{t+1} | Y^S)$ è possibile se e solo se i coefficienti della (10.9) e (10.10) non dipendono dalle osservazioni passate ma solo dal tempo e se:

$$\begin{aligned} a_0(t,y) &= a_0(t) + a_2(t) y(t) \\ a_1(t,y) &= a_1(t) \\ A_0(t,y) &= A_0(t) + A_2(t) y(t) \\ A_1(t,y) &= A_1(t) \end{aligned}$$

Indicando con

$$n_1(t,s) = E(E(\theta_{t+1} | y^t) | y^s) = E(m_{t+1} | y^s)$$

$$n_2(t,s) = E(E(y_{t+1} | y^t) | y^s) = E(y_{t+1} | y^s)$$

avremo dalla (10.9) e (10.10) i seguenti risultati:

$$n_2(t+1,s) = A_0(t) + A_2 n_2(t,s) + A_1 n_1(t,s) \quad (10.11)$$

$$n_1(t+1,s) = a_0(t) + a_2 n_2(t,s) + a_1 n_1(t,s) \quad (10.12)$$

Qualora fossimo interessati solo alla stima di θ_t è necessario porre $a_2(t) = 0$ e si ha:

$$n_1(t+1,s) = a_0(t) + a_1(t) n_1(t,s)$$

con

$$n_1(t,t) = m_t$$

BIBLIOGRAFIA

- B. Abraham - J. Ledolter (1982) : *Statistical Methods for Forecasting*, J. Wiley, New York.
- H. Akaike (1974) : Stochastic theory of minimal realizations, IEEE trans. on Automatic Control, AC. 19, 667-674.
- H. Akaike (1974) : Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes, Annals of the Institute of Statistical Mathematics, 363 - 387.
- B.D.O. Anderson-J.B. Moore (1979) : *Optimal Filtering*, Prentice-Hall, Inc. Englewood Cliffs, New Jersey.
- T.W. Anderson (1958) : *An Introduction to Multivariate Statistical Analysis*, J. Wiley, New York.
- M. Aoki (1967) : *Optimization of Stochastic Systems*, Academic Press, New York.
- M. Athans (1974): The importance of Kalman filtering methods for economic systems, Annals of Economic and Social Measurement, 3/1, 49-64.
- X AU. VV. (1983) : Special Issue on Applications of Kalman Filtering,
IEEE Trans. on Automatic Control.
- D.A. Belsley (1973): The applicability of the Kalman filter in the determination, of systematic parameter variation, Annals of Economic and Social Measurement 2/4, 531-533.

- B.J.N. Blight (1976) : discussione sul lavoro di Harrison e Stevens , J. Roy. Statist. Soc. Ser. B, 205- 247.
- G.E.P. Box - G.M. Jenkins (1970): Time Series Analysis, Forecasting and Control, Hodden-Bay, San Francisco.
- R.G. Brown (1963) : Smoothing, Forecasting and Prediction, Englewood Cliffs: Prentice Hall.
- N. Cantarelis- F.R. Johnston (1979): On line variance estimation for the steady-state bayesian forecasting model, Working paper n. 34, University of Warwick.
- N. Cantarelis - F.R. Johnston (1982): On line variance estimation for the steady-state bayesian forecasting model, Journal of Time series analysis, v. 3, 225-234.
- G. Chow (1975): Analysis and Control of Dynamic Economic Systems, J. Wiley, New York.
- W. Conrad - C. Corrado (1979): Application of the Kalman filter to revisions in monthly retail sales estimates, Journal of Economic Dynamics and Control, 177-198.
- M.H. De Groot (1970), Optimal Statistical Decisions, Mc Graw -Hill, New York.
- P. De Jong-B. Zehnworth (1983), Claims reserving, state-space models and the Kalman filter, Journal of the Institute of Actuaries, v. 110, 157-181.

- J.J. Deyst - C.F. Price (1968) : Conditions for asymptotic stability of the discrete minimum variance linear estimator, IEEE Trans. on Automatic Control , AC-13, 702-705.
- J.L. Doob (1944) : The elementary Gaussian processes, Ann. Math. Statist. ,15, 229-282.
- R.B. Gathercole - J.Q. Smith (1981) : Approximations for bayesian forecasting, Research Report, Dept.Statistical Science, University College London.
- R.B. Gathercole (1982) , A note for bayesian forecasters, in O. D. Anderson(ed.) Time Series Analysis:Theory and Practice 1, North-Holland.
- E.J. Godolphin - P.J. Harrison (1975) : Equivalence theorems for polynomial projecting predictors. J. Roy.Statist. Soc. Ser. B, 205- 215.
- P.J. Harrison (1965) : Short - term sales forecasting. J. Roy. Statist.Soc. , C, Appl. Statist. ,14, 102-139.
- P.J. Harrison (1967): Exponential smoothing and short-term sales forecasting, Management Science, 13, 821- 842.
- P.J. Harrison - C.F. Stevens (1971): A Bayesian approach to short term forecasting, Oper. Res. Quart. ,22, 341-362.
- P.J. Harrison- C.F. Stevens (1975) : Bayesian forecasting in action: case studies. Warwick: Statistics Research Report n.14

- P. J. Harrison- C.F. Stevens (1976): Bayesian forecasting
(with discussion) , J. Roy. Statist. Soc. Ser. B.,
205- 247.
- A.C. Harvey (1981) : The Kalman filter and it s application
in econometrics and time series analysis,
Methods of Operations Research, 3-18.
- A. C. Harvey (1981): Time Series Models
Phillip Allan, Dedington.
- X H. Heffes (1966): The effects of erroneous models on the Kalman
filter response, IEEE Trans. on Automatic Control,
AC.11, 541-543.
- S. Huyberegts (1979): Estimation des paramètres dans le modèle
linéaire générale dynamique. R.A.I.R.O. Operations
Research 13, 143-149.
- A. H. Jazwinski (1970): Stochastic Processes and Filtering Theory,
Academic Press, New York.
- F.R. Johnston-P.J.Harrison (1980): An application of forecasting
in the alcoholic drinks industry, J. Opl. Res. Soc.
v. 31, 699- 709.
- T. Kailath (1968): An innovations approach to least-squares
estimation. Part.I. Linear filtering in additive white
noise, IEEE Trans. on Automatic Control, AC 13, 646-655.
- T. Kailath (1981) : Lectures on Wiener and Kalman filtering,
Springer-Verlag.

- G. Kallianpur-C. Streibel (1968) : Estimation of stochastic processes with additive white noise observations errors; Ann. Math. Statist. 39, 785-801.
- G. Kallianpur- C. Streibel (1969): Stochastic differential equations in the estimation of continuous parameter stochastic processes, Theory Probab. Its Applic. 14, 567-594.
- G. Kallianpur (1980): Stochastic Filtering Theory, Springer Verlag.
- R. E. Kalman(1960): A new approach to linear filtering and prediction problems, J. Basic Eng. 82, 340-345.
- R.E. Kalman (1963): New methods in Wiener filtering theory, in G.L. Bogdanoff - J.Kozin (eds.) Proc. of First Symp.on Eng. Appl.of Random Function Theory and Probability, J. Wiley, New York.
- R.E. Kalman (1965): Linear stochastic filtering reappraisal and outlook. , Proc. Symp. on System Theory, Polytechnic Institute of Brooklyn, 197-205.
- R.E.Kalman-R.S.Bucy (1961): New results in linear filtering and prediction theory, Trans. ASME, Ser. D., Journal of Basic Eng. 83,95-107.
- A.N. Kolmogorov (1941): Interpolation and extrapolation of stationary random sequences, Bull. Moscow University, Ser. Math. 5.

R. Leoni (1979): Su un metodo di stima ricorsiva dello stato di un modello lineare dinamico, *Statistica*, 343-349.

D.V. Lindley (1965): *Introduction to Probability and Statistics from a bayesian wiewpoint*, vol.I,II.
Cambridge University Press.

D.V.Lindley - A.F.M. Smith (1972): Bayes estimates for the linear model (with discussion) . *J. Roy. Statist. Soc. Ser.B.* , 34, 1- 41.

A. Lindquist - G. Picci (1979): On the stocastic realization problem. *S.I.A.M. J. Control* , 365-384.

R.S. Liptser- A.N. Shyryayev (1972) : Statistics of conditionally Gaussian random sequences, *Proc. Sixtth Berkeley Symposium Math. Statistics and Probability (1970)*, vol. II, University of California Press, 389-422.

R.S. Liptser- A.N. Shyryayev (1977) : *Statistics of Random Processes*.
I. General theory ; II. Applications. Springer-Verlag.

X U.E. Makov (1980): Approximations to unsupervised filters,
IEEE Trans. on Automatic Control, AC.25, 842-847.

U.E. Makov-A.F.M. Smith (1976): Quasi-Bayes procedures for unsupervised learning. *Proc. IEEE conf. Decision and Control*.

X K. Mehra (1970): Identification of variances and adaptive Kalman filtering, IEEE Trans. on Automatic Control, AC.15, 2, 175-184.

R.K. Mehra (1971): Identification of linear dynamic systems,
AIAA Journal.

R.K. Mehra(1974): Identification and control and econometric systems,
similarities and differences, Annals of Economic and
Social Measurement,3,

R.K. Mehra (1976) : Identification and estimation of the error in
variables model in structural form, in Proc. of Symp.
on Stochastic Systems, North-Holland.

R.K. Mehra (1979): Kalman filters and their applications to
forecasting , in Makridakis-Wheelwright (eds.)
Special Issue on Forecasting , TIMS Studies in the
Management Sciences, 75- 94, North-Holland.

R. J. Meinhold - N.D. Singpurwalla (1983): Understanding the Kalman
filter, The American Statistician , 37, 123- 127.

G. Mélard (1980): Some links between the Harrison-Stevens and
Box-Jenkins methods for univariate time series, in
O.D. Anderson(ed.) Time Series, North-Holland.

G.W. Morrison-D.H.Pike (1967) : Kalman Filtering applied to
statistical forecasting, Manegement Science ,23,
7, 1768-1774.

T. Nishimura (1967) Error bounds of continuous Kalman filter and
the application to orbit determination problems,
IEEE Trans. on Automatic Control, AC.12 , 268-275.

A. O'Hagan (1976) : discussione sul lavoro di Harrison -Stevens
J. Roy.Statist.Soc. Ser.B, 38, 205-247.

G. Picci. (1976): Stochastic realizations of Gaussian processes
Proc. IEEE, 64, 112-122.

Plackett(1950) : Some theorems on least squares, Biometrika,
37, 149-157.

M.B. Prestley (1976) : discussione sul lavoro di Harrison-Stevens.
J. Roy. Statist. Soc. Ser. B. , 205-247.

C.R. Rao (1973) : Linear Statistical Inference and Its Applications,
2nd.ed. J. Wiley, New York.

Y.A. Rozanov (1976) : On two selected topics connected with
stochastic system theory. Applied Math. and
Optimization, 3, 73-80.

G. Ruckebusch (1977): On the theory of Markovian representation.
Proc. Measure theory applications to stochastic analysis.
Lecture Notes in Mathematics n. 695. Springer-Verlag.

A.H. Sarris (1973) : Kalman filter models a bayesian approach to
estimation of time varying regression coefficients ,
Annals of Economic and Sociel Measurement, 2/4, 501-523.

F.C. Schweppe (1965), Evaluation of likelihood functions for
Gaussian signals, IEEE Trans. Infomation Theory,
IT, 11, 61-70.

R.H. Shumway-D.S. Stoffer (1982): An approach to time series smoothing
and forecasting using the EM algorithm, Journal of
Time Series Analysis, v. 3, 253-263.

- A.F.M. Smith(1973):A General bayesian linear model,J.Roy. Statist. Soc. Ser.B,35, 67-75.
- A.F.M. Smith (1977) : A Bayesian analysis of some time varying models in J. Barra et al. (eds.) Recent Developments in Statistics, North-Holland.
- A.F.M. Smith (1980): Metodi bayesiani nelle serie temporali , Seminari presso l'Istituto di Calcolo delle Probabilità, Università di Roma.
- A.F.M. Smith (1980) : Change-point problems : approaches and applications, Trab. Estadist. 31, 83-98.
- A.F.M. Smith - U.E. Makov (1977): A quasi-Bayes approximation to unsupervised filters , Proc. Conf. on Measurement and Control, 102-106.
- A.F.M. Smith - U. E. Makov (1980): Bayesian detection and estimation of jumps in linear systems , in R.L. Jacobs et al. (eds.) Proc. of the IMA Conf. on " The Analysis and Optimization of Stochastic System", 333-346.
- A.F.M. Smith- D.G. Cook (1980): Straight lines with a change point: a Bayesian analysis of some renal transplant data, Applied Statistics 29, 130-189.
- AF.M. Smith - M. West (1983): Monitoring renal transplants: an application of the multiprocess Kalman filter, Biometrics, 39, 867-878.
- J.Q. Smith (1979): A generalization of the Bayesian Steady forecasting model, J. Roy. Statist. Soc. Ser. B, 41, 3, 375-387.

- J.Q. Smith - P.J. Harrison (1980): Discontinuity , decision and conflict in J.M. Bernardo et al.(eds.) Bayesian Statistics, University press, Valencia, 99-140.
- J.Q. Smith- P.J. Harrison-E.C. Zeeman (1981) : The analysis of some discontinuous decision processes, Eur. Jur. Oper. Res. 7,30-43.
- R.C. Souza (1982) A bayesian entropy approach to forecasting, the Binomial Beta model in O.D. Anderson (ed.) Time Series Analysis: Theory and Practice I. North-Holland.
- R.C. Souza - P.J. Harrison (1979) Steady-State System forecasting : A Bayesian Entropy approach. Inter. Report, Dept. of Statistics, University of Warwick.
- H.W. Sorenson (1966): Kalman filtering techniques advances in Control Systems, vol. 3. Academic press.
- R.J. Verral (1983) : Forecasting : the Bayesian approach, Journal of the Institute of Actuaries, 110,183-203.
- N. Wiener (1949): The Extrapolation , Interpolation and Smoothing of Stationary Time Series with Engineering Applications, J. Wiley, New York.
- H. Wold (1938): A study in the Analysis of Stationary Time Series, Almquist and Wiksell, Uppsala.
- W.H. Wonham (1970), Random differential equations in control theory, in A.T. Blarucha - Reis (eds.) Probabilistic Methods in Applied Mathematics, vol.2, Academic press, New York, 131-212.

65 66 07

70 71 72 73

76 77