# Serie storiche: Complementi ed Esercizi

Luisa Bisaglia Tommaso Di Fonzo Francesco Lisi Marco Marini

17 novembre 2020

# Indice

**Prefazione** **1**

**1 Concetti introduttivi** **3**

**2 Funzioni deterministiche del tempo** **5**

**3 Le medie mobili** **25**

**4 Il lisciamento esponenziale** **47**

**5 Processi stocastici** **65**

**6 Processi ARIMA** **85**

**7 Previsioni ARIMA** **115**

**8 Soluzioni agli esercizi supplementari** **121**

**Appendice 1. Uno sguardo ad alcuni risultati presentati in Di Fonzo e Lisi (2005)** **131**

**Appendice 2. Scomposizione del trinomio $ax^2 + bx + c$** **137**

**Appendice 3. La regione di stazionarietà di un processo AR(2)** **139**

# Prefazione

Questa raccolta di esercizi viene messa a disposizione

**ESCLUSIVAMENTE**

agli studenti iscritti nell'a.a. 2020/21 al corso di *Serie storiche* dei corsi di laurea in Statistica (SEI e STS) della Scuola di Scienze dell'Università di Padova.
Si tratta di materiale in corso di sistemazione, e perciò largamente incompleto e non immune da errori.
I docenti del corso di *Serie storiche*, Luisa Bisaglia e Tommaso Di Fonzo, e gli autori della raccolta, Francesco Lisi e Marco Marini, pregano gli studenti che individueranno refusi, imprecisioni ed errori, di segnalarli all'indirizzo di posta elettronica

`difonzo@stat.unipd.it`

Nell'intento di arrivare presto ad una versione definitiva del testo, sono graditi anche commenti ed osservazioni di carattere più generale, da inviare all'indirizzo di cui sopra.

Padova, settembre 2020

# 1
# Concetti introduttivi

### Le componenti di una serie storica

**1.1** Discutere, spiegandone i motivi, quali componenti ci si può aspettare nella serie storica

$$\{y_t\}_{t=1}^n$$

nei seguenti quattro casi specificando, se è il caso, il tipo di trend ed il periodo della componente stagionale: (a) Numero giornaliero di biglietti venduti allo sportello di una stazione ferroviaria negli ultimi due anni. (b) Numero di italiani che hanno un accesso ad Internet, rilevato mensilmente negli ultimi 8 anni. (c) Serie trimestrale del prodotto interno lordo in volumi concatenati negli ultimi 15 anni. (d) Serie annuale delle vendite di uno specifico prodotto dalla sua commercializzazione alla sua maturità.

□ (a) Il numero esiguo di anni osservati rende ipotizzabile la presenza di un trend costante nella serie $\{y_t\}_{t=1}^n$. La frequenza è giornaliera, pertanto è presumibile che vi sia una forte componente stagionale di periodo settimanale: a differenti giorni della settimana corrispondono infatti differenti quantità di biglietti ferroviari acquistati. Inoltre, i movimenti turistici possono provocare una stagionalità di periodo inferiore, possibilmente mensile. Scioperi, condizioni meteorologiche ed eventi particolari possono determinare la presenza di valori anomali. La componente erratica dovrebbe essere piuttosto stabile in relazione alle altre componenti.

(b) L'utilizzo di Internet nelle famiglie italiane è cresciuto esponenzialmente dalla seconda metà degli anni '90, è quindi plausibile la presenza di un trend esponenziale nella serie. Potrebbero essere evidenti delle flessioni stagionali nei mesi estivi, certamente con un'intensità minore rispetto alla tendenza di fondo del fenomeno. La componente erratica dovrebbe assumere una variabilità contenuta.

(c) La serie del PIL è la risultante dell'intera attività economica di un paese, pertanto è naturale attendersi una molteplicità di componenti in essa. Considerando il PIL di un paese industrializzato, è immaginabile una componente di crescita strutturale alla quale si affiancano marcate variazioni stagionali e cicliche. Il trend può essere colto da funzioni polinomiali del tempo, spesso anche del primo ordine (trend lineare), oppure si possono ipotizzare trend di tipo stocastico. La stagionalità è piuttosto regolare nel tempo, con picchi negativi in corrispondenza del terzo trimestre quando, nel periodo estivo, l'attività economica diminuisce. In 15 anni è inoltre possibile il completamento di 1-2 cicli economici, per cui possono risultare evidenti nella serie delle fasi di espansione e di recessione congiunturale.

(d) La curva di crescita in questo caso specifico è di tipo logistico, caratterizzata da un limite superiore che si presume non venga oltrepassato dalla serie che, dopo un periodo di forte crescita, prima rallenta e poi va ad assumere un andamento costante. La serie è annuale e quindi priva di stagionalità. La variabilità della componente erratica dipende molto dal tipo di prodotto in questione.

## Esercizi supplementari

**C1.1** In cosa consistono e a cosa servono le operazioni preliminari di pulizia di una serie storica?

# 2
# Funzioni deterministiche del tempo

### Modelli di regressione con funzioni deterministiche del tempo

**2.1** La serie storica quadrimestrale $\{y_t\}_{t=2}^{10}$, in cui $t = 1$ corrisponde al primo quadrimestre del primo anno di osservazione, è generata dal modello

$$Y_t = f(t) + \varepsilon_t,$$

con $f(t)$ pari alla somma tra un trend lineare ed una stagionalità costante rappresentata da variabili *dummy*.

(a) Si espliciti la componente deterministica $f(t)$ e si scriva la matrice dei regressori da usare per la stima simultanea delle componenti tendenziale e stagionale.

(b) Indicata con $S_t$ la componente stagionale definita al punto precedente, si scriva l'espressione per calcolare la serie destagionalizzata $y_t^d = y_t - S_t$, $t = 2, \ldots, 10$.

□ Il trend lineare è dato da

$$T_t = \alpha + \beta t$$

per $t = 2, \ldots, 10$. La stagionalità può essere rappresentata combinando le 3 variabili *dummy*

$$d_1 = \begin{cases} 1 & \text{nel primo quadrimestre} \\ 0 & \text{altrimenti} \end{cases}$$

$$d_2 = \begin{cases} 1 & \text{nel secondo quadrimestre} \\ 0 & \text{altrimenti} \end{cases}$$

$$d_3 = \begin{cases} 1 & \text{nel terzo quadrimestre} \\ 0 & \text{altrimenti} \end{cases}$$

mediante l'espressione

$$S_t = \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t}.$$

La componente deterministica sarebbe dunque pari a

$$f(t) = T_t + S_t = \alpha + \beta t + \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t}.$$

In questo caso, però, la matrice dei regressori $\mathbf{X}$ sarebbe di rango non pieno, il che renderebbe impossibile calcolare le stime dei parametri con il metodo dei minimi quadrati (si cade infatti nella cosiddetta

"trappola delle *dummy*"). Infatti, la matrice dei regressori, ottenuta affiancando le osservazioni relative, nell'ordine, alla costante, alla variabile $t$ ed alle *dummy* $d_{1t}$, $d_{2t}$ e $d_{3t}$, è data da

$$\mathbf{X} = \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 1 & 3 & 0 & 0 & 1 \\ 1 & 4 & 1 & 0 & 0 \\ 1 & 5 & 0 & 1 & 0 \\ 1 & 6 & 0 & 0 & 1 \\ 1 & 7 & 1 & 0 & 0 \\ 1 & 8 & 0 & 1 & 0 \\ 1 & 9 & 0 & 0 & 1 \\ 1 & 10 & 1 & 0 & 0 \end{bmatrix}$$

di dimensioni $(9 \times 5)$ e di rango $4$. È evidente che le colonne non sono linearmente indipendenti: ad esempio, la prima colonna è pari alla somma delle ultime tre.

Per non cadere nella trappola delle *dummy* conviene definire la componente deterministica come somma tra il trend lineare (privo di intercetta) e 3 variabili *dummy*:

$$f(t) = \beta t + \gamma_1^* d_{1t} + \gamma_2^* d_{2t} + \gamma_3^* d_{3t} \tag{2.1}$$

con matrice dei regressori pari a:

$$\mathbf{X}^* = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \\ 4 & 1 & 0 & 0 \\ 5 & 0 & 1 & 0 \\ 6 & 0 & 0 & 1 \\ 7 & 1 & 0 & 0 \\ 8 & 0 & 1 & 0 \\ 9 & 0 & 0 & 1 \\ 10 & 1 & 0 & 0 \end{bmatrix}.$$

Per calcolare la serie destagionalizzata, va anzitutto ricordato che, in generale, nel modello così formulato la somma dei coefficienti di regressione delle variabili *dummy* non è pari a zero (come invece richiede l'assunto che la componente stagionale si annulla nell'arco dell'anno). È dunque necessario calcolare i coefficienti 'ideali' di stagionalità, sottraendo ai coefficienti $\gamma_j^*$ la loro media $\bar{\gamma}^* = \frac{1}{3}\sum_{j=1}^{3} \gamma_j^*$:

$$\hat{\gamma}_j^* = \gamma_j^* - \bar{\gamma}^* \quad j = 1,2,3.$$

Possiamo così riscrivere la componente deterministica $f(t)$, secondo una parametrizzazione equivalente alla (2.1) e che può essere usata per il calcolo dei valori destagionalizzati:

$$\begin{aligned} f(t) &= \bar{\gamma}^* + \beta t + (\gamma_1^* - \bar{\gamma}^*) d_{1t} + (\gamma_2^* - \bar{\gamma}^*) d_{2t} + (\gamma_3^* - \bar{\gamma}^*) d_{3t} \\ &= \bar{\gamma}^* + \beta t + \hat{\gamma}_1 d_{1t} + \hat{\gamma}_2 d_{2t} + \hat{\gamma}_3 d_{3t}. \end{aligned}$$

I valori destagionalizzati si ottengono pertanto tramite l'espressione

$$y_t^d = y_t - (\hat{\gamma}_1 d_{1t} + \hat{\gamma}_2 d_{2t} + \hat{\gamma}_3 d_{3t}).$$

**Osservazione.** Per evitare la *trappola delle dummy* si sarebbe potuto scegliere un modello diverso dal (2.1). Ad esempio, se si considerasse un modello con l'intercetta e privo della terza variabile *dummy*, ossia:

$$f(t) = \tilde{\alpha} + \beta t + \tilde{\gamma}_1 d_{1t} + \tilde{\gamma}_2 d_{2t}, \tag{2.2}$$

si avrebbe la seguente matrice dei regressori:

$$\tilde{\mathbf{X}} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 3 & 0 & 0 \\ 1 & 4 & 1 & 0 \\ 1 & 5 & 0 & 1 \\ 1 & 6 & 0 & 0 \\ 1 & 7 & 1 & 0 \\ 1 & 8 & 0 & 1 \\ 1 & 9 & 0 & 0 \\ 1 & 10 & 1 & 0 \end{bmatrix}.$$

Le espressioni (2.1) e (2.2) sono parametrizzazioni diverse della medesima $f(t)$. Esse sono equivalenti in pratica (in particolare, in entrambe compare il termine $\beta t$), ma comportano (i) differenti chiavi di lettura dei coefficienti di regressione stagionali, e (ii) diversi modi di calcolo dei valori destagionalizzati. La costante $\tilde{\alpha}$ nel modello (2.2), ad esempio, comprende anche l'effetto stagionale del terzo quadrimestre, mentre i coefficienti $\tilde{\gamma}_1$ e $\tilde{\gamma}_2$ descrivono l'effetto differenziale del quadrimestre di pertinenza rispetto al terzo. Le relazioni tra i coefficienti dei due modelli sono date da:

$$\begin{aligned} \gamma_1^* &= \tilde{\alpha} + \tilde{\gamma}_1 \\ \gamma_2^* &= \tilde{\alpha} + \tilde{\gamma}_2 \\ \gamma_3^* &= \tilde{\alpha} \end{aligned}$$

ovvero

$$\begin{aligned} \tilde{\alpha} &= \gamma_3^* \\ \tilde{\gamma}_1 &= \gamma_1^* - \gamma_3^* \\ \tilde{\gamma}_2 &= \gamma_2^* - \gamma_3^* \end{aligned}.$$

Per calcolare la serie destagionalizzata secondo la parametrizzazione (2.2), osserviamo che

$$\bar{\gamma}^* = \frac{1}{3}\sum_{j=1}^{3} \gamma_j^* = \tilde{\alpha} + \frac{\tilde{\gamma}_1 + \tilde{\gamma}_2}{3}$$

e quindi i coefficienti ideali di stagionalità possono essere ottenuti nel modo seguente:

$$\begin{aligned} \hat{\gamma}_1 &= \gamma_1^* - \bar{\gamma}^* = \frac{2\tilde{\gamma}_1 - \tilde{\gamma}_2}{3} \\ \hat{\gamma}_2 &= \gamma_2^* - \bar{\gamma}^* = \frac{2\tilde{\gamma}_2 - \tilde{\gamma}_1}{3} \\ \hat{\gamma}_3 &= \gamma_3^* - \bar{\gamma}^* = -\frac{\tilde{\gamma}_1 + \tilde{\gamma}_2}{3} \end{aligned}.$$

**2.2** Una serie storica di dati mensili $y_{t,j}$, $t = 1, \ldots, n$, $j = 1, \ldots, 12$, è descritta dalla somma di un trend polinomiale quadratico e di una componente stagionale, modellata con variabili *dummy*. La stima dei coefficienti ideali di stagionalità, ottenuti a partire dai coefficienti di regressione delle variabili *dummy*, ed i valori osservati in un certo anno $t_0$ sono riportati nella tabella seguente:

| $j$ | $\hat{\gamma}_j$ | $y_{t_0,j}$ |
|---|---|---|
| 1 | -0,431 | 120,1 |
| 2 | 2,273 | 121,0 |
| 3 | ? | 133,4 |
| 4 | 0,804 | 116,5 |
| 5 | 7,474 | 128,5 |
| 6 | 6,587 | 123,6 |
| 7 | 6,252 | 125,9 |
| 8 | -51,630 | 60,2 |
| 9 | 8,594 | 121,7 |
| 10 | 10,525 | 132,6 |
| 11 | 6,063 | 123,1 |
| 12 | -7,998 | 105,0 |

Si descriva il modello sottostante e si ricavino $\hat{\gamma}_3$ ed i valori della serie destagionalizzata per i mesi di marzo, giugno, ottobre e dicembre dell'anno $t_0$.

□ La formulazione del modello con indice $t$ per l'anno e indice $j$ per il mese può essere ottenuta costruendo appropriati regressori per il trend polinomiale e per la stagionalità. La variabile di trend $T_{t,j}$ è descritta dalla relazione

$$T_{t,j} = 12(t-1) + j$$

mentre le variabili *dummy* $d_{jt}$ sono derivate come

$$d_{jt} = \begin{cases} 1 & \text{nel mese } j \\ 0 & \text{altrimenti} \end{cases}.$$

Con questi regressori il modello per $y_{t,j}$ può essere scritto come

$$y_{t,j} = \beta_1 T_{t,j} + \beta_2 T_{t,j}^2 + \sum_{j=1}^{12} \gamma_j d_{jt} + \varepsilon_{t,j}$$

per $t = 1, \ldots, n$ e $j = 1, \ldots, 12$. Si noti che per evitare la trappola delle *dummy* il termine di intercetta non è stato appositamente inserito.

Per definizione, i coefficienti ideali di stagionalità godono della proprietà

$$\sum_{j=1}^{12} \hat{\gamma}_j = 0.$$

Quindi, poiché

$$\sum_{\substack{j=1 \\ j \neq 3}}^{12} \hat{\gamma}_j = -11{,}487$$

allora $\hat{\gamma}_3 = 11{,}487$.

I valori destagionalizzati con il metodo della regressione si ottengono sottraendo ai valori originali i coefficienti ideali di stagionalità. Per l'anno $t_0$ questi sono calcolabili mediante l'espressione

$$\hat{y}^d_{t_0,j} = y_{t_0,j} - \hat{\gamma}_j, \quad j = 1, \ldots, 12.$$

Di seguito si riportano i valori destagionalizzati per i mesi richiesti:

$$\begin{array}{lll} \text{Marzo} & : & \hat{y}^d_{t_0,3} = 133{,}40 - 11{,}487 = 121{,}913 \\ \text{Giugno} & : & \hat{y}^d_{t_0,6} = 123{,}60 - 6{,}587 = 117{,}013 \\ \text{Ottobre} & : & \hat{y}^d_{t_0,10} = 132{,}60 - 10{,}525 = 122{,}075 \\ \text{Dicembre:} & & \hat{y}^d_{t_0,12} = 105{,}00 + 7{,}998 = 112{,}998 \end{array}$$

**2.3** Si è adattato alla serie storica trimestrale $y_t$, per $t = 1, \ldots, n$, il seguente modello additivo:

$$y_t = 14d_{1t} + 3d_{2t} - 5d_{3t} - 10d_{4t} + 0{,}5t + \varepsilon_t,$$

dove $Var\{\varepsilon_t\} = 4$ e $d_{1t}$, $d_{2,t}$, $d_{3t}$, $d_{4t}$ sono le variabili binarie indicatrici rispettivamente del primo, del secondo, del terzo e del quarto trimestre.

(a) A quanto ammonta la variazione della serie nell'arco di un anno dovuta alla componente tendenziale? E quella dovuta alla componente stagionale?

(b) Supponiamo che i primi quattro valori della serie, $t = 1{,}2{,}3{,}4$, siano relativi ai quattro trimestri di un certo anno e supponiamo che in tali periodi i valori osservati della serie siano $16$, $4$, $-2$, $-9$. Si calcolino i corrispondenti valori destagionalizzati.

(c) Si applichi alla serie l'operatore $(1 - B^4)$ ottenendo la serie

$$y^*_t = (1 - B^4)y_t.$$

Quanto vale mediamente la serie $y^*_t$? (Si ricordi che $E(\varepsilon_t) = 0$, $\forall t$).

□ (a) La variazione della serie trimestrale nell'arco di un anno è data dalla differenza $y_t - y_{t-4}$. Se la serie risulta dalla somma di trend, stagionalità e componente accidentale, tale differenza risulta pari a

$$\begin{aligned} y_t - y_{t-4} = (T_t - T_{t-4}) + (S_t - S_{t-4}) \\ +(\varepsilon_t - \varepsilon_{t-4}). \end{aligned} \tag{2.3}$$

Nel caso in esame la variazione della serie nell'arco di un anno dovuta alla componente tendenziale è pari al primo termine al secondo membro della (2.3):

$$T_t - T_{t-4} = 0{,}5t - 0{,}5(t-4) = 2.$$

La variazione della serie nell'arco di un anno dovuta alla componente stagionale è pari al secondo termine della (2.3). Essendo $S_t$ una funzione periodica di periodo $4$, tale termine è nullo.

(b) La serie destagionalizzata si ottiene sottraendo alla serie osservata la componente stagionale. Nel caso in questione, la componente stagionale

$$S_t = 14d_{1t} + 3d_{2t} - 5d_{3t} - 10d_{4t}$$

pur rappresentando i movimenti infra-annuali dovuti alla stagionalità, dà luogo ad oscillazioni la cui media è diversa da zero all'interno di uno stesso anno. Infatti,

$$\frac{14 + 3 - 5 - 10}{4} = 0{,}5.$$

Pertanto, se si utilizza $S_t$ per destagionalizzare è necessario ripristinare il livello della serie aggiungendo alla serie destagionalizzata la media della funzione $S_t$:

$$\hat{y}_t^d = y_t - S_t + 0{,}5.$$

Avremo quindi:

$$\begin{array}{ll} t=1: & \hat{y}_1^d = 16 - 14 + 0{,}5 = 2{,}5 \\ t=2: & \hat{y}_2^d = 4 - 3 + 0{,}5 = 1{,}5 \\ t=3: & \hat{y}_3^d = -2 + 5 + 0{,}5 = 3{,}5 \\ t=4: & \hat{y}_4^d = -9 + 10 + 0{,}5 = 1{,}5. \end{array}$$

In alternativa si possono calcolare preliminarmente i coefficienti ideali di stagionalità, sottraendo ai coefficienti grezzi il valore $0{,}5$. Per essi vale infatti la proprietà

$$\frac{13{,}5 + 2{,}5 - 5{,}5 - 10{,}5}{4} = 0.$$

Gli stessi valori destagionalizzati possono quindi essere calcolati sottraendo ai valori $y_t$ i coefficienti ideali.

(c) L'espressione per la serie $y_t^*$ coincide con quella fornita nella (2.3). Allora si era applicato implicitamente l'operatore differenza stagionale $(1 - B^4)$. Calcolando ora il valore atteso dei singoli termini otteniamo

$$\begin{aligned} E(y_t^*) = E(T_t - T_{t-4}) + E(S_t - S_{t-4}) \\ + E(\varepsilon_t - \varepsilon_{t-4}). \end{aligned}$$

Il secondo termine si annulla per quanto detto al punto (a). Pertanto il valore medio di $y_t^*$ è pari a

$$E(y_t^*) = T_t - T_{t-4} - E(\varepsilon_t - \varepsilon_{t-4}).$$

Considerato il fatto che il trend è di tipo deterministico e quindi non è una variabile casuale, tenendo conto che il valore atteso di una somma è pari alla somma dei valori attesi e che il valore atteso della componente accidentale è nullo, si ottiene che in media la variazione annuale della serie $y_t^*$ è pari a

$$E(y_t^*) = T_t - T_{t-4} = 0{,}5t - 0{,}5(t-4) = 2.$$

**2.4** Si è adattato alla serie storica semestrale $y_t$, per $t = 1, \ldots, T$, il seguente modello additivo:

$$y_t = 14 d_{1t} + 3 d_{2t} + 0{,}5t + \varepsilon_t, \quad Var\{\varepsilon_t\} = 4$$

con la variabile $d_{1t}$ uguale ad uno se il periodo $t$ cade nel primo semestre ed uguale a zero altrimenti e, analogamente, la variabile $d_{2t}$ uguale ad uno se il periodo $t$ cade nel secondo semestre ed uguale a zero altrimenti.

(a) A quanto ammonta mediamente la differenza dovuta alla componente stagionale tra il valore $y_t$ in un primo semestre e il valore di $y_t$ in un secondo semestre?

(b) Supponiamo che i primi due valori della serie ($t = 1{,}2$) siano relativi al primo e al secondo semestre di un certo anno e supponiamo che in tali periodi i valori osservati siano, rispettivamente, 16 e 2. Si calcolino i corrispondenti valori destagionalizzati.

(c) Supponiamo che i valori osservati arrivino fino al periodo $T = 20$ e che tale periodo cada in un secondo semestre. Calcolare i valori previsti di $y_t$ per i periodi $t = 21{,}22$.

□ (a) Indichiamo con $S_t$ la componente stagionale

$$S_t = 14d_{1t} + 3d_{2t}$$

e calcoliamo il valore atteso della differenza fra $S_{t_0}$ e $S_{t_0-1}$, dove con $t_0$ indichiamo un generico primo semestre:

$$E(S_{t_0} - S_{t_0-1}) = E(S_{t_0}) - E(S_{t_0-1}) = \\ 14d_{1t_0} - 3d_{2t_0-1} = 11.$$

L'ultimo passaggio è stato ottenuto sulla base della natura deterministica delle variabili *dummy* $d_{1t}$ e $d_{2t}$: il valore atteso di una componente deterministica è infatti la componente stessa. Si noti, inoltre, che il risultato non cambia se si considerano dei semestri non contigui tra loro, poiché la struttura dei coefficienti stagionali è fissata nel tempo.

(b) Come ormai noto, il calcolo dei valori destagionalizzati sottintende l'utilizzo dei coefficienti ideali di stagionalità. Anche in questo caso il modello presenta coefficienti grezzi, in quanto la loro somma è differente da zero. Nell'ipotesi di un modello additivo, deriviamo quindi i coefficienti ideali di stagionalità sottraendo la media $\bar{\gamma}^*$ ai coefficienti grezzi

$$\hat{\gamma}_1 = \gamma_1^* - \bar{\gamma}^* = 14 - 8{,}5 = 5{,}5 \\ \hat{\gamma}_2 = \gamma_2^* - \bar{\gamma}^* = 3 - 8{,}5 = -5{,}5.$$

Ovviamente, nel caso di un modello semestrale i coefficienti ideali devono essere uguali e di segno opposto affinché la loro somma risulti nulla. I valori destagionalizzati per le osservazioni fornite sono dunque

$$y_1^d = y_1 - \hat{\gamma}_1 = 16 - 5{,}5 = 10{,}5 \\ y_2^d = y_2 - \hat{\gamma}_2 = 2 - (-5{,}5) = 7{,}5.$$

(c) Infine, l'esercizio richiede il calcolo delle previsioni per la serie $y_t$ per i due semestri $t = 21, 22$, tempi riferiti rispettivamente ad un primo e ad un secondo semestre. Applicando l'operatore valore atteso condizionato[1] all'informazione posseduta al tempo 20, si ottengono le seguenti previsioni per i due semestri

$$\hat{y}_{21} = E(Y_{21}|I_{20}) = E(14d_{1,21} + 3d_{2,21} + 0{,}5(21) + \varepsilon_{21}) = \\ 14 + 10{,}5 = 24{,}5 \\ \hat{y}_{22} = E(Y_{22}|I_{20}) = E(14d_{1,22} + 3d_{2,22} + 0{,}5(22) + \varepsilon_{22}) = \\ 3 + 11 = 14$$

per i quali si è assunto $E(\varepsilon_{21}) = E(\varepsilon_{22}) = 0$.

1. L'operatore valore atteso condizionato sarà più ampiamente trattato nel capitolo 7, riguardante le previsioni di un modello ARIMA.

**2.5** La serie storica $\{y_t\}_{t=1}^{21}$ si riferisce ad una variabile economica osservata trimestralmente nel periodo 1997.2-2002.2. Per essa è stato stimato il seguente modello:

$$y_t = 5d_{1t} - 3{,}5d_{2t} + 2d_{3t} + 6d_{4t} + 0{,}8t + 0{,}02t^2 + \varepsilon_t, \;\; \varepsilon_t \sim WN(0{,}4)$$

dove $d_{1t}$, $d_{2t}$, $d_{3t}$ e $d_{4t}$ sono variabili *dummy* relative rispettivamente al primo, al secondo, al terzo ed al quarto trimestre.

(a) Calcolare le previsioni per il periodo 2002.3-2003.2.

(b) Sono stati osservati i valori $y_1 = 5{,}5$, $y_2 = 2{,}8$, $y_3 = 13$, $y_4 = 21{,}5$. Fornire i corrispondenti valori destagionalizzati per queste quattro osservazioni.

(c) Riscrivere il modello inserendo la costante tra i regressori.

□ (a) Le previsioni per il periodo 2002.3-2003.2 si ottengono applicando i coefficienti del modello ai valori assunti dalle variabili nei tempi corrispondenti ai quattro trimestri, ovvero per $t = 22, 23, 24, 25$:

$$\begin{array}{ll} t = 22: & \hat{y}_{22} = 2 + 0{,}8(22) + 0{,}02(22^2) = 29{,}28 \\ t = 23: & \hat{y}_{23} = 6 + 0{,}8(23) + 0{,}02(23^2) = 34{,}98 \\ t = 24: & \hat{y}_{24} = 5 + 0{,}8(24) + 0{,}02(24^2) = 35{,}72 \\ t = 25: & \hat{y}_{25} = -3{,}5 + 0{,}8(25) + 0{,}02(25^2) = 29{,}00. \end{array}$$

Come si nota, le variabili *dummy* stagionali producono effetti solo nei trimestri di pertinenza, attraverso i corrispettivi coefficienti di regressione. I regressori $t$ e $t^2$, che rappresentano l'evoluzione di fondo della variabile, influiscono invece in ciascun passo di previsione. Si noti inoltre come estendendo il periodo di previsione la stagionalità prevista sia sempre la stessa, mentre la tendenza di fondo cresce in progressione quadratica col procedere del tempo, ossia all'aumentare dell'indice temporale $t$.

(b) La serie destagionalizzata si calcola sulla base dei coefficienti ideali di stagionalità. I coefficienti riportati, $\gamma_j^*$, $j = 1, \ldots, 4$, sono in realtà grezzi, in quanto per essi non vale la proprietà

$$\sum_{j=1}^{4} \gamma_j^* = 0.$$

La somma dei coefficienti stagionali è infatti pari a 9,5. Sottraendo ai coefficienti grezzi la media ($\bar{\gamma}^* = \frac{9{,}5}{4} = 2{,}375$) si ottengono i coefficienti ideali

$$\begin{aligned} \hat{\gamma}_1 &= \gamma_1^* - \bar{\gamma}^* = 5 - 2{,}375 = 2{,}625 \\ \hat{\gamma}_2 &= \gamma_2^* - \bar{\gamma}^* = -3{,}5 - 2{,}375 = -5{,}875 \\ \hat{\gamma}_3 &= \gamma_3^* - \bar{\gamma}^* = 2 - 2{,}375 = -0{,}375 \\ \hat{\gamma}_4 &= \gamma_4^* - \bar{\gamma}^* = 6 - 2{,}375 = 3{,}625. \end{aligned}$$

Con questi coefficienti è possibile derivare i valori destagionalizzati richiesti:

$$\begin{array}{ll} t = 1: & \hat{y}_1^d = 5{,}5 - (-5{,}875) = 11{,}375 \\ t = 2: & \hat{y}_2^d = 2{,}8 - (-0{,}375) = 3{,}175 \\ t = 3: & \hat{y}_3^d = 13 - 3{,}625 = 9{,}375 \\ t = 4: & \hat{y}_4^d = 21{,}5 - 2{,}625 = 18{,}875. \end{array}$$

Si noti che per tali valori vale la relazione

$$\sum_{t=1}^{4} y_t = \sum_{t=1}^{4} \hat{y}_t^d = 42{,}8,$$

una proprietà soddisfatta dall'utilizzo dei coefficienti ideali di stagionalità.

(c) Il modello di regressione con variabili *dummy* stagionali può essere riscritto inserendo un termine costante ed eliminando una di queste a scelta, evitando così il problema della collinearità tra i regressori più volte ricordata nei precedenti esercizi. Scegliendo di eliminare $d_{1t}$ il modello assume la seguente forma:

$$y_t = 5 - 8{,}5d_{2t} - 3d_{3t} + d_{4t} + 0{,}8t + 0{,}02t^2 + \varepsilon_t.$$

In questo caso la costante $\tilde{\alpha} = 5$ ingloba anche l'effetto stagionale (grezzo) relativo al primo trimestre, mentre i coefficienti $\tilde{\gamma}_2, \tilde{\gamma}_3$ e $\tilde{\gamma}_4$, relativi a $d_{2t}$, $d_{3t}$ e $d_{4t}$, rispettivamente, sono gli effetti stagionali differenziali rispetto al primo trimestre, ovvero

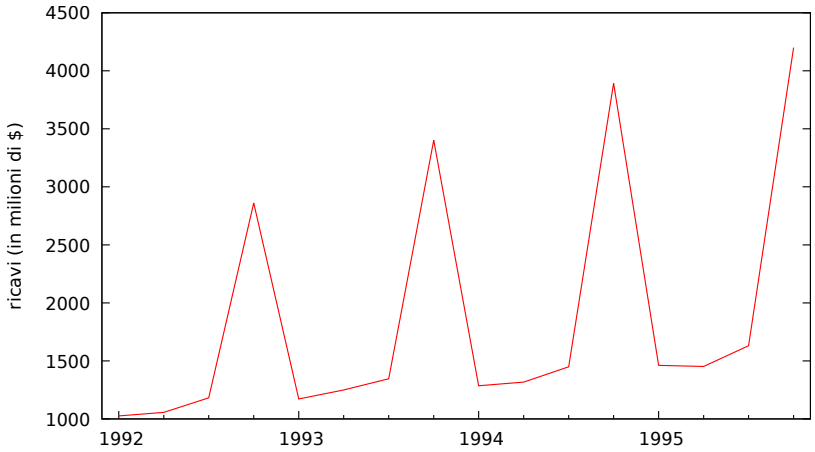
$$\begin{aligned}
\tilde{\alpha} &= \gamma_1^* = 5 \\
\tilde{\gamma}_2 &= \gamma_2^* - \gamma_1^* = -3{,}5 - 5 = -8{,}5 \\
\tilde{\gamma}_3 &= \gamma_3^* - \gamma_1^* = 2 - 5 = -3 \\
\tilde{\gamma}_4 &= \gamma_4^* - \gamma_1^* = 6 - 5 = 1.
\end{aligned}$$

Tuttavia, come mostrato nell'esercizio 2.1, per il calcolo dei valori destagionalizzati è sempre meglio riferirsi alla formulazione del modello senza intercetta e con le quattro variabili *dummy* stagionali, perché questa formulazione consente di ottenere direttamente gli effetti di ciascuna stagione sulla serie.

**2.6** Si consideri la seguente serie storica trimestrale, $y_t$, relativa ai ricavi delle vendite di giocattoli in USA dal primo trimestre del 1992 al quarto trimestre del 1995 (Shmueli e Lichtendahl Jr., 2016, *Practical Time Series Forecasting with R: A Hands-On Guide*, 2nd ed.):

| **t** | 92:1 | 92:2 | 92:3 | 92:4 | 93:1 | 93:2 | 93:3 | 93:4 | 94:1 | 94:2 | 94:3 | 94:4 | 95:1 | 95:2 | 95:3 | 95:4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $y_t$ | 1026 | 1056 | 1182 | 2861 | 1172 | 1249 | 1346 | 3402 | 1286 | 1317 | 1449 | 3893 | 1462 | 1452 | 1631 | 4200 |

il cui grafico è riportato nella seguente figura:




(a) Dire, giustificando la risposta, quale fra i seguenti modelli di regressione dovrebbe adattarsi meglio a questi dati:

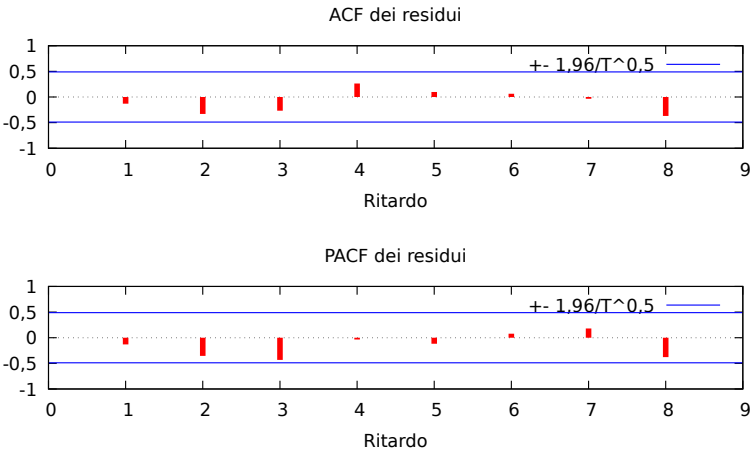
- Modello con trend lineare
- Modello con trend lineare e dummy stagionali
- Modello con trend quadratico
- Modello con trend quadratico e dummy stagionali

(b) Di seguito viene riportato l'output relativo al modello con trend lineare e variabili dummy e i grafici delle autocorrelazioni dei residui del modello.

```
Modello 1: OLS, usando le osservazioni 1992:1-1995:4 (T = 16)
Variabile dipendente: y

coefficiente   errore std.   rapporto t   p-value
-------------------------------------------------------------
time          53,9750       10,4983        5,141     0,0003   ***
dq1          858,675       119,237         7,201     1,75e-05 ***
dq2          836,700       125,979         6,642     3,65e-05 ***
dq3          916,225       133,208         6,878     2,66e-05 ***
dq4         3049,25        140,849        21,65      2,28e-10 ***

Media var. dipendente   1874,000   SQM var. dipendente     1066,721
Somma quadr. residui    387951,8   E.S. della regressione  187,7987
R-quadro                0,977271   R-quadro corretto       0,969006
F(4, 11)                118,2394   P-value(F)              5,83e-09
Log-verosimiglianza    -103,4714   Criterio di Akaike      216,9428
Criterio di Schwarz     220,8057   Hannan-Quinn            217,1406
rho                    -0,165786   Durbin-Watson           2,015442
Note: SQM = scarto quadratico medio; E.S. = errore standard
```




1. Dire, giustificando la risposta, se il modello si adatta bene ai dati.
2. Calcolare i valori destagionalizzati relativi all'anno 1995.
3. Prevedere i valori per l'anno 1996.

□ Osservando i grafici appare chiaro quanto segue:

(a) la serie contiene un trend lineare ed una componente stagionale, sicché il modello più adatto è quello con trend lineare e dummy stagionali.

(b.1) Tutti i coefficienti stimati risultano significativamente diversi da zero. Il valore della statistica $R^2$ è molto alto (circa 0,98), ad indicare che le variabili indipendenti del modello spiegano gran parte della variabilità della serie delle vendite di giocattoli. Inoltre, il grafico dei residui sembra mostrare che questi sono non correlati, visto che le autocorrelazioni empiriche sono tutte interne alle bande di confidenza (bisogna comunque osservare che abbiamo solo 16 osservazioni, e quindi le bande di confidenza sono molto larghe). Tutto questo ci fa ritenere che il modello adottato sia un buon modello.

(b.2) Gli indici grezzi di stagionalità corrispondono ai quattro coefficienti delle variabili dummy e sono pertanto:

$$\gamma_1^* = 858{,}675 \quad \gamma_2^* = 836{,}700 \quad \gamma_3^* = 916{,}225 \quad \gamma_4^* = 3049{,}25$$

Evidentemente questi coefficienti non sono quelli ideali, poiché sono tutti positivi ed hanno media pari a $\bar{\gamma}^* = 1415{,}2125$. I coefficienti ideali di stagionalità si ricavano sottraendo tale media ai coefficienti grezzi, ossia:

$$\hat{\gamma}_1 = -556{,}5375 \quad \hat{\gamma}_2 = -578{,}5125 \quad \hat{\gamma}_3 = -498{,}9875 \quad \hat{\gamma}_4 = 1634{,}0375$$

I valori destagionalizzati si trovano sottraendo ai dati originali i pertinenti coefficienti ideali di stagionalità:

$$y^d_{95:1} = 1462 + 556{,}5375 = 2018{,}5375 \quad y^d_{95:2} = 1452 + 578{,}5125 = 2030{,}5125$$

$$y^d_{95:3} = 1631 + 498{,}9875 = 2129{,}9875 \quad y^d_{95:4} = 4200 - 1634{,}0375 = 2565{,}9625$$

(b.3) Essendo $n = 16$, si ha:

$$\hat{y}_{96:1} = 53{,}975 \cdot 17 + 858{,}675 = 1776{,}25 \quad \hat{y}_{96:2} = 53{,}975 \cdot 18 + 836{,}700 = 1808{,}25$$

$$\hat{y}_{96:3} = 53{,}975 \cdot 19 + 916{,}225 = 1941{,}75 \quad \hat{y}_{96:4} = 53{,}975 \cdot 20 + 3049{,}25 = 4128{,}75.$$

## Scelta del grado del polinomio

**2.7** Per stimare il trend della serie storica $\{y_t\}_{t=1}^{20}$ si prendono in considerazione le seguenti relazioni stimate con il metodo dei minimi quadrati ordinari (tra parentesi compaiono le stime delle deviazioni standard):

A: $\hat{y}_t = \underset{(18)}{1.077} + \underset{(5,40)}{52,16}t - \underset{(0,44)}{2,27}t^2 + \underset{(0,01)}{0,04}t^3$

B: $\hat{y}_t = \underset{(23)}{1.046} + \underset{(10,94)}{71,21}t - \underset{(1,55)}{5,20}t^2 + \underset{(0,08)}{0,20}t^3 - \underset{(0,0035)}{0,0025}t^4$

Inoltre, per il modello A si ha che il coefficiente di determinazione $R_A^2$ è pari a $0{,}9851$, mentre per il modello B si ha $R_B^2 = 0{,}9859$. Indicare quale dei due modelli è preferibile e motivare brevemente la scelta.

□ I modelli stimati per il trend sono di tipo polinomiale:

$$y_t = \sum_{j=0}^{p} \beta_j t^j + \varepsilon_t.$$

Il grado del polinomio per il modello A è $p = 3$, mentre nel modello B è stato impiegato un polinomio del quarto ordine ($p = 4$). La scelta tra i due modelli, e quindi la scelta del grado del polinomio, può essere condotta attraverso il confronto delle statistiche fornite dall'esercizio. Si intuisce che la relazione A è migliore della relazione B. Tale affermazione può essere motivata sulla base della significatività statistica dei coefficienti e del coefficiente di determinazione corretto per il numero di regressori utilizzati.

L'esercizio riporta i valori dei coefficienti con le relative deviazioni standard. La significatività statistica della stima di un coefficiente $\beta$ si valuta attraverso il test $t_{\hat{\beta}}$

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}}$$

che, in ipotesi di normalità dei disturbi $\varepsilon_t$, segue una distribuzione standard $t$ di Student con $n - k$ gradi di libertà, dove $n$ è il numero di osservazioni utilizzate e $k$ il numero di regressori inclusi nella stima. La tabella seguente riporta le statistiche $t$ per i coefficienti dei due modelli.

| Modello | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\beta_3$ | $\beta_4$ |
|---|---|---|---|---|---|
| A | 59,83 | 9,66 | $-5{,}16$ | 4,00 | - |
| B | 45,48 | 6,51 | $-3{,}35$ | 2,50 | -0,71 |

I coefficienti fino al terzo ordine sono in entrambi i casi significativi. Il coefficiente del regressore $t^4$ incluso nel modello B ha invece una statistica $t$ il cui valore assoluto è inferiore a 2. Pertanto, esso non è significativamente diverso da zero ad un livello di probabilità del 5% e può essere escluso dalla specifica finale del modello.

Un altro criterio che può aiutare nella scelta del grado del polinomio è il calcolo del coefficiente di determinazione corretto $\bar{R}^2$. È noto infatti che il coefficiente $R^2$ è una funzione non decrescente del numero dei regressori, per cui nel nostro caso è naturale avere $R_B^2 > R_A^2$. Il coefficiente $\bar{R}^2$ tiene conto del numero dei regressori attraverso la seguente formulazione

$$\bar{R}^2 = 1 - \frac{n-1}{n-k}(1 - R^2).$$

Questa misura penalizza la perdita di gradi di libertà che deriva dall'aggiungere variabili esplicative nella specifizazione del modello: si può dimostrare che $\bar{R}^2$ *aumenta* (*diminuisce*) quando una variabile aggiunta al modello ha una statistica $t$ *superiore* (*inferiore*) ad 1 in valore assoluto (Greene, 2003, p. 35). Nel nostro caso otteniamo i valori

$$\bar{R}^2_A = 1 - \frac{19}{16}(1 - 0{,}9851) = 0{,}9823$$
$$\bar{R}^2_B = 1 - \frac{19}{15}(1 - 0{,}9859) = 0{,}9821$$

quindi il coefficiente corretto per il modello B è in effetti diminuito, considerata la significatività statistica del coefficiente di regressione per $t^4$ pari a -0,71. Prendendo come misura di confronto il coefficiente di determinazione corretto, vale quindi la relazione $\bar{R}^2_A > \bar{R}^2_B$. Questo criterio fornisce quindi un ulteriore elemento a supporto per la scelta del modello A.

**2.8** Si consideri la seguente serie storica dei ricavi di una ditta di commercio all'ingrosso

| *Anno* | *Ricavi* |
|---|---|
| 2003 | 132 |
| 2004 | 247 |
| 2005 | 465 |
| 2006 | 594 |
| 2007 | 650 |
| 2008 | 803 |
| 2009 | 925 |

Supponendo che il fenomeno sia combinazione di una componente di trend e di una componente accidentale,

(a) scrivere il modello valido per un trend esponenziale semplice, completo di ipotesi sulla componente accidentale, e stimare i parametri;

(b) calcolare la previsione per le vendite del 2010.

□ (a) Un trend esponenziale è descritto dalla funzione del tempo

$$f(t) = \alpha \mathrm{e}^{\beta t}, \quad \alpha > 0.$$

L'andamento della curva è crescente per $\beta > 0$, mentre se $\beta < 0$ la curva è decrescente. Il trend esponenziale ha il vantaggio di poter essere trasformato in una funzione lineare del tempo attraverso la trasformata logaritmica. Infatti si ha

$$\log f(t) = \log \alpha + \beta t.$$

Se il modello è di tipo moltiplicativo, i parametri $\alpha^* = \log \alpha$ e $\beta$ possono essere stimati attraverso il metodo dei minimi quadrati lineari. Infatti, moltiplicando un termine di errore $\varepsilon_t$ si ha

$$y_t = \alpha \mathrm{e}^{\beta t} \cdot \varepsilon_t$$

da cui

$$\begin{aligned} \log(y_t) &= \log \alpha + \beta t + \log \varepsilon_t \\ &= \alpha^* + \beta t + \varepsilon^*_t. \end{aligned} \tag{2.4}$$

Affinché si possa impiegare il metodo di stima lineare bisogna riferire le assunzioni del modello di regressione a $\varepsilon_t^*$, e quindi a $\log \varepsilon_t$. Questo implica l'ipotesi di un disturbo $\varepsilon_t$ log-normale nel modello originario. Si può dimostrare che questa ipotesi comporta una distorsione nei valori stimati $\hat{y}_t$, anche in presenza di perfetta conoscenza dei parametri (si veda Di Fonzo e Lisi, 2005, p. 49).

Il modello linearizzato nella (2.4) può dunque essere stimato mediante il metodo dei minimi quadrati. La tabella seguente riporta i calcoli necessari per la stima del coefficiente $\beta$, dove si è indicato $y_t^* = \log y_t$ , $\tilde{x}_t = t - E(t)$ e $\tilde{y}_t^* = y_t^* - E(y_t^*)$. Per semplificare i calcoli si è utilizzato l'asse dei tempi con origine $t = 1$.

| Anno | $t$ | $y_t$ | $y_t^*$ | $\tilde{x}_t$ | $\tilde{y}_t^*$ | $\tilde{x}_t\tilde{y}_t^*$ | $\tilde{x}_t^2$ |
|---|---|---|---|---|---|---|---|
| 2003 | 1 | 132 | 4,8828 | -3 | -1,2481 | 3,7443 | 9 |
| 2004 | 2 | 247 | 5,5094 | -2 | -0,6215 | 1,2430 | 4 |
| 2005 | 3 | 465 | 6,1420 | -1 | 0,0111 | -0,0111 | 1 |
| 2006 | 4 | 594 | 6,3869 | 0 | 0,2560 | 0,0000 | 0 |
| 2007 | 5 | 650 | 6,4770 | 1 | 0,3461 | 0,3461 | 1 |
| 2008 | 6 | 803 | 6,6884 | 2 | 0,5575 | 1,1149 | 4 |
| 2009 | 7 | 925 | 6,8298 | 3 | 0,6989 | 2,0967 | 9 |
| $E(\cdot)$ | 4 | | 6,1309 | | | 1,2191 | 4 |

La stima del coefficiente $\beta$ secondo i minimi quadrati è quindi data da

$$\hat{\beta} = \frac{E(\tilde{x}_t\tilde{y}_t)}{E(\tilde{x}_t^2)} = \frac{1{,}2191}{4} = 0{,}3048.$$

La stima di $\alpha^*$ è invece fornita dalla relazione

$$\hat{\alpha}^* = E(\tilde{y}_t) - \hat{\beta}E(\tilde{x}_t) = 6{,}1309 - 0{,}3048(4) = 4{,}9118.$$

Una stima di $\alpha$ può dunque essere ottenuta dalla trasformazione inversa

$$\hat{\alpha} = \mathrm{e}^{\hat{\alpha}^*} = \mathrm{e}^{4{,}9118} = 135{,}8796.$$

Si noti che il valore $\hat{\alpha}$ non è stato ottenuto come funzione lineare delle osservazioni $y_t$, per cui esso non è una stima ai minimi quadrati lineari del parametro $\alpha$.

(b) Sulla base del modello stimato

$$y_t = 135{,}8796\mathrm{e}^{0{,}3048t}\varepsilon_t$$

si può ora compiere una previsione dei ricavi per l'anno 2010. La previsione per $t = 8$ indica un ricavo atteso pari a

$$E(y_8) = 135{,}8796\mathrm{e}^{0{,}3048\cdot 8} = 1556{,}2123,$$

di quasi il 70% superiore a quelli registrati nel 2009. Si ricordi tuttavia che il trend esponenziale semplice cresce indefinitamente al crescere di $t$; è bene dunque fare attenzione all'utilizzo di un modello di questo tipo se l'obiettivo è la previsione di valori futuri. In alternativa possono essere impiegate curve non lineari caratterizzate dalla presenza di un asintoto superiore costante (ad esempio la curva esponenziale modificata o la curva logistica).

### Trend logistico

**2.9** Alla serie storica annuale $y_t$ ($t = 1$ corrisponde all'anno 1980) riportata nella tabella seguente si vuole adattare un trend logistico.

| $t$ | $y_t$ | $t$ | $y_t$ | $t$ | $y_t$ | $t$ | $y_t$ |
|---|---|---|---|---|---|---|---|
| 1 | 0,116 | 9 | 0,620 | 17 | 1,593 | 25 | 1,947 |
| 2 | 0,166 | 10 | 0,717 | 18 | 1,796 | 26 | 1,943 |
| 3 | 0,230 | 11 | 0,930 | 19 | 1,713 | 27 | 2,050 |
| 4 | 0,221 | 12 | 1,144 | 20 | 1,786 | 28 | 2,025 |
| 5 | 0,347 | 13 | 1,236 | 21 | 1,867 | 29 | 1,943 |
| 6 | 0,388 | 14 | 1,361 | 22 | 1,993 | 30 | 2,105 |
| 7 | 0,468 | 15 | 1,515 | 23 | 1,824 | 31 | 2,024 |
| 8 | 0,618 | 16 | 1,535 | 24 | 1,888 | | |

(a) Ottenere le stime dei parametri con il metodo delle somme.

(b) Utilizzando Gretl, stimare i parametri del modello mediante i minimi quadrati non lineari usando come valori iniziali (i) la terna $(0{,}1; 0{,}1; 0{,}1)$ e (ii) le stime ottenute al punto precedente.

(c) Confrontare i risultati con quelli ottenuti adattando alla serie storica un trend polinomiale opportunamente scelto.

□ La funzione

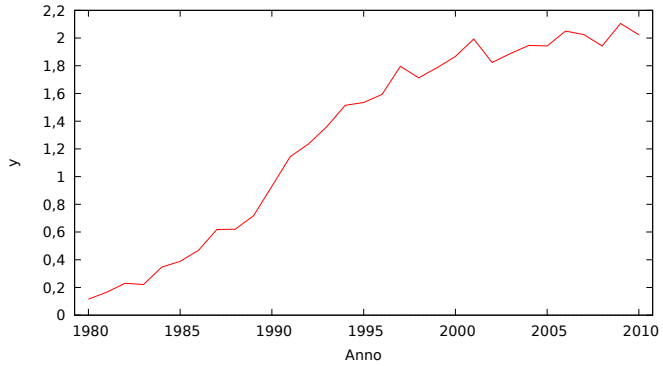
$$f(t) = \frac{\alpha}{1 + \beta \mathrm{e}^{-kt}} \tag{2.5}$$

prende il nome di curva *logistica*. Essa è una funzione non lineare dei suoi parametri $\alpha$, $\beta$ e $k$. Questo tipo di funzione è utile quando l'evoluzione temporale di un fenomeno manifesta una fase di crescita molto accelerata all'inizio del periodo di osservazione, seguita poi da una fase di stasi che persiste lungo l'asse temporale. Un'applicazione tipica nel campo economico è l'andamento delle vendite nel corso del ciclo di vita di un prodotto ad elevato contenuto tecnologico: se il prodotto è di successo, non appena viene introdotto sul mercato si avrà un rapido sviluppo delle vendite, a cui seguirà una inevitabile decelerazione nella fase di maturità del prodotto e ad una stabilizzazione nel lungo periodo.

La Figura 2.1 mostra il grafico della serie presentata in tabella, che sembra possedere le caratteristiche tipiche di un trend logistico. La stima dei parametri della (2.5) non può essere effettuata secondo il metodo dei minimi quadrati lineari, poiché le equazioni normali non sono funzioni lineari delle osservazioni. Esistono due metodi per la stima dei parametri di un trend logistico: il metodo delle somme e quello dei minimi quadrati non lineari. Il primo metodo ha il vantaggio di essere facilmente applicabile, tuttavia le stime che ne derivano non godono di particolari proprietà statistiche. Stime più soddisfacenti in tal senso possono essere derivate dalla soluzione offerta dai minimi quadrati non lineari: il sistema di equazioni che ne deriva non è facilmente risolvibile per via analitica, pertanto nella maggior parte dei casi si deve ricorrere a procedure iterative di ottimizzazione (come verrà mostrato nel seguito).

(a) Una stima approssimata dei coefficienti $\alpha$, $\beta$ e $k$ può essere ottenuta mediante il cosiddetto metodo delle somme. Come primo passo si suddivide l'intero periodo campionario in tre sottoperiodi di pari lunghezza $m$. Dal momento che abbiamo 31 osservazioni, trascuriamo per semplicità la prima osservazione in modo da avere 3 sotto-campioni da 10 elementi ciascuno. Si pone quindi la condizione che la somma dei valori osservati sia pari alla somma dei valori calcolati tramite la funzione

Figura 2.1
Grafico della serie $\{y_t\}_{t=1}^{31}$




(2.5), ovvero

$$S_1 = \sum_{i=2}^{11} y_t = \sum_{i=2}^{11} f(t)$$
$$S_2 = \sum_{i=12}^{21} y_t = \sum_{i=12}^{21} f(t)$$
$$S_3 = \sum_{i=22}^{31} y_t = \sum_{i=22}^{31} f(t).$$

Per semplificare la procedura di stima, si utilizza in realtà il reciproco della funzione $f(t)$

$$\frac{1}{f(t)} = \frac{1+\beta \mathrm{e}^{-kt}}{\alpha} = \frac{1}{\alpha} + \frac{\beta}{\alpha}\mathrm{e}^{-kt}.$$

Ridefiniamo le variabili in modo da ottenere la curva esponenziale modificata

(2.6) $$f^*(t) = \beta_0 + \beta_1 \beta_2^t,$$

con $\beta_0 = \dfrac{1}{\alpha}$, $\beta_1 = \dfrac{\beta}{\alpha}$ e $\beta_2 = \mathrm{e}^{-k}$. Di Fonzo e Lisi (2005, p. 69-70) dimostrano che le stime per questi parametri sono date da

$$\hat{\beta}_2 = \sqrt[m]{\frac{S_3 - S_2}{S_2 - S_1}}$$
$$\hat{\beta}_1 = (S_2 - S_1)\frac{\hat{\beta}_2 - 1}{\hat{\beta}_2(\hat{\beta}_2^m - 1)^2}$$
$$\hat{\beta}_0 = \frac{1}{m}\left[S_1 - \hat{\beta}_1\hat{\beta}_2\left(\frac{\hat{\beta}_2^m - 1}{\hat{\beta}_2 - 1}\right)\right].$$

Dopo aver trovato i seguenti valori per $S_1$, $S_2$ e $S_3$

$$S_1 = 28{,}194 \quad S_2 = 6{,}593 \quad S_3 = 5.073$$

utilizzando come termini delle sommatorie i valori reciproci dei valori $y_t$, le stime per i parametri del modello (2.6) mediante il metodo delle somme sono

$$\hat{\beta}_0 = 0{,}496 \quad \hat{\beta}_1 = 7{,}597 \quad \hat{\beta}_2 = 0{,}767.$$

Dalle stime $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ si risale immediatamente alla stima dei parametri della curva logistica applicando le trasformazioni inverse:

$$\begin{aligned}
\hat{\alpha} &= \frac{1}{\hat{\beta}_0} = \frac{1}{0{,}496} = 2{,}017 \\
\hat{\beta} &= \hat{\beta}_1 \hat{\alpha} = 7{,}597 \cdot 2{,}017 = 15{,}322 \\
\hat{k} &= -\log(\hat{\beta}_2) = -\log(0{,}767) = 0{,}265.
\end{aligned}$$

(b) Un approccio con migliori proprietà statistiche per la stima dei parametri della curva logistica (2.5) è senza dubbio quello dei minimi quadrati non lineari. L'assenza di linearità nei parametri richiede procedure di calcolo numerico per la ricerca del minimo della funzione obiettivo (la somma del quadrato dei residui): per tale compito è pertanto necessario ricorrere all'utilizzo di un software econometrico. Gretl permette di calcolare le stime dei parametri col metodo dei minimi quadrati non lineari usando una versione modificata dell'algoritmo di Levenberg-Marquardt. È necessario che l'utente specifichi il modello e che i parametri di tale modello vengano dichiarati ed inizializzati (usando il comando `genr`). In via opzionale, per velocizzare i calcoli ed ottenere risultati più affidabili, è possibile fornire al programma le derivate parziali della funzione rispetto ai parametri. Per una curva esponenziale modificata si ha:

$$\begin{aligned}
\frac{\partial f(t)}{\partial \alpha} &= \frac{1}{1+\beta \mathrm{e}^{-kt}} \\
\frac{\partial f(t)}{\partial \beta} &= -\frac{\alpha \mathrm{e}^{-kt}}{(1+\beta \mathrm{e}^{-kt})^2} \\
\frac{\partial f(t)}{\partial k} &= \frac{\alpha \beta \mathrm{e}^{-kt} t}{(1+\beta \mathrm{e}^{-kt})^2}
\end{aligned}$$

Supponendo di aver caricato i dati della serie storica nella variabile `y` di Gretl, di seguito vengono riportati i comandi per stimare il modello (da inserire nella finestra che appare in Gretl quando si apre in sequenza Modello/NLS - Minimi quadrati non lineari...).

```
genr alfa = 0.1
genr beta = 0.1
genr kappa = 0.1
y = alfa/(1+beta*exp(-kappa*time))
deriv alfa = 1/(1+beta*exp(-kappa*time))
deriv beta = (-alfa*exp(-kappa*time))/(1+beta*exp(-kappa*time))^2
deriv kappa = (alfa*beta*exp(-kappa*time)*time)/(1+beta*exp(-kappa*time))^2
```

Nella Tabella 2.1 sono riportate le stime dei coefficienti ottenute da Gretl con alcune diagnostiche sul modello.

Se invece non si forniscono le derivate analitiche, le ultime tre istruzioni vanno sostituite come segue

Tabella 2.1
Stima del trend logistico con i minimi quadrati non lineari. Output di Gretl

```
Sono state usate derivate analitiche
Tolleranza = 1,81899e-012
Convergenza raggiunta dopo 21 iterazioni
```

Modello 1: NLS, usando le osservazioni 1980–2010 ($T$ = 31)
`y = alfa/(1+beta*exp(-kappa*time))`

| | Stima | Errore Std. | rapporto $t$ | p-value |
|---|---|---|---|---|
| alfa | 2,02107 | 0,0217029 | 93,1245 | 0,0000 |
| $\beta$ | 20,7218 | 2,33371 | 8,8793 | 0,0000 |
| kappa | 0,264036 | 0,0106818 | 24,7182 | 0,0000 |

| | | | |
|---|---|---|---|
| Media var. dipendente | 1,293839 | SQM var. dipendente | 0,698388 |
| Somma quadr. residui | 0,087733 | E.S. della regressione | 0,055976 |
| $R^2$ | 0,994004 | $R^2$ corretto | 0,993576 |
| Log-verosimiglianza | 46,95825 | Criterio di Akaike | −87,91650 |
| Criterio di Schwarz | −83,61454 | Hannan–Quinn | −86,51417 |
| $\hat{\rho}$ | −0,083529 | Durbin–Watson | 2,164108 |

```
params alfa beta kappa
```

In questo caso le stime ottenute saranno esattamente le stesse, ma il numero di iterazioni necessarie sale considerevolmente (questo risultato viene lasciato alla verifica del lettore).

In ogni caso, si è pervenuti al modello stimato

$$\hat{y}_t^{NL} = \frac{2{,}021}{1 + 20{,}722\mathrm{e}^{(-0{,}264t)}}$$

con $\bar{R}^2 = 0{,}9936$.

Mentre per $\alpha$ e $k$ i valori sono pressoché identici, le stime per $\beta$ sono marcatamente differenti: il valore stimato secondo i minimi quadrati è notevolmente superiore a quello trovato con il metodo delle somme. La Figura 2.2 confronta il grafico della serie $y_t$ con le serie stimate secondo i parametri ottenuti con il metodo delle somme ($\hat{y}_t^{MS}$) e con il metodo dei minimi quadrati non lineari ($\hat{y}_t^{NL}$). La curva più vicina ai dati osservati è senza dubbio la serie ($\hat{y}_t^{NL}$), che evidentemente è basata su una migliore stima di $\beta$. In effetti, la curva ($\hat{y}_t^{NL}$) è traslata verso il basso rispetto a ($\hat{y}_t^{MS}$) ed è più vicina alle osservazioni $y_t$. D'altronde, il parametro $\beta$ determina proprio il punto d'incontro della curva con l'asse verticale: maggiore è il suo valore, minore sarà il valore di $f(0)$.

Utilizzando le stime ricavate con il metodo delle somme come valori iniziali dei parametri $\alpha$, $\beta$ e $k$ non si rileva alcuna differenza nella stima dei parametri secondo il metodo dei minimi quadrati non lineari. In questo caso la scelta arbitraria della terna di valori (0,1; 0,1; 0,1) ha consentito di ottenere risultati del tutto soddisfacenti.

(c) L'esercizio richiede ora di stimare un trend polinomiale di grado opportuno e di confrontare i risultati con la stima del trend logistico.

Figura 2.2
Grafico della serie $\{y_t\}_{t=2}^{31}$ e dei valori stimati $\hat{y}_t^{MS}$ e $\hat{y}_t^{NL}$

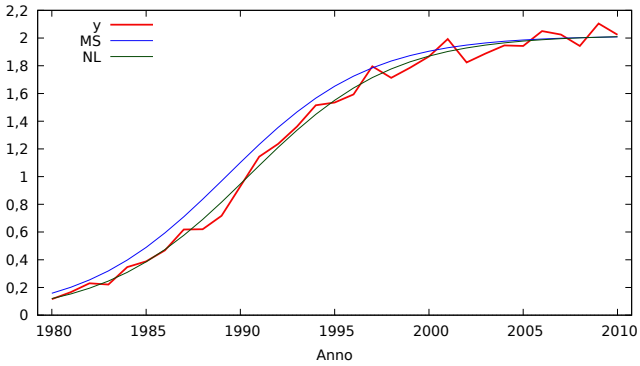
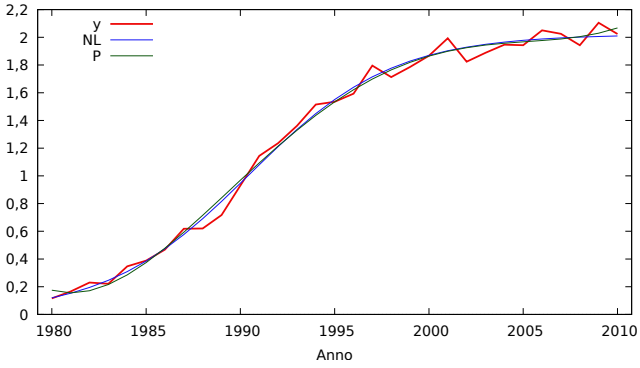



Figura 2.3
Grafico della serie $\{y_t\}_{t=1}^{31}$ e dei valori stimati $\hat{y}_t^{NL}$ e $\hat{y}_t^{P}$




Utilizzando il criterio degli $\bar{R}^2$ e della significatività statistica dei coefficienti, il seguente polinomio del quarto grado è stato adattato alla serie per $t = 1, \ldots, 31$ con il metodo dei minimi quadrati:

$$\begin{array}{cccccccccc} \hat{y}_t^P = & 0{,}234401 & - & 0{,}0820894t & + & 0{,}023363t^2 & - & 0{,}00105214t^3 & + & 1{,}43693 \cdot 10^{-5}t^4 \\ & (0{,}0694999) & & (0{,}0290698) & & (0{,}00361284) & & (0{,}000168374) & & (2{,}61178 \cdot 10^{-6}) \end{array}$$

con un $\bar{R}^2$ pari a $0{,}9919$. Il grado di accostamento del modello polinomiale risulta quindi inferiore a quello ottenuto dal trend logistico. La perdita di aderenza con i valori osservati è particolarmente evidente all'inizio ed alla fine del periodo di stima, come si può notare nella Figura 2.3. Se l'obiettivo fosse la previsione dei valori futuri, l'utilizzo di un trend polinomiale risulterebbe quindi poco appropriato.

## Esercizi supplementari

**C2.1** Discutere il problema della scelta del grado del polinomio per la stima del trend di una serie storica.

**C2.2** Sia data una serie storica semestrale, $\{y_t\}_{t=1}^{n}$, composta da un trend parabolico, da oscillazioni stagionali e da una componente erratica. Specificare, utilizzando il metodo delle variabili *dummy*, un modello opportuno per tale serie.

**C2.3** Si consideri la serie storica trimestrale generata dal seguente modello:

$$y_t = 10t + 0{,}5t^2 + 5d_{1t} - 10d_{2t} - 7d_{3t} + 12d_{4t} + \varepsilon_t$$

con $\varepsilon_t \sim N(0{,}16)$, $t = 1, \ldots 12$ e $d_{1t}$, $d_{2t}$, $d_{3t}$ e $d_{4t}$ variabili *dummy* introdotte per rappresentare, rispettivamente, la prima, la seconda, la terza e la quarta stagione dell'anno.

(a) Sulla base delle informazioni fornite è possibile calcolare i valori della serie destagionalizzata?

(b) Si consideri il valore della serie al tempo 1, corrispondente al primo trimestre di un anno. Quanto vale mediamente la serie in tale periodo? E quanto ci si aspetta sia variabile il valore della serie in tale periodo attorno alla sua media?

(c) È possibile eliminare la componente di trend applicando alla serie l'operatore differenza prima?

**C2.4** Risolvere l'esercizio 2.1 ipotizzando che la serie $y_t$, $t = 2, \ldots, 10$, sia una serie trimestrale in cui $t = 1$ corrisponde al primo trimestre del primo anno di osservazione. Usare un modello con intercetta e privo della variabile *dummy* relativa al secondo trimestre.

**C2.5** Usare Gretl per stimare il trend esponenziale per la serie dell'esercizio 2.7 col metodo dei minimi quadrati non lineari. Confrontare i risultati con quelli ottenuti stimando il modello linearizzato e commentare.

# 3
# Le medie mobili

### Proprietà delle medie mobili

**3.1** Quali condizioni devono essere soddisfatte perché una media mobile di ordine $2m+1$ conservi un trend lineare? Giustificare la risposta.

□ Per lasciare inalterata una generica componente di trend $T_t$ i coefficienti $\vartheta_i$ di una media mobile, per $i = -m, \ldots, m$, devono soddisfare la seguente relazione

$$\sum_{i=-m}^{m} \vartheta_i T_{t+i} = T_t.$$

Nel caso specifico di un trend lineare si avrà

$$\sum_{i=-m}^{m} \vartheta_i [a + b(t+i)] = a + bt.$$

È immediato verificare che

$$\sum_{i=-m}^{m} \vartheta_i [a + b(t+i)] = a \sum_{i=-m}^{m} \vartheta_i + bt \sum_{i=-m}^{m} \vartheta_i + b \sum_{i=-m}^{m} i\vartheta_i. \tag{3.1}$$

Affinché il secondo membro dell'espressione (3.1) risulti pari a $a + bt$ i coefficienti della media mobile devono soddisfare le seguenti due relazioni

$$\sum_{i=-m}^{m} \vartheta_i = 1 \tag{3.2}$$

$$\sum_{i=-m}^{m} i\vartheta_i = 0, \tag{3.3}$$

che rappresentano le condizioni necessarie per il mantenimento di un trend di tipo lineare.

**3.2** Si dispone di una serie storica giornaliera che presenta una componente stagionale costante di periodo 7. (a) Stabilire se una media mobile aritmetica semplice a 7 termini elimina tale stagionalità. (b) Derivare l'associato rapporto di riduzione della varianza (c) Mostrare che la media mobile al punto (a) conserva trend polinomiali di primo grado ma non quelli di secondo grado.

□ (a) Sì, è una proprietà fondamentale della media mobile aritmetica: elimina componenti stagionali (costanti) di periodo *pari* al suo ordine. Si noti che in questo caso essendo il periodo della stagionalità dispari non c'è alcun bisogno di ricorrere alla centratura della media mobile.

(b) Il rapporto di riduzione della varianza di una media mobile aritmetica è pari al reciproco del suo ordine. In questo caso risulta pari a 1/7.

(c) La somma dei coefficienti della media mobile è pari ad 1. Si tratta inoltre di una media mobile simmetrica. Ne segue che conserva i trend lineari. Il fatto che la somma

$$\sum_{i=-3}^{3} i^2 \vartheta_i$$

risulti diversa da 0 comporta che la media mobile in questione *non* conservi trend parabolici.

**3.3** Si derivi la media mobile a 3 termini che minimizza il rapporto di riduzione della varianza tra quelle che conservano un trend polinomiale di secondo grado. (a) Oltre a quella trovata, quante altre medie mobili a tre termini conservano trend polinomiali di secondo grado? (b) La media mobile trovata conserva i trend lineari? (c) A quanto ammonta il rapporto di riduzione della varianza? (d) La media mobile a tre termini che minimizza il rapporto di riduzione dsella varianza tra quelle che conservano trend lineari, dà luogo ad un rapporto di riduzione della varianza maggiore, minore o uguale a quello trovato?

□ L'insieme delle medie mobili a 3 termini che conservano polinomi di secondo grado è costituito da quelle che soddisfano le seguenti condizioni:

$$\begin{aligned} \vartheta_{-1} + \vartheta_0 + \vartheta_1 &= 1 \\ -\vartheta_{-1} + \vartheta_1 &= 0 \\ (-1)^2\vartheta_{-1} + (1)^2\vartheta_1 &= 0 \end{aligned}$$

È immediato verificare che il sistema risultante ammette una ed una sola soluzione: $\vartheta_{-1} = \vartheta_1 = 0$ , $\vartheta_0 = 1$. La media mobile che ne risulta lascia inalterata la serie storica alla quale viene applicata.

(a) Per quanto detto al punto precedente, non vi sono altre medie mobili a 3 termini che conservano polinomi di secondo grado.

(b) Sì, la media mobile trovata conserva i trend lineari in quanto soddisfa le condizioni

$$\begin{aligned} \vartheta_{-1} + \vartheta_0 + \vartheta_1 &= 1 \\ -\vartheta_{-1} + \vartheta_1 &= 0. \end{aligned}$$

(c) Il rapporto di riduzione della varianza per una generica media mobile di ordine $2m + 1$ risulta pari a

$$\frac{\sigma^{2*}}{\sigma^2} = \sum_{i=-m}^{m} \vartheta_i^2. \tag{3.4}$$

In questo caso il rapporto è pari a 1. Si tratta di un'ovvia conseguenza del fatto che la media mobile in questione non modifica la serie storica alla quale viene applicata.

(d) La media mobile a 3 termini che conserva polinomi di primo grado è la media mobile aritmetica, alla quale corrisponde un rapporto di riduzione della varianza pari a 1/3, inferiore a quello che si trova imponendo il vincolo di conservazione dei polinomi di secondo grado.

**3.4** Si considerino le due medie mobili

$$M_1 = -y_{t-2} + \frac{1}{2}y_{t-1} + \frac{5}{4}y_t + \frac{3}{4}y_{t+1} - \frac{1}{2}y_{t+2}$$
$$M_2 = \frac{1}{4}y_{t-1} + \frac{1}{2}y_t + \frac{1}{4}y_{t+1}$$

e si supponga di volerle applicare (i) ad una serie storica incorrelata al fine di ridurne la variabilità e (ii) ad una serie storica generata da un modello con trend evolutivo e componente erratica, al fine di stimarne il trend.

(a) Quale delle due medie mobili sarebbe preferibile e perché?

(b) Ci sono motivi per ritenere che una delle due medie mobili non debba essere utilizzata?

□ (a) La riduzione di variabilità di una serie a seguito dell'applicazione di una media mobile si misura con il rapporto definito nella (3.4). Tuttavia, per $M_1$ tale quantità risulta pari a 3,625. Pertanto, questa media mobile produce una serie storica la cui componente accidentale ha una variabilità che è oltre tre volte e mezzo rispetto a quella della serie originale! È chiaro quindi che l'obiettivo di riduzione della componente erratica nel caso (i) può essere perseguito solo mediante l'utilizzo di $M_2$, con un rapporto di riduzione pari al 37,5%.

Anche per il caso (ii) la media $M_2$ risulta più appropriata. Per essa infatti valgono le condizioni (3.2) e (3.3), mentre per $M_1$ si ha

$$\sum_{i=-2}^{2} i\vartheta_i = \frac{8-2+3-4}{4} = \frac{5}{4} \neq 0.$$

Ne consegue che l'applicazione di $M_1$ non lascerebbe inalterato un trend di tipo lineare.

(b) Entrambe le medie mobili sono centrate, mentre solo $M_2$ è simmetrica. La proprietà di simmetria dei coefficienti è fondamentale per ottenere i risultati desiderabili dall'applicazione di una media mobile, ovvero la riduzione della componente d'errore, la conservazione del trend o la depurazione della componente stagionale. L'utilizzo della media mobile $M_1$ è quindi sconsigliabile.

**3.5** Dimostrare che la media mobile semplice minimizza il rapporto di riduzione della varianza residua e conserva trend lineari.

□ La dimostrazione può essere condotta ricordando che i coefficienti di una media mobile possono essere derivati dalla soluzione dal seguente problema di minimo vincolato

$$\begin{cases} \text{Min}_{\boldsymbol{\vartheta}} & \boldsymbol{\vartheta}'\boldsymbol{\Omega}\boldsymbol{\vartheta} \\ \text{soggetto a} & \mathbf{R}'\vartheta = \boldsymbol{\alpha} \end{cases} \tag{3.5}$$

dove $\boldsymbol{\vartheta}$ è il vettore di dimensione $2m+1$ contenente i coefficienti della media mobile, $\boldsymbol{\Omega}$ è una matrice simmetrica e positiva definita di dimensioni $2m+1$, $\mathbf{R}$ è una matrice $(2m+1 \times k)$ ed $\boldsymbol{\alpha}$ un vettore

($k \times 1$). Questi ultimi due oggetti definiscono i vincoli del sistema da risolvere. La soluzione al sistema (3.5) è data da

$$\boldsymbol{\vartheta} = \boldsymbol{\Omega}^{-1}\mathbf{R}(\mathbf{R}'\boldsymbol{\Omega}^{-1}\mathbf{R})^{-1}\boldsymbol{\alpha}. \tag{3.6}$$

Nel caso in esame la matrice $\boldsymbol{\Omega}$ è la matrice identità, mentre

$$\mathbf{R} = \begin{bmatrix} 1 & -m \\ 1 & -m+1 \\ \vdots & \vdots \\ 1 & m-1 \\ 1 & m \end{bmatrix} \quad \text{e} \quad \boldsymbol{\alpha} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Il prodotto matriciale $\mathbf{R}'\boldsymbol{\vartheta} = \boldsymbol{\alpha}$ impone le condizioni (3.2-3.3). Sostituendo gli elementi noti nella (3.6) e risolvendo si ottiene

$$\boldsymbol{\vartheta} = \mathbf{R}(\mathbf{R}'\mathbf{R})^{-1}\boldsymbol{\alpha} = \begin{bmatrix} \frac{1}{2m+1} \\ \frac{1}{2m+1} \\ \vdots \\ \frac{1}{2m+1} \end{bmatrix}$$

ovvero la media mobile semplice di ordine $2m+1$.

**3.6** Derivare la media mobile di ordine 5 con rapporto di riduzione della varianza residua minimo e che conserva un trend quadratico.

□ La soluzione può essere facilmente derivata utilizzando la formulazione matriciale introdotta nell'esercizio precedente. La condizione di minimo sul rapporto di riduzione della varianza residua si ottiene assumendo $\boldsymbol{\Omega} = \mathbf{I}$, e poiché l'ordine della media mobile è 5, abbiamo $m = 2$. La conservazione di un trend quadratico è invece garantita se si definiscono $\mathbf{R}$ e $\boldsymbol{\alpha}$ nel seguente modo:

$$\mathbf{R} = \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \quad \text{e} \quad \boldsymbol{\alpha} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

I coefficienti della media mobile si ottengono quindi tramite la (3.6). Sostituendo si ottiene:

$$\begin{aligned} \boldsymbol{\vartheta} &= \mathbf{R}(\mathbf{R}'\mathbf{R})^{-1}\boldsymbol{\alpha} \\ &= \begin{bmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 0{,}4857 & 0 & -0{,}1429 \\ 0 & 0{,}1 & 0 \\ -0{,}1429 & 0 & 0{,}0714 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} -0{,}0857 \\ 0{,}3429 \\ 0{,}4857 \\ 0{,}3429 \\ -0{,}0857 \end{bmatrix}. \end{aligned}$$

La media mobile ricercata è pertanto simmetrica e può essere espressa nella notazione usuale

$$M = \left\{ [5]; \frac{1}{35}[-3,12,\mathbf{17}] \right\}.$$

**3.7** Derivare le condizioni che devono essere soddisfatte da una media mobile di ordine 5 per conservare il trend cubico $T_t = 1 - t^3$.

□ I coefficienti $\vartheta_i$, per $i = -2, -1,0,1,2$, della media mobile devono soddisfare la relazione

$$\sum_{i=-2}^{2} \vartheta_i T_{t+i} = T_t$$

ovvero

$$\sum_{i=-2}^{2} \vartheta_i \left[1 - (t+i)^3\right] = 1 - t^3.$$

Poiché $(t+i)^3 = t^3 + 3t^2 i + 3ti^2 + i^3$, si ha:

$$\sum_{i=-2}^{2} \vartheta_i \left[1 - t^3 - 3t^2 i - 3ti^2 - i^3\right] =$$
$$= \sum_{i=-2}^{2} \vartheta_i - t^3 \sum_{i=-2}^{2} \vartheta_i - 3t^2 \sum_{i=-2}^{2} i\vartheta_i - 3t \sum_{i=-2}^{2} i^2\vartheta_i - \sum_{i=-2}^{2} i^3\vartheta_i$$

Affinché questa espressione sia pari a $1 - t^3$ è necessario che

$$\sum_{i=-2}^{2} \vartheta_i = 1$$
$$\sum_{i=-2}^{2} i\vartheta_i = \sum_{i=-2}^{2} i^2\vartheta_i = \sum_{i=-2}^{2} i^3\vartheta_i = 0$$

**3.8** Si considerino le due serie $V_t$ e $Z_t$

$$V_t = \sum_{i=-2}^{2} \vartheta_i Y_{t+i} = M_1 Y_t \qquad Z_t = \sum_{i=-2}^{2} \psi_i Y_{t+i} = M_2 Y_t$$

dove $\{Y_t,\, t = 1,2,\cdots\}$ è una sequenza di variabili casuali indipendenti e identicamente distribuite di media nulla e varianza $\sigma^2$ ed $M_1$ e $M_2$ sono medie mobili con coefficienti pari a

$$\vartheta_2 = \vartheta_{-2} = -\frac{21}{286},\ \vartheta_1 = \vartheta_{-1} = \frac{84}{286} \ \text{ e } \ \vartheta_0 = \frac{160}{286};$$

$$\psi_2 = \psi_{-2} = -\frac{3}{35},\ \ \psi_1 = \psi_{-1} = \frac{12}{35} \ \text{ e } \ \psi_0 = \frac{17}{35}.$$

(a) Per quale delle due serie $V_t$ e $Z_t$ l'autocorrelazione del primo ordine $\rho_1$ è più elevata?

(b) Quale delle due medie mobili $M_1$ e $M_2$ conserva un trend parabolico?

(c) Quale delle due medie mobili dovrebbe indurre una dinamica meno irregolare nelle serie risultanti?

□ (a) Per calcolare $\rho_1$ è necessario derivare la media e la varianza delle serie $V_t$ e $Z_t$. Per la serie $V_t$ si ha

$$E(V_t) = E\left(\sum_{i=-2}^{2} \vartheta_i Y_{t+i}\right) = \sum_{i=-2}^{2} \vartheta_i E(Y_{t+i}) = 0$$
$$Var(V_t) = \left[E\left(\sum_{i=-2}^{2} \vartheta_i Y_{t+i}\right)^2\right] = \sigma^2 \sum_{i=-2}^{2} \vartheta_i^2 = 0{,}4963\sigma^2.$$

Analogamente,

$$E(Z_t) = E\left(\sum_{i=-2}^{2} \psi_i Y_{t+i}\right) = \sum_{i=-2}^{2} \psi_i E(Y_{t+i}) = 0$$
$$Var(Z_t) = \left[E\left(\sum_{i=-2}^{2} \psi_i Y_{t+i}\right)^2\right] = \sigma^2 \sum_{i=-2}^{2} \psi_i^2 = 0{,}4857\sigma^2.$$

Calcoliamo ora $\rho_1$ per $V_t$. Essa è pari a

$$\rho_1^V = \frac{E(V_t V_{t-1})}{Var(V_t)}$$

dove

$$E(V_t V_{t-1}) = E\left[(\vartheta_{-2}Y_{t-2} + \vartheta_{-1}Y_{t-1} + \vartheta_0 Y_t + \vartheta_1 Y_{t+1} + \vartheta_2 Y_{t+2}) \times (\vartheta_{-2}Y_{t-3} + \vartheta_{-1}Y_{t-2} + \vartheta_0 Y_{t-1} + \vartheta_1 Y_t + \vartheta_2 Y_{t+1})\right].$$

Poiché $E(Y_t Y_{t+s}) = 0$ per $s \neq 0$, il calcolo precedente si semplifica in

$$E(V_t V_{t-1}) = [\vartheta_{-2}\vartheta_{-1}E(Y_{t-2}^2) + \vartheta_{-1}\vartheta_0 E(Y_{t-1}^2) + \vartheta_0\vartheta_1 E(Y_t^2) + \vartheta_1\vartheta_2 E(Y_{t+1}^2)].$$

Sostituendo il valore dei coefficienti all'espressione precedente si ottiene

$$E(V_t V_{t-1}) = \sigma^2 \frac{1}{286^2}[-21 \cdot 84 + 84 \cdot 160 + 160 \cdot 84 - 84 \cdot 21] = 0{,}2855\sigma^2.$$

Quindi,

$$\rho_1^V = \frac{0{,}2855\sigma^2}{0{,}4963\sigma^2} = 0{,}5753.$$

Seguendo lo stesso procedimento per la serie $Z_t$ si ottiene

$$\rho_1^Z = \frac{0{,}2743\sigma^2}{0{,}4857\sigma^2} = 0{,}5648.$$

Pertanto, la serie storica che presenta autocorrelazione del primo ordine più elevata è $V_t$.

(b) Verifichiamo ora se le due medie mobili $M_1$ ed $M_2$ conservano un trend parabolico. Indichiamo tale trend con

$$T_t = a + bt + ct^2$$

e calcoliamo $M_1 T_t$ e $M_2 T_t$. Conviene esplicitare il trend in funzione del periodo di ritardo:

$$\begin{aligned}
T_{t-2} &= a + b(t-2) + c(t-2)^2 = a - 2b + 4c + (b - 4c)t + ct^2 \\
T_{t-1} &= a + b(t-1) + c(t-1)^2 = a - b + c + (b - 2c)t + ct^2 \\
T_t &= a + bt + ct^2 \\
T_{t+1} &= a + b(t+1) + c(t+1)^2 = a + b + c + (b + 2c)t + ct^2 \\
T_{t+2} &= a + b(t+2) + c(t+2)^2 = a + 2b + 4c + (b + 4c)t + ct^2.
\end{aligned}$$

Applicando $M_1$ a $T_t$ si ottiene

$$\begin{aligned}
\vartheta_{-2} T_{t-2} &= \frac{1}{286}[-21a + 42b - 84c - 21(b - 4c)t - 21ct^2] \\
\vartheta_{-1} T_{t-1} &= \frac{1}{286}[84a - 84b + 84c + 84(b - 2c)t + 84ct^2] \\
\vartheta_0 T_t &= \frac{1}{286}[160a + 160bt + 160ct^2] \\
\vartheta_1 T_{t+1} &= \frac{1}{286}[84a + 84b + 84c + 84(b + 2c)t + 84ct^2] \\
\vartheta_2 T_{t+2} &= \frac{1}{286}[-21a - 42b - 84c - 21(b + 4c)t - 21ct^2]
\end{aligned}$$

Sommando i 5 termini e dopo alcuni semplici passaggi si ricava

$$\sum_{i=-2}^{2} \vartheta_i T_{t+i} = \frac{1}{286}(286a + 286bt + 286ct^2) = a + bt + ct^2 = T_t.$$

Operando allo stesso modo con $M_2$ si ottiene

$$\sum_{i=-2}^{2} \psi_i T_{t+i} = \frac{1}{35}(35a + 35bt + 35ct^2) = a + bt + ct^2 = T_t.$$

Entrambe le medie mobili conservano quindi un trend parabolico. Si noti che $M_1$ è una media mobile polinomiale di secondo grado di ordine 3 mentre $M_2$ è una media mobile di Henderson di ordine 5.

(c) La media mobile $M_2$ dovrebbe comportare una dinamica meno irregolare di $M_1$ in quanto il suo rapporto di riduzione della varianza è minore (0,4857 contro 0,4963).

**3.9** Si consideri la media mobile

$$y_t^* = \frac{1}{4}(y_{t-2} + y_{t-1} + y_t + y_{t+1})$$

e si ipotizzi di applicarla ad una serie storica generata dalla somma di una componente stagionale costante di periodo 4 e di un trend lineare.

(a) Dire se questa media mobile è in grado di eliminare la componente periodica.

(b) Dire se questa media mobile è adatta a stimare il trend.

(c) Se la media mobile in questione venisse applicata ad un *white noise*, la serie perequata risulterebbe più o meno variabile di quella originale? Quantificare il miglioramento o il peggioramento.

□ (a) Una media mobile semplice di ordine 4 elimina una componente stagionale di periodo 4. Tale risultato è garantito anche da una media mobile di tipo asimmetrico, come quella incontrata in questo esercizio.

(b) Definiamo il trend lineare $T_t = a + bt$ e applichiamo la media mobile in questione. Avremo

$$\begin{aligned} T_t^* &= \frac{1}{4}T_{t-2} + \frac{1}{4}T_{t-1} + \frac{1}{4}T_t + \frac{1}{4}T_{t+1} \\ T_t^* &= \frac{1}{4}(a + b(t-2)) + \frac{1}{4}(a + b(t-1)) \\ &\quad + \frac{1}{4}(a + bt) + \frac{1}{4}(a + b(t+1)) \\ T_t^* &= a + bt - \frac{1}{2}b \end{aligned}$$

La stima del trend $T_t^*$ differisce da quello originale $T_t$ per la differente intercetta ($a - \frac{1}{2}b$). Pur mantenendo lo stesso coefficiente angolare, non si può affermare che tale media mobile risulti adatta per la stima di un trend lineare. Più consona risulterebbe l'applicazione di una media mobile semplice centrata e simmetrica.

(c) Il rapporto di riduzione della varianza è pari a $4 \cdot 0{,}25^2 = 0{,}25$, per cui la serie perequata risulterebbe meno variabile di quella originaria (la variabilità del suo termine di disturbo $\varepsilon_t^*$ si riduce ad un quarto di quella del disturbo $\varepsilon_t$).

### Composizioni di medie mobili

**3.10** È noto che la serie storica $y_t$ risulta dalla somma delle due componenti non osservabili $x_t$ e $z_t$. Siano $M_1$ e $M_2$ due medie mobili tali che $x_t$ è invariante rispetto a $M_1$ ed appartiene al nucleo di $M_2$, mentre $z_t$ è invariante rispetto a $M_2$ ed appartiene al nucleo di $M_1$. Stabilire le proprietà della serie $M_1M_2Y_t$.

□ In base al testo si ha:

$$M_1x_t = x_t \qquad M_2x_t = 0 \qquad M_1z_t = 0 \qquad M_2z_t = z_t.$$

Pertanto,

$$M_1M_2y_t = M_1M_2(x_t + z_t) = M_1(M_2x_t + M_2z_t) = M_1z_t = 0.$$

Si noti che il risultato non dipende dall'ordine con cui le medie mobili vengono applicate. Infatti:

$$M_2M_1y_t = M_2M_1(x_t + z_t) = M_2(M_1x_t + M_1z_t) = M_2x_t = 0.$$

**3.11** Indicare l'ordine e derivare i coefficienti della media mobile che si ottiene componendo una media mobile aritmetica semplice di ordine 4 ed una media mobile aritmetica semplice di ordine 6.

□ Indicata con $M$ la media mobile di interesse, si ha $M = M_2M_1$, con

$$M_1 = \left\{ [4]; \frac{1}{4}[1,1,1,1] \right\} \text{ e } M_2 = \left\{ [6]; \frac{1}{6}[1,1,1,1,1,1] \right\}.$$

L'ordine di $M$ è 4+6-1=9. Possiamo dunque scrivere $y^{**} = My = M_2M_1y = M_2y^*$. Il primo valore calcolabile per la serie $y^{**}$ è il quinto, di seguito si mostra in dettaglio una possibile procedura di calcolo del valore $y_5^{**}$. Il primo passaggio riguarda il calcolo dei valori $y^* = M_1y$ d'interesse:

$$\begin{aligned}
y^*_{[2,3]} &= \frac{1}{4}(y_1 + y_2 + y_3 + y_4) \\
y^*_{[3,4]} &= \frac{1}{4}(y_2 + y_3 + y_4 + y_5) \\
y^*_{[4,5]} &= \frac{1}{4}(y_3 + y_4 + y_5 + y_6) \\
y^*_{[5,6]} &= \frac{1}{4}(y_4 + y_5 + y_6 + y_7) \\
y^*_{[6,7]} &= \frac{1}{4}(y_5 + y_6 + y_7 + y_8) \\
y^*_{[7,8]} &= \frac{1}{4}(y_6 + y_7 + y_8 + y_9)
\end{aligned}$$

Successivamente si applica $M_2$ ai 6 termini calcolati:

$$y_5^{**} = \frac{1}{6}(y^*_{[2,3]} + y^*_{[3,4]} + y^*_{[4,5]} + y^*_{[5,6]} + y^*_{[6,7]} + y^*_{[7,8]}).$$

Organizzando i termini sotto forma matriciale il calcolo dei coefficienti della media mobile $M$ diventa molto più semplice. Infatti,

$$\begin{aligned}
y_5^{**} = \frac{1}{6}\frac{1}{4}\,(y_1+\;\; & y_2+\;\; y_3+\;\; y_4+ \\
& y_2+\;\; y_3+\;\; y_4+\;\; y_5+ \\
& \quad\; y_3+\;\; y_4+\;\; y_5+\;\; y_6+ \\
& \quad\quad\; y_4+\;\; y_5+\;\; y_6+\;\; y_7+ \\
& \quad\quad\quad\; y_5+\;\; y_6+\;\; y_7+\;\; y_8+ \\
& \quad\quad\quad\quad\; y_6+\;\; y_7+\;\; y_8+ y_9) \\
y_5^{**} = \frac{1}{24}\,(y_1+ 2y_2+ 3y_3+ 4y_4+ & 4y_5+ 4y_6+ 3y_7+ 2y_8+ y_9)\,.
\end{aligned}$$

Generalizzando, si ha

$$y_t^{**} = \frac{1}{24}(y_{t-4} + 2y_{t-3} + 3y_{t-2} + 4y_{t-1} + 4y_t + 4y_{t+1} + 3y_{t+2} + 2y_{t+3} + y_{t+4}).$$

Pertanto $M = \left\{ [9]; \frac{1}{24}[1,2,3,4,\mathbf{4}] \right\}$.

**3.12** Indicare l'ordine e derivare i coefficienti della media mobile che si ottiene componendo

$$M_1 = \left\{ [3]; \frac{1}{3}[1,\mathbf{1}] \right\} \text{ e } M_2 = \left\{ [5]; \frac{1}{9}[1,2,\mathbf{3}] \right\}$$

□ L'ordine di $M = M_1 M_2$ è 3+5-1=7. Ripetendo i passi del precedente esercizio possiamo scrivere

$$y^{**} = My = M_2 M_1 y = M_2 y^*$$

$$\begin{aligned}
y_2^* &= \frac{1}{3}(y_1 + y_2 + y_3) \\
y_3^* &= \frac{1}{3}(y_2 + y_3 + y_4) \\
y_4^* &= \frac{1}{3}(y_3 + y_4 + y_5) \\
y_5^* &= \frac{1}{3}(y_4 + y_5 + y_6) \\
y_6^* &= \frac{1}{3}(y_5 + y_6 + y_7)
\end{aligned}$$

Applicando $M_2$ ai 5 termini calcolati otteniamo:

$$y_4^{**} = \frac{1}{9}(y_2^* + 2y_3^* + 3y_4^* + 2y_5^* + y_6^*).$$

Ridisponendo i termini si ottiene

$$\begin{aligned}
y_4^{**} = \frac{1}{9}\frac{1}{3}\,(y_1+\ \ y_2+\ \ y_3+ & \\
2y_2+\ 2y_3+\ 2y_4+ & \\
3y_3+\ 3y_4+\ 3y_5+ & \\
2y_4+\ 2y_5+\ 2y_6+ & \\
y_5+\ \ y_6+\ y_7) & \\
y_4^{**} = \frac{1}{27}\,(y_1+\ 3y_2+\ 6y_3+\ 7y_4+\ 6y_5+\ 3y_6+\ y_7)\,. &
\end{aligned}$$

Generalizzando, si ha

$$y_t^{**} = \frac{1}{27}(y_{t-3} + 3y_{t-2} + 6y_{t-1} + 7y_t + 6y_{t+1} + 3y_{t+2} + y_{t+3})$$

con $M = \left\{[7]; \frac{1}{27}[1,3,6,\mathbf{7}]\right\}$.

**3.13** Si considerino le due medie mobili $M_1$ ed $M_2$ date da

$$M_1 = \left\{[5]; \left[-\frac{21}{286}, \frac{84}{286}, \frac{\mathbf{160}}{\mathbf{286}}\right]\right\}$$

$$M_2 = \left\{[15]; \frac{1}{320}[-3, -6, -5, 3, 21, 46, 67, \mathbf{74}]\right\}.$$

(a) Qual è il massimo grado di un trend polinomiale che $M_1$ ed $M_2$ riescono a conservare?

(b) Se si fosse interessati solo ad avere una serie perequata il più regolare possibile, quale delle due medie mobili è preferibile? Spiegare il perché.

□ (a) Le due medie mobili conservano polinomi di grado minore od uguale a 3. Ad esempio, per $M_1$ si ha:

$$\begin{aligned}
\sum_{i=-2}^{2} \vartheta_i &= \left(-\frac{21}{286} + \frac{84}{286} + \frac{160}{286} + \frac{84}{286} - \frac{21}{286}\right) = 1 \\
\sum_{i=-2}^{2} i\vartheta_i &= \left(-2(-\frac{21}{286}) - 1(\frac{84}{286}) \right. \\
&\quad \left. +0(\frac{160}{286}) + 1(\frac{84}{286}) + 2(-\frac{21}{286})\right) = 0 \\
\sum_{i=-2}^{2} i^2\vartheta_i &= \left((-2)^2(-\frac{21}{286}) + (-1)^2(\frac{84}{286}) \right. \\
&\quad \left. +0^2(\frac{160}{286}) + 1^2(\frac{84}{286}) + 2^2(-\frac{21}{286})\right) = 0 \\
\sum_{i=-2}^{2} i^3\vartheta_i &= \left(-2^3(-\frac{21}{286}) - 1^3(\frac{84}{286}) \right. \\
&\quad \left. +0^3(\frac{160}{286}) + 1^3(\frac{84}{286}) + 2^3(-\frac{21}{286})\right) = 0 \\
\sum_{i=-2}^{2} i^4\vartheta_i &= \left((-2)^4(-\frac{21}{286}) + (-1)^4(\frac{84}{286}) \right. \\
&\quad \left. +0^4(\frac{160}{286}) + 1^4(\frac{84}{286}) + 2^4(-\frac{21}{286})\right) \neq 0
\end{aligned}$$

(b) La media mobile da preferire è quella che ha il rapporto di riduzione della varianza residua più basso in quanto ciò comporta una maggiore riduzione della componente irregolare. Per $M_1$ si ha

$$\begin{aligned}
\sum_{i=-2}^{2} \vartheta_i^2 &= \left(-\frac{21}{286}\right)^2 + \left(\frac{84}{286}\right)^2 + \left(\frac{160}{286}\right)^2 \\
&\quad + \left(\frac{84}{286}\right)^2 + \left(-\frac{21}{286}\right)^2 = 0.496
\end{aligned}$$

Invece, per $M_2$ si ha

$$\begin{aligned}
\sum_{i=-7}^{7} \vartheta_i^2 &= 2\left(-\frac{3}{320}\right)^2 + 2\left(-\frac{6}{320}\right)^2 + 2\left(-\frac{5}{320}\right)^2 \\
&\quad +2\left(\frac{3}{320}\right)^2 + 2\left(\frac{21}{320}\right)^2 + 2\left(\frac{46}{320}\right)^2 \\
&\quad +2\left(\frac{67}{320}\right)^2 + \left(\frac{74}{320}\right)^2 = 0.193
\end{aligned}$$

La media da preferire è dunque $M_2$. Si noti, per inciso, che $M_1$ è una media mobile di Henderson adatta per serie storiche trimestrali, in quanto elimina approssimativamente la componente stagionale e minimizza una quantità che esprime lo scarto esistente tra una parabola e la curva descritta dai coefficienti $\vartheta_i$ (si veda Di Fonzo e Lisi, 2005, p. 116-117). Invece, $M_2$ è una media mobile di Spencer, anch'essa utile per serie trimestrali. Essa è il risultato della composizione di 4 medie

mobili (di cui 3 semplici), ed ha la proprietà di annullare una stagionalità di periodo 4 o 5 con ampiezza variabile linearmente.

### Stima delle componenti con medie mobili

**3.14** Delineare una strategia (elementare) di destagionalizzazione mediante medie mobili di una serie storica trimestrale che si suppone sia stata generata da un modello additivo.

□ Una semplice strategia di stima della componente stagionale (e conseguente destagionalizzazione) per una serie storica trimestrale generata da un modello additivo

$$y_t = T_t + S_t + \varepsilon_t$$

consiste nello stimare i valori del trend-ciclo con medie mobili centrate di ordine 5,

$$\hat{T}_t = y_t^{**} = \frac{1}{8}(y_{t-2} + 2y_{t-1} + 2y_t + 2y_{t+1} + y_{t+2}),$$

e dalla successiva individuazione dei valori stimati per la componente stagionale e aleatoria, che indichiamo con $\widehat{(S+\varepsilon)}_t = y_t - \hat{T}_t$. Il passaggio successivo consiste nel verificare se la componente stagionale è significativamente presente nei valori $\widehat{(S+\varepsilon)}_t$. In caso di risposta affermativa, si calcolano 4 coefficienti grezzi di stagionalità, ottenuti considerando le medie dei valori di $\widehat{(S+\varepsilon)}_t$ appartenenti ai medesimi trimestri:

$$S_j^* = \frac{1}{N}\sum_{T=1}^{N} \widehat{(S+\varepsilon)}_{T,j} \quad j = 1,2,3,4$$

in cui $N$ è il numero di anni. L'ultimo passaggio da compiere è la trasformazione dei coefficienti grezzi in modo che soddisfino la condizione di avere somma nulla all'interno di ciascun anno (coefficienti *ideali*). Un semplice modo per ottenere ciò consiste nel considerare gli scarti dalla media:

$$\hat{S}_j = S_j^* - \bar{S}^*, \quad j = 1,2,3,4$$

con $\bar{S}^* = \frac{1}{4}\sum_{j=1}^{4} S_j^*$.

Infine, la serie destagionalizzata si ottiene semplicemente sottraendo $\hat{S}_j$ al pertinente valore $y_t$, ossia

$$y_{T,j}^d = y_{T,j} - \hat{S}_j, \quad T = 1, \cdots, N \;\; j = 1,2,3,4.$$

**3.15** Per modellare la serie storica trimestrale $y_t$ è stato usato un modello additivo del tipo $y_t = T_t + S_t + \varepsilon_t$. Utilizzando un'opportuna media mobile, i valori dei coefficienti specifici di stagionalità per i quattro trimestri degli anni $T-4$, $T-3$, $T-2$ e $T-1$ sono risultati:

| *Anno* | I trim | II trim | III trim | IV trim |
|---|---|---|---|---|
| $T-4$ | -6.951 | -2.808 | -12.517 | 21.679 |
| $T-3$ | -6.940 | -3.543 | -10.572 | 23.600 |
| $T-2$ | -9.896 | -5.064 | -11.364 | 29.938 |
| $T-1$ | -9.613 | -2.799 | -13.075 | 27.265 |

Per i quattro trimestri dell'anno $T$ sono stati osservati i seguenti valori:

$y_{T,1} = 367.861 \;\; y_{T,2} = 380.308 \;\; y_{T,3} = 377.246 \;\; y_{T,4} = 424.881$

Sulla base di questi dati, relativamente all'anno $T$, determinare:

(a) i valori assunti dalla componente stagionale

(b) i valori destagionalizzati mediante i coefficienti grezzi di stagionalità

(c) i valori destagionalizzati mediante i coefficienti ideali di stagionalità

□ (a) La tabella mostra i coefficienti specifici di stagionalità, che rappresentano una stima della componente stagionale congiunta ai fattori accidentali della serie ($\widehat{S_t + \varepsilon_t}$). L'andamento dei coefficienti fa presumere la presenza di una componente stagionale, con un forte picco positivo nel quarto trimestre ed un calo stagionale nel terzo. Per ridurre i fattori accidentali si calcolano delle medie stagionali dei coefficienti specifici:

$$\hat{S}_j^* = \frac{1}{N}\sum_{T=1}^{N} CS_{T,j} \qquad j = 1, \cdots, 4 \tag{3.7}$$

dove $N$ è il numero di anni e $CS_{T,j}$ è il coefficiente specifico corrispondente al periodo $j$ dell'anno $T$. In questo esercizio abbiamo

$\hat{S}_1^* = -8.350 \;\; \hat{S}_2^* = -3.553 \;\; \hat{S}_3^* = -11.882 \;\; \hat{S}_4^* = 25.620.$

I fattori stagionali dovrebbero compensarsi in un certo anno. Nel caso di un modello additivo tale proprietà implica l'annullamento dell'aggregazione annuale della componente stagionale, ovvero

$$\sum_{j=1}^{4} \hat{S}_j^* = 0.$$

Tuttavia le medie stagionali $\hat{S}_j^*$ non godono di tale proprietà. Infatti, si ha

$$\frac{1}{4}\sum_{j=1}^{4} \hat{S}_j^* = \bar{S}^* = 458{,}75.$$

Per tale motivo i coefficienti $\hat{S}_j^*$ sono detti *grezzi*, mentre i coefficienti *ideali* si ottengono sottraendo ad essi la loro media $\bar{S}^*$. I coefficienti ideali sono allora pari a

$$\begin{aligned}
\hat{S}_1 &= -8.808{,}75 \\
\hat{S}_2 &= -4.011{,}75 \\
\hat{S}_3 &= -12.340{,}75 \\
\hat{S}_4 &= 25.161{,}25.
\end{aligned}$$

I coefficienti ideali $\hat{S}_j$ costituiscono la componente stagionale calcolata mediante il metodo delle medie mobili, ipotizzata costante nel periodo di osservazione.

(b) I valori destagionalizzati mediante i coefficienti grezzi di stagionalità per l'anno $T$ si ottengono sottraendo ai valori $y_{T,j}$ i coefficienti grezzi $\hat{S}_j^*$:

$$\begin{aligned}
\hat{y}_{T,1}^{d*} &= y_{T,1} - \hat{S}_1^* = 376.211 \\
\hat{y}_{T,2}^{d*} &= y_{T,2} - \hat{S}_2^* = 383.862 \\
\hat{y}_{T,3}^{d*} &= y_{T,3} - \hat{S}_3^* = 389.128 \\
\hat{y}_{T,4}^{d*} &= y_{T,4} - \hat{S}_4^* = 399.261.
\end{aligned}$$

(c) I valori destagionalizzati mediante i coefficienti ideali di stagionalità per l'anno $T$ si ottengono sottraendo ai valori $y_{T,j}$ i coefficienti ideali $\hat{S}_j$:

$$\hat{y}^d_{T,1} = y_{T,1} - \hat{S}_1 = 376.679{,}75$$
$$\hat{y}^d_{T,2} = y_{T,2} - \hat{S}_2 = 384.329{,}75$$
$$\hat{y}^d_{T,3} = y_{T,3} - \hat{S}_3 = 389.586{,}75$$
$$\hat{y}^d_{T,4} = y_{T,4} - \hat{S}_4 = 399.719{,}75.$$

**3.16** Si consideri la serie storica dell'indice generale dei prezzi al consumo per le famiglie di operai ed impiegati (base 1995=100)

| *Anno* | *Mese* | *Indice generale* |
|---|---|---|
| **1995** | Dicembre | 102,2 |
| **1996** | Gennaio | 102,4 |
| | Febbraio | 102,7 |
| | Marzo | 103,0 |
| | Aprile | 103,6 |
| | Maggio | 104,0 |
| | Giugno | 104,2 |
| | Luglio | 104,0 |
| | Agosto | 104,1 |
| | Settembre | 104,4 |
| | Ottobre | 104,5 |
| | Novembre | 104,8 |
| | Dicembre | 104,9 |
| **1997** | Gennaio | 105,1 |
| | Febbraio | 105,2 |
| | Marzo | 105,3 |
| | Aprile | 105,4 |

(a) Si stimi la componente tendenziale mediante una media mobile aritmetica semplice di ordine 3.

(b) Nell'ipotesi che la serie storica sia del tipo $y_t = T_t + \varepsilon_t$, con $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$, si valuti se tale procedura, anche empiricamente, introduce un effetto Slutsky-Yule sulla serie stimata dei residui $\hat{\varepsilon}_t$.

□ La stima del trend mediante una media mobile aritmetica semplice di ordine 3 è data da

$$\hat{T}_t = \frac{y_{t-1} + y_t + y_{t+1}}{3}.$$

I calcoli necessari per la soluzione dell'esercizio sono riportati nella tabella seguente.

| $t$ | $y_t$ | $\hat{T}_t$ | $\hat{\varepsilon}_t$ | $(\hat{\varepsilon}_t - \bar{\hat{\varepsilon}})^2$ | $(\hat{\varepsilon}_t - \bar{\hat{\varepsilon}})(\hat{\varepsilon}_{t+1} - \bar{\hat{\varepsilon}})$ |
|---|---|---|---|---|---|
| 1 | 102,2 | - | - | - | - |
| 2 | 102,4 | 102,43 | -0,0333 | 0,0013 | 0,0001 |
| 3 | 102,7 | 102,70 | 0,0000 | 0,0000 | 0,0002 |
| 4 | 103,0 | 103,10 | -0,1000 | 0,0104 | -0,0066 |
| 5 | 103,6 | 103,53 | 0,0667 | 0,0042 | 0,0042 |
| 6 | 104,0 | 103,93 | 0,0667 | 0,0042 | 0,0084 |
| 7 | 104,2 | 104,07 | 0,1333 | 0,0172 | -0,0134 |
| 8 | 104,0 | 104,10 | -0,1000 | 0,0104 | 0,0070 |
| 9 | 104,1 | 104,17 | -0,0667 | 0,0047 | -0,0044 |
| 10 | 104,4 | 104,33 | 0,0667 | 0,0042 | -0,0044 |
| 11 | 104,5 | 104,57 | -0,0667 | 0,0047 | -0,0044 |
| 12 | 104,8 | 104,73 | 0,0667 | 0,0042 | -0,0023 |
| 13 | 104,9 | 104,93 | -0,0333 | 0,0013 | -0,0011 |
| 14 | 105,1 | 105,07 | 0,0333 | 0,0010 | -0,0001 |
| 15 | 105,2 | 105,20 | 0,0000 | 0,0000 | 0,0000 |
| 16 | 105,3 | 105,30 | 0,0000 | - | - |
| 17 | 105,4 | - | - | - | - |
| **Somma** | | | 0,0333 | 0,0678 | -0,0168 |
| **Media** | | | 0,0022 | | |

La stima dell'autocorrelazione di ordine 1 della serie dei residui stimati è dunque pari a

$$\hat{\rho}_1 = \frac{\hat{\gamma}_1}{\hat{\gamma}_0} = \frac{\sum_{t=2}^{15}(\hat{\varepsilon}_t - \bar{\hat{\varepsilon}})(\hat{\varepsilon}_{t+1} - \bar{\hat{\varepsilon}})}{\sum_{t=2}^{15}(\hat{\varepsilon}_t - \bar{\hat{\varepsilon}})^2} = \frac{-0{,}0168}{0{,}0678} = -0{,}2478$$

La serie dei residui stimati presenta dipendenza lineare negativa allo sfasamento temporale 1. Ciò potrebbe indicare che la stima del trend non ha avuto esito positivo, non essendo la serie dei residui riconducibile ad un white noise. Tuttavia, l'affermazione andrebbe verificata alla luce della significatività dell'autocorrelazione stimata, nonché dei valori fatti registrare dalle autocorrelazioni per sfasamenti temporali superiori.

**3.17** Si dispone di una serie storica trimestrale $z_t$ risultante dalla somma di un trend parabolico, di una componente stagionale di periodo 4 e di una componente accidentale di varianza pari a $\sigma^2_\varepsilon$. Il modello per $z_t$ risulta quindi:

$$z_t = a + bt + ct^2 + S_t + \varepsilon_t, \quad Var(\varepsilon_t) = \sigma^2_\varepsilon.$$

Siano $y_t$ e $x_t$ le serie storiche ottenute da $z_t$ nel modo seguente:

$$y_t = (1 - B)(1 - B^4)z_t$$
$$x_t = (1 - B)^2 \frac{1}{4} \sum_{l=0}^{3} z_{t-l}.$$

(a) Derivare le caratteristiche del trend, della stagionalità e della componente accidentale di $y_t$. Quanto vale l'autocorrelazione della componente accidentale di $y_t$?

(b) Derivare le caratteristiche del trend, della stagionalità e della componente accidentale di $x_t$. Quanto vale l'autocorrelazione della componente accidentale di $x_t$?

(c) Stabilire che relazione corre tra $y_t$ e $x_t$.

□ (a) Per verificare cosa succede alle componenti di $z_t$ nella serie $y_t$ conviene riscrivere il filtro nell'operatore di ritardo $B$ utilizzato nella costruzione della serie. Per il polinomio stagionale $(1 - B^4)$ vale l'equivalenza

$$(1 - B^4) = (1 - B)(1 + B + B^2 + B^3)\,,$$

ovvero esso può essere scritto come il prodotto tra l'operatore differenza $(1-B)$ e l'operatore sommatoria stagionale $(1 + B + B^2 + B^3)$. La serie $y_t$ può essere allora equivalentemente ottenuta come

$$y_t = (1 - B)^2(1 + B + B^2 + B^3)z_t.$$

Applichiamo ora tale polinomio alle singole componenti di $z_t$. Come è noto, l'applicazione di un'operatore differenza di ordine $q$ riduce il grado di un polinomio di $q$ unità. L'operatore $(1 - B)^2$ implica quindi la rimozione del trend parabolico, che ha ordine 2. Verifichiamo tale proprietà:

$$\begin{aligned}(1 - B)^2(a + bt + ct^2) &= \\ (1 - 2B + B^2)(a + bt + ct^2) &= a + bt + ct^2 \\ &\quad -2a - 2b(t - 1) - 2c(t - 1)^2 \\ &\quad a + b(t - 2) + c(t - 2)^2 \\ &= 2c \quad .\end{aligned}$$

Il trend parabolico si trasforma nella costante $2c$, che corrisponde al doppio del coefficiente del termine di secondo grado. Applicando l'operatore sommatoria si ottiene

$$(1 + B + B^2 + B^3)2c = 8c\,;$$

pertanto il trend in $y_t$ scompare a favore di un termine costante pari a $8c$.

Per stabilire cosa succede alla componente stagionale applichiamo il filtro di partenza ad $S_t$:

$$\begin{aligned}(1 - B)(1 - B^4)S_t &= (1 - B - B^4 + B^5)S_t \\ &= S_t - S_{t-1} - S_{t-4} + S_{t-5} = 0.\end{aligned} \tag{3.8}$$

Il risultato discende dalla proprietà di una componente stagionale deterministica di periodo $p$ per cui vale $S_t = S_{t-p}$. Come si nota, l'impiego della differenza stagionale $(1 - B^4)$ è da solo sufficiente ad eliminare $S_t$.

Infine, analizziamo le proprietà della componente accidentale $\varepsilon_t$. Essendo l'unica componente stocastica presente in $z_t$, valuteremo la serie risultante in termini di valore atteso, varianza ed autocovarianza della serie. Indichiamo con $\eta_t$ la nuova componente accidentale:

$$\eta_t = (1 - B)(1 - B^4)\varepsilon_t = \varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-4} + \varepsilon_{t-5}.$$

Il valore atteso di $\eta_t$ è sempre nullo. La varianza, invece, assume un valore superiore. Si ha infatti che

$$E(\eta_t^2) = E[(\varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-4} + \varepsilon_{t-5})^2] = 4\sigma_\varepsilon^2$$

data l'assenza di correlazione fra le $\varepsilon$ a qualsiasi ritardo. In pratica, questo vuol dire che la serie $y_t$ possiede una componente accidentale che in media è quattro volte superiore a quella originale contenuta in $z_t$.

Gli effetti del filtro sulla componente erratica non si esauriscono qui. Infatti, oltre all'aumento della varianza esso introduce anche correlazione seriale negli errori $\eta_t$. Tale effetto spurio è noto come effetto di Slutsky-Yule: l'applicazione di un qualsiasi polinomio nell'operatore di ritardo $B$ ad un processo temporalmente incorrelato induce correlazioni non nulle fra elementi successivi del processo risultante. Per il nostro caso verifichiamo ora cosa succede al processo $\eta_t$ esplicitandone le funzioni di autocovarianza $\gamma_k$ e di autocorrelazione $\rho_k$.

Per $k = 1$ abbiamo

$$\begin{aligned} E(\eta_t\eta_{t-1}) &= E[(\varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-4} + \varepsilon_{t-5}) \times \\ &\qquad (\varepsilon_{t-1} - \varepsilon_{t-2} - \varepsilon_{t-5} + \varepsilon_{t-6})] \\ &= -E(\varepsilon_{t-1}^2) - E(\varepsilon_{t-5}^2) \\ &= -2\sigma_\varepsilon^2 \ . \end{aligned}$$

Applicando lo stesso procedimento ai vari ritardi si ottiene la seguente funzione di autocovarianza

$$\gamma_k = \begin{cases} -2\sigma_\varepsilon^2 & k = 1 \\ 0 & k = 2 \\ \sigma_\varepsilon^2 & k = 3 \\ -2\sigma_\varepsilon^2 & k = 4 \\ \sigma_\varepsilon^2 & k = 5 \\ 0 & k > 5 \ . \end{cases}$$

Dividendo per $\gamma_0$ otteniamo la corrispondente funzione di autocorrelazione:

$$\rho_k = \begin{cases} -0{,}5 & k = 1 \\ 0 & k = 2 \\ 0{,}25 & k = 3 \\ -0{,}5 & k = 4 \\ 0{,}25 & k = 5 \\ 0 & k > 5 \ . \end{cases}$$

Il processo $\eta_t$ risulta quindi negativamente correlato ai ritardi 1 e 4 e positivamente in prossimità del ritardo stagionale.

(b) I risultati ottenuti per la serie $y_t$ sono validi anche per la serie $x_t$. Infatti, sfruttando l'operatore di ritardo B il modello utilizzato per la costruzione di questa serie può essere riscritto come

$$\begin{aligned} x_t &= (1 - B)^2 \frac{1}{4}(1 + B + B^2 + B^3)z_t \\ &= (1 - B)\frac{1}{4}(1 - B^4)z_t. \end{aligned}$$

A meno della costante $\dfrac{1}{4}$, il filtro utilizzato nella costruzione di $y_t$ e $x_t$ è identico. Per analizzare le proprietà di $x_t$ resta da verificare come questo fattore influisca sulle componenti di $z_t$ e quali implicazioni esso comporti nella struttura di autocovarianza della componente accidentale.

Per quanto riguarda il trend, seguendo i passaggi già svolti nel punto precedente si ottiene

$$(1-B)^2\frac{1}{4}(1+B+B^2+B^3)T_t = 2c\ .$$

La relazione nella (3.8) è ancora valida, per cui la stagionalità scompare anche in $x_t$. Le differenze sostanziali emergono invece nella trasformazione delle $\varepsilon$. Infatti, indicando con $\xi_t$ il nuovo errore

$$\xi_t = \frac{1}{4}(\varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-4} + \varepsilon_{t-5}),$$

per questo si avrà una varianza di quattro volte inferiore a quella originaria, ovvero

$$E(\xi_t^2) = \frac{1}{4}\sigma_\varepsilon^2.$$

Parallelamente, la funzione di autocovarianza assume valori ridotti rispetto a quella di $y_t$:

$$\gamma_k = \begin{cases} -\frac{1}{8}\sigma_\varepsilon^2 & k=1 \\ 0 & k=2 \\ \frac{1}{16}\sigma_\varepsilon^2 & k=3 \\ -\frac{1}{8}\sigma_\varepsilon^2 & k=4 \\ \frac{1}{16}\sigma_\varepsilon^2 & k=5 \\ 0 & k>5\ . \end{cases}$$

Al contrario, la funzione di autocorrelazione $\rho_k$ è identica a quella calcolata per $y_t$. È evidente, quindi, che l'utilizzo di una media mobile di ordine 4 al posto di una somma mobile dello stesso ordine abbia un effetto spianante sulla varianza della componente erratica di una serie; la correlazione spuria introdotta dall'aggregazione degli errori ai vari ritardi è tuttavia identica nei due casi.

(c) Riassumendo i risultati ottenuti, le serie $y_t$ e $x_t$ possono essere descritte dalle seguenti relazioni

$$\begin{aligned} y_t &= 8c + \varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-4} + \varepsilon_{t-5} \\ x_t &= 2c + \frac{1}{4}(\varepsilon_t - \varepsilon_{t-1} - \varepsilon_{t-4} + \varepsilon_{t-5}). \end{aligned}$$

Vale pertanto la relazione $x_t = \frac{1}{4}y_t$ .

**3.18** Determinare la media mobile che conserva trend locali quadratrici su intervalli del tipo $[t-3, t+3]$.

□ Particolari medie mobili possono essere utilizzate quando si vuole approssimare una serie $y_t$ mediante polinomi di grado ridotto di tipo *locale*. Una serie storica è assimilabile ad un polinomio di grado $p$ ($p \leq 2m$) nell'intervallo temporale

$$(t-m, t-m+1, \cdots, t, \cdots, t+m)$$

se approssimando la serie su ciascuno di questi intervalli con un polinomio di grado $p$ stimato con il metodo dei minimi quadrati il valore stimato è uguale al valore osservato nel punto centrale dell'intervallo, cioè $y_t^* = y_t$ (Di Fonzo e Lisi, 2005, p. 109).

Questo esercizio richiede di trovare i coefficienti della media mobile che conserva trend quadratici su intervalli simmetrici di ampiezza 3. Indicando con $p$ il grado del polinomio e con $m$ l'ampiezza, fissiamo $p = 2$ e $m = 3$. Si tratta quindi di risolvere il problema di minimizzazione

$$\sum_{i=-3}^{3} (y_{t+i} - a_0 - a_1 i - a_2 i^2)^2, \tag{3.9}$$

ovvero l'espressione dei minimi quadrati ordinari con valori interpolati da un polinomio di secondo grado nell'intervallo $[t-3, t+3]$. Si può dimostrare che la condizione prima citata è soddisfatta per

$$y_t^* = \hat{a}_0.$$

Poiché le stime dei minimi quadrati sono funzioni lineari delle osservazioni, si ha

$$\hat{a}_0 = \sum_{i=-3}^{3} \vartheta_i y_{t+i},$$

che è una media mobile di ordine $2m+1$ con coefficienti pari a $\vartheta_i$. La soluzione del problema (3.9) è data dal seguente sistema di equazioni normali

$$\begin{cases} \hat{a}_0 \sum_{i=-3}^{3} 1 + \hat{a}_1 \sum_{i=-3}^{3} i + \hat{a}_2 \sum_{i=-3}^{3} i^2 = \sum_{i=-3}^{3} y_{t+i} \\ \hat{a}_0 \sum_{i=-3}^{3} i + \hat{a}_1 \sum_{i=-3}^{3} i^2 + \hat{a}_2 \sum_{i=-3}^{3} i^3 = \sum_{i=-3}^{3} i y_{t+i} \\ \hat{a}_0 \sum_{i=-3}^{3} i^2 + \hat{a}_1 \sum_{i=-3}^{3} i^3 + \hat{a}_2 \sum_{i=-3}^{3} i^4 = \sum_{i=-3}^{3} i^2 y_{t+i} \end{cases}$$

I termini che contengono sommatorie di potenze dispari di $i$ sono nulle, per cui il sistema si semplifica in

$$\begin{cases} \hat{a}_0 \sum_{i=-3}^{3} 1 + \hat{a}_2 \sum_{i=-3}^{3} i^2 = \sum_{i=-3}^{3} y_{t+i} \\ \hat{a}_1 \sum_{i=-3}^{3} i^2 = \sum_{i=-3}^{3} i y_{t+i} \\ \hat{a}_0 \sum_{i=-3}^{3} i^2 + \hat{a}_2 \sum_{i=-3}^{3} i^4 = \sum_{i=-3}^{3} i^2 y_{t+i} \end{cases}$$

Per determinare il valore di $\hat{a}_0$ si può risolvere il sistema dato dalla prima e dalla terza equazione

$$\begin{cases} 7\hat{a}_0 + 28\hat{a}_2 = \sum_{i=-3}^{3} y_{t+i} \\ 28\hat{a}_0 + 196\hat{a}_2 = \sum_{i=-3}^{3} i^2 y_{t+i} \end{cases}$$

da cui si deriva la soluzione

$$\hat{a}_0 = \frac{1}{21} \sum_{i=-3}^{3} (7 - i^2) y_{t+i}.$$

La media mobile ricercata è simmetrica e pertanto può essere espressa nella notazione fin qui utilizzata

$$M = \left\{ [7]; \frac{1}{21}[-2,3,6,\mathbf{7}] \right\}.$$

## Esercizi supplementari

**C3.1** In cosa consiste l'operazione di 'centratura' di una media mobile?

**C3.2** Cos'è l'effetto di Slutsky-Yule?

**C3.3** Descrivere le caratteristiche delle m.m. di Henderson, evidenziandone i possibili utilizzi.

**C3.4** Descrivere le principali caratteristiche della stima del trend-ciclo di una serie storica mensile con la procedura Census X11.

**C3.5** Derivare le condizioni che devono essere soddisfatte dai coefficienti di una media mobile di ordine 7 per conservare il trend parabolico

$$T_t = 6 + 0{,}2t^2$$

**C3.6** Indicare l'ordine e derivare i coefficienti della media mobile che si ottiene componendo una media mobile di ordine 4 con coefficienti pari a

$$\left(\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}\right)$$

ed una media mobile aritmetica semplice di ordine 4.

**C3.7** Indicare l'ordine e derivare i coefficienti della media mobile che si ottiene componendo una media mobile aritmetica semplice di ordine 3 con la media mobile

$$M_1 = \left\{ [5]; \frac{1}{4}[-3,3,\mathbf{4}] \right\}.$$

**C3.8** Indicare l'ordine e derivare i coefficienti della media mobile che si ottiene componendo una media mobile aritmetica semplice di ordine 3 con la media mobile

$$M_1 = \left\{ [5]; \frac{1}{13}[1,3,\mathbf{5}] \right\}.$$

**C3.9** Si considerino le tre medie mobili

$$M_0 = \frac{1}{24}(y_{t-3} + 3y_{t-2} + 5y_{t-1} + 6y_t + 5y_{t+1} + 3y_{t+2} + y_{t+3})$$
$$M_1 = \frac{1}{6}(y_{t-2} + 2y_{t-1} + 2y_t + y_{t+1})$$
$$M_2 = \frac{1}{4}(y_{t-1} + y_t + y_{t+1} + y_{t+2}) \quad .$$

Mostrare che $M_0$ si può ottenere componendo $M_1$ e $M_2$.

**C3.10** Determinare i coefficienti di una media mobile a 7 termini che minimizzi il rapporto di riduzione della varianza residua e conservi un trend parabolico. Successivamente si confronti tale rapporto con quello della media mobile aritmetica semplice dello stesso ordine.

**C3.11** È data la seguente serie storica:

$$y_t = \{10{,}13{,}11{,}9{,}12{,}8{,}11{,}14{,}10{,}9\}.$$

Si procede al lisciamento di tale serie applicando separatamente una media mobile aritmetica semplice di ordine 3 ($M_1$) e la media mobile centrata simmetrica

$$M_2 = \left\{ [3]; \frac{1}{4}[1{,}\mathbf{2}] \right\}.$$

Stabilire quale tra le due medie mobili liscia maggiormente la serie storica.

**C3.12** Sia data la serie storica $\{y_t\}_{t=1}^{20}$

| *Anno* | *I trim* | *II trim* | *III trim* | *IV trim* |
|---|---|---|---|---|
| 1 | 105,6 | 103,5 | 88,3 | 100,1 |
| 2 | 100,1 | 101,3 | 89,0 | 100,7 |
| 3 | 99,3 | 98,6 | 88,6 | 100,7 |
| 4 | 100,0 | 101,4 | 88,6 | 99,4 |
| 5 | 96,1 | 100,3 | 87,9 | 97,8 |

Si derivi la serie destagionalizzata secondo il metodo delle medie mobili nell'ipotesi di un modello di tipo additivo.

# 4
# Il lisciamento esponenziale

### Il lisciamento esponenziale semplice

**4.1** Data la serie storica $\{y_t\}_{t=1}^{20}$

| $t$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $y_t$ | 0,955 | 0,594 | 1,366 | 2,148 | 1,511 |
| $t$ | 6 | 7 | 8 | 9 | 10 |
| $y_t$ | 1,488 | 0,823 | -0,725 | -0,701 | 0,722 |
| $t$ | 11 | 12 | 13 | 14 | 15 |
| $y_t$ | 1,130 | 2,696 | 1,211 | -1,564 | -2,553 |
| $t$ | 16 | 17 | 18 | 19 | 20 |
| $y_t$ | -1,842 | -0,828 | -0,330 | 0,428 | -0,187 |

(a) prevedere $y_{21}$ e $y_{22}$ mediante il lisciamento esponenziale semplice con $\delta = 0{,}3$;

(b) individuare il numero di ritardi necessario per ottenere la previsione di $y_{21}$ con un'approssimazione dell'ordine di $10^{-5}$.

□ Data una serie $y_t$, il metodo del lisciamento esponenziale semplice deriva una previsione $\hat{y}_{t+1}$ attraverso la seguente espressione:

$$\hat{y}_{t+1} = (1-\delta)\sum_{j=0}^{t-1} \delta^j y_{t-j} \tag{4.1}$$

dove $\delta$, detta *costante di lisciamento*, è compresa tra 0 e 1 ed $n$ è il numero di osservazioni. L'ipotesi sottostante tale formulazione è che le osservazioni più distanti dal periodo di previsione abbiano via via un'influenza minore sulla previsione stessa. La costante di lisciamento determina l'intensità del peso da attribuire all'informazione passata: maggiore il suo valore, maggiore l'influenza delle osservazioni passate sulle previsioni future.

L'espressione (4.1) può essere riscritta in forma approssimata nel seguente modo:

$$\hat{y}_{t+1} = \delta\hat{y}_t + (1-\delta)y_t. \tag{4.2}$$

L'interpretazione della costante $\delta$ è ora più evidente. In particolare, sono indicativi i casi limite dell'intervallo definito per $\delta$: per $\delta = 0$ la previsione $\hat{y}_{t+1}$ corrisponde all'ultimo valore osservato, cioè $\hat{y}_{t+1} = y_t$. La previsione, in questo caso, diventa estremamente *flessibile*. Al contrario, se $\delta = 1$ si ha $\hat{y}_{t+1} = \hat{y}_t$, ovvero la previsione resta fissa in modo rigido per ogni periodo temporale (generalmente al valore medio delle osservazioni $\bar{y}$). Per utilizzare la formula (4.2) si ha bisogno di determinare la costante $\delta$ e la stima $\hat{y}_1$ per inizializzare la procedura ricorsiva.

(a) L'esercizio fornisce un valore per $\delta$ pari a 0,3; una volta determinato $\hat{y}_1$ le previsioni richieste possono essere agevolmente calcolate utilizzando l'espressione ricorsiva nella (4.2). Come valore

iniziale si può utilizzare come detto la media campionaria delle osservazioni $y_t$: in questo caso si ottiene $\hat{y}_1 = \bar{y} = 0{,}3171$. Per il valore al tempo $t = 2$ otteniamo

$$\begin{aligned}\hat{y}_2 &= 0{,}3(\hat{y}_1) + 0{,}7(y_1) \\ &= 0{,}3(0{,}3171) + 0{,}7(0{,}955) = 0{,}764.\end{aligned}$$

Con la nuova previsione $\hat{y}_2$ si può ora derivare la previsione per il tempo $t = 3$:

$$\begin{aligned}\hat{y}_3 &= 0{,}3(\hat{y}_2) + 0{,}7(y_2) \\ &= 0{,}3(0{,}764) + 0{,}7(0{,}594) = 0{,}645.\end{aligned}$$

La seguente tabella mostra le stime ottenute secondo la (4.2) usando le osservazioni da $t = 1$ a $t = 19$.

| $t$ | $y_t$ | $\hat{y}_t$ | $t$ | $y_t$ | $\hat{y}_t$ |
|---|---|---|---|---|---|
| 1 | 0,955 | 0,317 | 11 | 1,130 | 0,340 |
| 2 | 0,594 | 0,764 | 12 | 2,696 | 0,893 |
| 3 | 1,366 | 0,645 | 13 | 1,211 | 2,155 |
| 4 | 2,148 | 1,150 | 14 | -1,564 | 1,494 |
| 5 | 1,511 | 1,849 | 15 | -2,553 | -0,647 |
| 6 | 1,488 | 1,612 | 16 | -1,842 | -1,981 |
| 7 | 0,823 | 1,525 | 17 | -0,828 | -1,884 |
| 8 | -0,725 | 1,034 | 18 | -0,330 | -1,145 |
| 9 | -0,701 | -0,197 | 19 | 0,428 | -0,574 |
| 10 | 0,722 | -0,550 | 20 | -0,187 | 0,127 |

Ora, per prevedere il valore per $t = 21$ si fa uso dell'osservazione del periodo immediatamente precedente ($y_{20}$) e della sua previsione fatta al tempo 19 ($\hat{y}_{20}$):

$$\hat{y}_{21} = 0{,}3(0{,}127) + 0{,}7(-0{,}187) = -0{,}093.$$

La previsione al tempo 22 è identica a $\hat{y}_{21}$: infatti, in assenza di nuove informazioni l'ultima previsione calcolata rimane fissa per ogni orizzonte previsivo. Seguendo il metodo del lisciamento esponenziale semplice vale quindi l'identità

$$\hat{y}_{21+h} = \hat{y}_{21} = -0{,}093 \quad \text{per ogni } h > 0.$$

(b) Nel punto precedente abbiamo utilizzato la forma approssimata (4.2) per ottenere i valori previsti $\hat{y}_{t+1}$. La formula (4.1) richiede infatti una maggiore complessità di calcolo, implicando ad ogni istante temporale l'utilizzo di tutti i termini della serie fino allora osservati. L'esercizio ora richiede di verificare quanti termini è necessario includere nella formula (4.2) per ottenere una previsione $\hat{y}_{21}$ con un approssimazione pari a $10^{-5}$. Indicando con $\hat{y}_{21}^k$ la previsione che considera come ultimo termine l'osservazione 20-$k$,

$$\hat{y}_{21}^k = (1-\delta)\sum_{j=0}^{k} \delta^j y_{20-j} \tag{4.3}$$

si chiede di determinare il valore $k$ tale per cui

$$\hat{y}_{21}^k - \hat{y}_{21} = \hat{y}_{21}^k - 0{,}092717 < 0{,}00001,$$

dove si è utilizzata un'approssimazione più fine del valore in precedenza previsto $\hat{y}_{21}$. La seguente tabella mostra le previsioni approssimate $\hat{y}_{21}^k$, per $k = 0{,}1, \cdots, 9$, la previsione esatta $\hat{y}_{21}$ e la differenza fra le due misure $e_{21}^k$. Nella tabella si riporta nella prima colonna anche l'indice temporale corrispondente alla prima osservazione della serie storica $y_t$ considerata nella formula (4.3).

| $t$ | $k$ | $\hat{y}_{21}^k$ | $\hat{y}_{21}$ | $e_{21}^k$ |
|---|---|---|---|---|
| 20 | 0 | -0,130900 | -0,092717 | 0,038183 |
| 19 | 1 | -0,041020 | -0,092717 | -0,051697 |
| 18 | 2 | -0,061810 | -0,092717 | -0,030907 |
| 17 | 3 | -0,077459 | -0,092717 | -0,015258 |
| 16 | 4 | -0,087903 | -0,092717 | -0,004814 |
| 15 | 5 | -0,092246 | -0,092717 | -0,000471 |
| 14 | 6 | -0,093044 | -0,092717 | 0,000327 |
| 13 | 7 | -0,092859 | -0,092717 | 0,000141 |
| 12 | 8 | -0,092735 | -0,092717 | 0,000018 |
| 11 | 9 | -0,092719 | -0,092717 | 0,000002 |

Come si nota, l'errore $e_{21}^k$ tende a ridursi al crescere del valore $k$, e si annulla se si considerano tutte le osservazioni disponibili. L'approssimazione risulta inferiore a 0,00001 per $k = 9$. Dunque, per commettere un errore inferiore a $10^{-5}$ rispetto alla previsione $\hat{y}_{21}$ è necessario considerare le ultime 10 osservazioni disponibili nella formula (4.3), corrispondenti ai periodi $t = 11, \cdots, 20$. Il valore di $k$ risulta quindi pari a 9. Il vantaggio nell'utilizzare l'espressione (4.2) al posto della (4.1) è evidente: tale formula infatti è una media ponderata di due soli valori, mentre nel secondo caso i calcoli coinvolgono ad ogni passaggio un maggior numero di osservazioni passate.

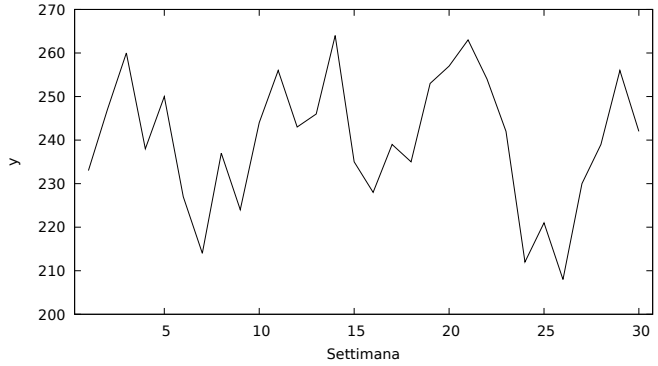
**4.2** Data la serie storica $\{y_t\}_{t=1}^{30}$

| settimana | litri | settimana | litri | settimana | litri |
|---|---|---|---|---|---|
| 1 | 233 | 11 | 256 | 21 | 263 |
| 2 | 247 | 12 | 243 | 22 | 254 |
| 3 | 260 | 13 | 246 | 23 | 242 |
| 4 | 238 | 14 | 264 | 24 | 212 |
| 5 | 250 | 15 | 235 | 25 | 221 |
| 6 | 227 | 16 | 228 | 26 | 208 |
| 7 | 214 | 17 | 239 | 27 | 230 |
| 8 | 237 | 18 | 235 | 28 | 239 |
| 9 | 224 | 19 | 253 | 29 | 256 |
| 10 | 244 | 20 | 257 | 30 | 242 |

riferita ai consumi settimanali di benzina, in litri, per le ambulanze di un pronto soccorso di un ospedale del Veneto, calcolare la previsione del valore $y_{31}$ mediante lisciamento esponenziale semplice usando la costante di lisciamento $\delta = 0{,}1$ e la condizione iniziale $\hat{y}_1 = \bar{y}$. Ripetere poi l'esercizio utilizzando come condizione iniziale $\hat{y}_1 = y_1$ e fornire un'interpretazione dei risultati.

□ Le caratteristiche della serie fornita dall'esercizio si comprendono meglio attraverso il grafico rispetto all'asse temporale, mostrato nella Figura 4.1. Nel periodo considerato il consumo di benzina non mostra alcun trend, ma manifesta un andamento stazionario con una variabilità che dipende dal numero e dalla dislocazione sul territorio delle chiamate ricevute. Il metodo del lisciamento esponenziale può essere considerato come una valida soluzione per la previsione del valore successivo all'ultimo periodo disponibile. L'esercizio richiede infatti di calcolare la previsione del valore $y_{31}$ con il parametro $\delta = 0{,}1$ e la condizione iniziale $\hat{y}_1 = \bar{y}$, ovvero il valor medio di tutte le osservazioni fornite.

Figura 4.1
Consumo settimanale di benzina in un pronto soccorso




Per ottenere la previsione $\hat{y}_{31}$ conviene, come detto nel precedente esercizio, applicare ricorsivamente la formula (4.2). È prioritario il calcolo dei valori "perequati" $\hat{y}_t$ nell'intervallo $t = 1, \cdots, 30$, considerando come condizione iniziale la media delle osservazioni nell'intero periodo. Considerato che $\bar{y} = 239{,}9$, abbiamo:

$$\begin{aligned}
\hat{y}_1 &= \bar{y} = 239{,}9 \\
\hat{y}_2 &= 0{,}1(239{,}9) + 0{,}9(233) = 233{,}69 \\
\hat{y}_3 &= 0{,}1(233{,}69) + 0{,}9(247) = 245{,}67 \\
\hat{y}_4 &= 0{,}1(245{,}67) + 0{,}9(260) = 258{,}57 \\
\hat{y}_5 &= 0{,}1(258{,}57) + 0{,}9(238) = 240{,}06 \\
&\vdots \\
\hat{y}_{29} &= 0{,}1(227{,}92) + 0{,}9(239) = 237{,}89 \\
\hat{y}_{30} &= 0{,}1(237{,}89) + 0{,}9(256) = 254{,}19.
\end{aligned}$$

La previsione per il valore $y_{31}$ è quindi data da

$$\hat{y}_{31} = 0{,}1(254{,}19) + 0{,}9(242) = 243{,}22.$$

In aggiunta si richiede di calcolare la stessa previsione usando come condizione iniziale $\hat{y}_1 = y_1$. La

nuova sequenza dei valori perequati è la seguente

$$\begin{aligned}
\hat{y}_1 &= y_1 = 233 \\
\hat{y}_2 &= 0{,}1(233) + 0{,}9(233) = 233 \\
\hat{y}_3 &= 0{,}1(233) + 0{,}9(247) = 245{,}60 \\
\hat{y}_4 &= 0{,}1(245{,}60) + 0{,}9(260) = 258{,}56 \\
\hat{y}_5 &= 0{,}1(258{,}56) + 0{,}9(238) = 240{,}06 \\
&\vdots \\
\hat{y}_{29} &= 0{,}1(227{,}92) + 0{,}9(239) = 237{,}89 \\
\hat{y}_{30} &= 0{,}1(237{,}89) + 0{,}9(256) = 254{,}19.
\end{aligned}$$

Come si nota, questi valori sono identici ai precedenti (alla seconda cifra decimale) a partire dal valore $\hat{y}_5$; ovviamente il valore previsto $\hat{y}_{31}$ sarà sempre pari a quello ottenuto in precedenza (243,22). In generale, una diversa condizione iniziale può modificare i valori perequati per le prime osservazioni ma da un certo punto in poi i risultati tendono a coincidere. Tale proprietà dipende soprattutto dal valore della costante di lisciamento $\delta$: a parità di condizioni, un minore valore di $\delta$ riduce l'influenza della condizione iniziale sul valore previsto.

**4.3** Data la seguente serie trimestrale del Prodotto Interno Lordo (PIL) italiano, espressa in milioni di euro concatenati con anno di riferimento 2000

| *Trimestre* | PIL |
|---|---|
| 2000:1 | 294.417 |
| 2000:2 | 296.582 |
| 2000:3 | 299.197 |
| 2000:4 | 301.724 |
| 2001:1 | 304.069 |
| 2001:2 | 303.364 |
| 2001:3 | 302.799 |
| 2001:4 | 302.234 |

calcolare la previsione per i quattro trimestri del 2002 utilizzando:

- l'interpolante lineare $y_s = 300.548 + 599{,}11s$, ottenuto per $s = -7, -5, -3, -1{,}1{,}3{,}5{,}7$;
- il lisciamento esponenziale semplice con $\delta = 0{,}1$.

Valutare inoltre la bontà delle previsioni, mediante un opportuno indice dell'errore medio di previsione, considerati i seguenti valori realizzati per il 2002:

| *Trimestre* | PIL |
|---|---|
| 2002:1 | 302.484 |
| 2002:2 | 303.940 |
| 2002:3 | 305.003 |
| 2002:4 | 305.196 |

□ L'esercizio richiede di derivare previsioni dei valori trimestrali del 2002 della serie del PIL italiano impiegando il metodo del lisciamento esponenziale semplice (con $\delta = 0{,}1$) ed il modello di interpolazione lineare

$$y_s = 300.548 + 599{,}11s \tag{4.4}$$

per $s = -7, -5, \cdots, 5, 7$.

Le previsioni secondo il lisciamento esponenziale semplice si ottengono dall'applicazione della (4.2), fissando la costante $\delta$ al valore indicato $0{,}1$. Per quanto riguarda la specificazione del valore iniziale $\hat{y}_1$ si può porre semplicemente $\hat{y}_1 = y_1$: il basso livello di $\delta$ rende infatti ininfluente le osservazioni lontane per le previsioni ai tempi richiesti.

Impiegando l'espressione (4.2) in modo ricorsivo si ottiene la sequenza di valori:

$$\begin{aligned}
\hat{y}_1 &= 294.417 \\
\hat{y}_2 &= 0{,}1(294.417) + 0{,}9(294.417) = 294.417 \\
\hat{y}_3 &= 0{,}1(294.417) + 0{,}9(296.582) = 296.365{,}5 \\
\hat{y}_4 &= 0{,}1(296.365{,}5) + 0{,}9(299.197) = 298.913{,}9 \\
&\cdots \\
\hat{y}_9 &= 0{,}1(303.543{,}8) + 0{,}9(302.234) = 302.296{,}6 \\
\hat{y}_{9+h} &= 302.296{,}6
\end{aligned}$$

per ogni valore $h$ positivo. Ipotizzando di non conoscere le osservazioni trimestrali del 2002, la previsione per il primo trimestre ($t = 9$) rimane costante per ogni orizzonte temporale.

Nel caso dell'interpolante lineare si tratta di applicare l'equazione del modello (4.4) ai valori $s = 9{,}11{,}13{,}15$, tempi che definiscono la sequenza dei quattro trimestri del 2002. Ad esempio, per il primo trimestre si ha

$$\hat{y}_9 = 300.348 + 599{,}11(9) = 305.940{,}2.$$

La tabella seguente riassume i risultati ottenuti dai due metodi di previsione. Nella seconda colonna sono indicati i valori osservati nei trimestri del 2002 ($y_t$), mentre nella terza e quarta colonna troviamo i valori previsti con il modello lineare ($\hat{y}_{1t}$) e con il metodo del lisciamento esponenziale semplice ($\hat{y}_{2t}$). Inoltre, nelle ultime due colonne sono riportati gli errori commessi dai due modelli, calcolati come differenza fra i valori osservati e quelli previsti.

| *Trimestre* | $y_t$ | $\hat{y}_{1t}$ | $\hat{y}_{2t}$ | $e_{1t}$ | $e_{2t}$ |
|---|---|---|---|---|---|
| 2002:1 | 302.484 | 305.940,2 | 302.296,6 | 3.456,2 | -187,4 |
| 2002:2 | 303.940 | 307.138,4 | 302.296,6 | 3.198,4 | -1643,4 |
| 2002:3 | 305.003 | 308.336,6 | 302.296,6 | 3.333,6 | -2706,4 |
| 2002:4 | 305.196 | 309.534,9 | 302.296,6 | 4.338,9 | -2899,4 |

Per valutare l'accuratezza complessiva delle quattro previsioni si può utilizzare l'errore assoluto medio di previsione, definito dalla formula

$$EAM = \frac{1}{n}\sum_{t=1}^{n} |e_t| \tag{4.5}$$

dove $e_t = \hat{y}_t - y_t$ è la differenza fra il valore previsto e quello vero, mentre in questo caso $n$ è il numero di previsioni effettuate. Per i nostri due modelli di previsione otteniamo

$$EAM_1 = \frac{1}{4}\sum_{t=1}^{4} |e_{1,t}| = 3581{,}8$$
$$EAM_2 = \frac{1}{4}\sum_{t=1}^{4} |e_{2,t}| = 1859{,}2.$$

Pertanto, in media l'errore commesso attraverso il lisciamento esponenziale è circa la metà del metodo interpolante. Infine, si fa presente che nel caso dell'interpolante lineare le previsioni effettuate sono tutte superiori ai valori osservati, mentre con il lisciamento esponenziale le previsioni sono sistematicamente più basse dei veri valori. Entrambi i metodi sembrano quindi soffrire di una *distorsione sistematica*. Il lisciamento esponenziale perde aderenza con l'aumentare dell'orizzonte previsivo, in quanto poco adatto a serie evolutive con forti componenti di trend. D'altro canto, l'utilizzo di un trend deterministico attraverso l'interpolante lineare ha prodotto un effetto distorsivo di segno opposto e di maggiore intensità.

Questa situazione è segnalata, per entrambi i metodi, dall'uguaglianza (in valore assoluto) tra l'errore assoluto medio di previsione $EAM$ e l'errore medio $EM$, definito dalla formula

$$EM = \frac{1}{n}\sum_{t=1}^{n} e_t. \tag{4.6}$$

Infatti si ha

$$EM_1 = \frac{1}{4}\sum_{t=1}^{4} e_{1,t} = 3581{,}8$$
$$EM_2 = \frac{1}{4}\sum_{t=1}^{4} e_{2,t} = -1859{,}2.$$

Affinché gli errori siano privi di sistematicità $EM$ deve essere invece di entità ridotta rispetto ad $EAM$.

**4.4** Per lisciare la serie $\{y_t\}_{t=1}^{10}$ si decide di applicare la media mobile simmetrica

$$M = \left\{ [3]; \frac{1}{5}[1{,}3] \right\}.$$

Si estrapoli il valore di $y_9$ mediante lisciamento esponenziale con $\delta = 0{,}1$ (per semplicità arrestato al terzo termine) e, per $y_8$, si calcoli il valore perequato con la media mobile $M$. Mostrare che questa procedura è equivalente ad applicare una media mobile asimmetrica ad $y_8$ e determinare i coefficienti di tale media mobile.

□ L'applicazione di una media mobile simmetrica di ordine 3 coinvolge tre termini di una serie storica per ciascuna osservazione. Per il primo e l'ultimo valore osservato non è possibile derivare il valore perequato, con la spiacevole conseguenza di ottenere una serie più corta di quella originale. Per risolvere tale problema si impiegano comunemente dei valori previsti in sostituzione di quelli mancanti ottenuti

sulla base di appropriate tecniche statistiche di previsione. Tra di esse rientrano i metodi del lisciamento esponenziale: in questo esercizio si richiede appunto di impiegare il lisciamento semplice per ottenere una previsione dell'osservazione $y_9$ e, successivamente, utilizzare questo valore per calcolare il valore perequato di $y_8$ attraverso la media mobile $M$.

Il valore previsto $\hat{y_9}$ si ottiene applicando la formula (4.1). La costante di lisciamento è fissata a 0,1 ($\delta = 0.1$), mentre si richiede l'utilizzo dei tre termini più vicini a ciascuna osservazione per il calcolo dei valori perequati ($n = 3$). Sostituendo all'espressione (4.1) abbiamo

$$\hat{y}_9 \simeq 0{,}9 \sum_{j=0}^{2} 0{,}1^j y_{8-j}$$

per cui un'approssimazione del valore previsto è data da

$$\hat{y}_9 = 0{,}9(y_8 + 0{,}1y_7 + 0{,}01y_6) \tag{4.7}$$

Ora, la media mobile $M$ può essere applicata anche per l'ultima osservazione $y_8$. Il valore perequato è

$$y_8^* = My_8 = \frac{1}{5}(y_7 + 3y_8 + \hat{y_9});$$

sostituendo a $\hat{y_9}$ l'espressione (4.7) si ottiene

$$\begin{aligned} y_8^* &= \frac{1}{5}[y_7 + 3y_8 + 0{,}9(y_8 + 0{,}1y_7 + 0{,}01y_6)] \\ &= \frac{1}{5}(0{,}009y_6 + 1{,}09y_7 + 3{,}9y_8) \\ &= 0{,}0018y_6 + 0{,}218y_7 + 0{,}78y_8 \end{aligned}$$

che corrisponde ad una media mobile asimmetrica degli ultimi tre termini della serie, con un forte peso assegnato all'ultimo valore $y_8$.

Pertanto, si è dimostrato che l'utilizzo di previsioni per sostituire i valori mancanti necessari per applicare una media mobile simmetrica equivale nella sostanza all'impiego di una media mobile asimmetrica delle osservazioni passate. Si noti che, data l'approssimazione utilizzata, non si ottiene esattamente una media mobile poichè la somma dei coefficienti non è pari ad 1, ma a 0,9998. Ovviamente, tale discrepanza diventa via via minore quanto più il numero $n$ di termini utilizzato nella (4.1) si avvicina al numero complessivo di osservazioni della serie storica.

**4.5** Si consideri la serie storica $\{y_t\}_{t=1}^{24}$ riportata nella tabella seguente:

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $y_t$ | 3,40 | 3,34 | 2,88 | 3,22 | 2,90 | 3,52 | 4,31 | 5,03 |
| $t$ | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| $y_t$ | 4,13 | 3,70 | 3,12 | 3,50 | 3,15 | 3,79 | 4,74 | 3,94 |
| $t$ | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| $y_t$ | 3,34 | 3,81 | 3,15 | 4,56 | 5,01 | 4,95 | 4,23 | 5,05 |

Per tale serie stimare il valore di $\delta$ necessario per l'applicazione del metodo del lisciamento esponenziale semplice seguendo il metodo dei minimi quadrati. Verificare poi la stima ottenuta con quella fornita da EViews.

□ Generalmente, la costante di lisciamento $\delta$ è fissata in un intervallo compreso tra 0,05 e 0,3: un valore superiore renderebbe troppo rigida la previsione dei valori futuri rispetto al valore delle osservazioni più recenti. Per evitare una scelta soggettiva si può affontare la stima di $\delta$ in un modo statisticamente più fondato. Un metodo semplice è quello dei minimi quadrati, secondo il quale il valore $\hat{\delta}$ ottimale è quello che rende minima la funzione

$$S(\delta) = \sum_{t=m}^{n-1} (y_{t+1} - \hat{y}_{t+1})^2 = \sum_{t=m}^{n-1} \left[ y_{t+1} - (1-\delta) \sum_{j=0}^{t-1} \delta^j y_{t-j} \right]^2 \tag{4.8}$$

avendo indicato con $m$ il numero di termini da trascurare per evitare gli effetti di un valore iniziale $\hat{y}_1$ predeterminato. Operativamente, si considera una griglia di valori possibili per $\delta$ e per ciascuno di essi si calcola la formula (4.8). Il valore di $\delta$ a cui è associato il minimo valore $S(\delta)$ equivale alla stima dei minimi quadrati.

A scopo esemplificativo utilizziamo la seguente griglia di valori

$$\delta = (0{,}1; 0{,}2; 0{,}3; 0{,}4; 0{,}5; 0{,}6; 0{,}7; 0{,}8; 0{,}9).$$

Per ciascuno dei valori selezionati si è applicato il lisciamento esponenziale semplice attraverso la formula (4.2) e calcolata la somma del quadrato degli scarti tra valori osservati e valori perequati $S(\delta)$. Si è utilizzato un valore di $m$ pari a 5. La tabella seguente mostra i valori di $S(\delta)$ ottenuti.

| $\delta$ | 0,1 | 0,2 | 0,3 | 0,4 | 0,5 |
|---|---|---|---|---|---|
| $S(\delta)$ | 9,364 | 9,350 | 9,408 | 9,514 | 9,641 |
| $\delta$ | 0,6 | 0,7 | 0,8 | 0,9 | |
| $S(\delta)$ | 9,779 | 9,950 | 10,278 | 11,353 | |

I risultati indicano quindi che il valore di $\delta$ che minimizza $S(\delta)$ è $\hat{\delta} = 0{,}2$. Per ottenere una stima più precisa di $\delta$ si può utilizzare una griglia più fitta di valori, ad esempio con passo 0,01. Ovviamente, questo aumenta notevolmente la quantità di calcoli necessari per coprire l'intervallo di valori possibili per $\delta$. Si può allora ricorrere ad EViews, in grado di stimare automaticamente il valore di $\delta$ che minimizza l'errore quadratico medio di previsione. In realtà, il metodo del lisciamento esponenziale in EViews segue l'impostazione alternativa

$$\hat{y}_{t+1} = (1-\alpha)\hat{y}_t + \alpha y_t$$

dove $\alpha$ corrisponde al parametro $1-\delta$ delle espressioni (4.1) e (4.2). La stima di $\alpha$ fornita da EViews è pari a 0,812, quindi abbiamo

$$\hat{\delta} = 1 - 0{,}812 = 0{,}188.$$

Per tale valore la funzione $S(\delta)$ è pari a 9,347, un valore effettivamente inferiore a quello trovato in precedenza per $\hat{\delta} = 0{,}2$.

## I metodi Holt-Winters

**4.6** Date le serie storiche del PIL per i due paesi A e B

| *Anno* | PIL paese A | PIL paese B |
|---|---|---|
| 1996 | 848.213 | 1.894.597 |
| 1997 | 865.400 | 1.921.068 |
| 1998 | 881.071 | 1.960.507 |
| 1999 | 895.275 | 1.991.034 |
| 2000 | 921.370 | 2.049.794 |

prevedere per ciascun paese il tasso di crescita nel 2001 usando il lisciamento esponenziale secondo il metodo di Holt-Winters non stagionale con $\alpha = \beta = 0{,}1$.

□ Una generalizzazione del lisciamento esponenziale semplice è dato dai metodi di Holt-Winters, dal nome dei due studiosi che li hanno proposti in letteratura. Tali metodi sono adatti quando la serie da prevedere è caratterizzata da trend non costanti o fluttuazioni marcate. Infatti, i valori previsti sono funzioni di stime ottenute per ciascuna delle componenti inosservate ipotizzate (livello, trend e, nel caso di variabili infra-annuali, stagionalità).

In questo esercizio si suppone che la serie sia composta da un livello costante e da un trend crescente. Si adotta pertanto il metodo Holt-Winters non stagionale, descritto dalle seguenti due equazioni:

$$\begin{cases} \hat{L}_t = \alpha(\hat{L}_{t-1} + \hat{T}_{t-1}) + (1-\alpha)y_t, & 0 < \alpha < 1 \\ \hat{T}_t = \beta\hat{T}_{t-1} + (1-\beta)(\hat{L}_t - \hat{L}_{t-1}), & 0 < \beta < 1 \end{cases} \tag{4.9}$$

dove con $\hat{L}_t$ e $\hat{T}_t$ si indicano le stime delle componenti di livello e di trend di $y_t$ per $t = 1, \cdots, n$.

A differenza del lisciamento esponenziale semplice vi è quindi una doppia ricorsione, una riguardante il livello della serie ed una riguardante la tendenza della stessa. Si intuisce comunque che la logica sottostante entrambe le equazioni di aggiornamento è la stessa utilizzata nel lisciamento esponenziale semplice: le due espressioni sono medie ponderate fra due informazioni, una stima effettuata al tempo precedente e l'osservazione che "entra" al tempo corrente. Le informazioni richieste per inizializzare le equazioni (4.9) riguardano le quantità $\hat{L}_2$ e $\hat{T}_2$. A questo scopo si assume in genere

$$\hat{L}_2 = y_2$$
$$\hat{T}_2 = y_2 - y_1.$$

Per i due paesi abbiamo calcolato le equazioni per $\hat{L}_t$ e $\hat{T}_t$ per $t = 2, 3, 4, 5$, corrispondenti agli anni dal 1997 al 2000. La previsione ad un passo della serie, che indichiamo qui con $\hat{y}_{t+1}$, è data dalla somma delle stime del livello e del trend al tempo $t$, ovvero

$$\hat{y}_{t+1} = \hat{L}_t + \hat{T}_t.$$

Le due tabelle seguenti riassumono i risultati ottenuti per i due paesi e derivano nell'ultima riga la previsione richiesta per l'anno 2001.

| *Anno* | $t$ | $y_t^A$ | $\hat{L}_t^A$ | $\hat{T}_t^A$ | $\hat{y}_{t+1\|t}^A$ |
|---|---|---|---|---|---|
| 1996 | 1 | 848.213 | | | |
| 1997 | 2 | 865.400 | 865.400 | 17.187 | |
| 1998 | 3 | 881.071 | 881.223 | 15.959 | 882.587 |
| 1999 | 4 | 895.275 | 895.466 | 14.415 | 897.182 |
| 2000 | 5 | 921.370 | 920.221 | 23.721 | 909.880 |
| 2001 | 6 | | | | 943.942 |

| *Anno* | $t$ | $y_t^B$ | $\hat{L}_t^B$ | $\hat{T}_t^B$ | $\hat{y}_{t+1\|t}^B$ |
|---|---|---|---|---|---|
| 1996 | 1 | 1.894.597 | | | |
| 1997 | 2 | 1.921.068 | 1.921.068 | 26.471 | |
| 1998 | 3 | 1.960.507 | 1.959.210 | 36.975 | 1.947.539 |
| 1999 | 4 | 1.991.034 | 1.991.549 | 32.803 | 1.996.185 |
| 2000 | 5 | 2.049.794 | 2.047.250 | 53.411 | 2.024.352 |
| 2001 | 6 | | | | 2.100.661 |

Il tasso di crescita (espresso in %) per i due paesi nel 2001 previsto dal metodo Holt-Winters si ottiene dal confronto con i corrispettivi dati osservati nel 2000:

$$t^A = \frac{943.942 - 909.880}{909.880} \times 100 = 2{,}450$$
$$t^B = \frac{2.100.661 - 2.049.794}{2.049.794} \times 100 = 2{,}482.$$

Si ricordi infine che per previsioni a più passi in avanti ($k > 1$) la formula di previsione da utilizzare è

$$\hat{y}_{t+k} = \hat{L}_t + k\hat{T}_t,$$

dove si tiene esplicitamente conto della tendenza evolutiva della componente di trend al crescere dell'orizzonte previsivo $k$. I valori previsti si pongono quindi su una linea retta avente intercetta $\hat{L}_n$ e coefficiente angolare $\hat{T}_n$.

**4.7** Si consideri la serie trimestrale del tasso di disoccupazione nel Veneto (1992.4-2001.2, fonte:ISTAT) riportata nella seguente tabella

| *Anno* | *Trim.* | $y_t$ | *Anno* | *Trim.* | $y_t$ |
|---|---|---|---|---|---|
| 1992 | 4 | 5,32 | 1997 | 1 | 5,93 |
| 1993 | 1 | 5,09 | | 2 | 4,84 |
| | 2 | 5,50 | | 3 | 4,35 |
| | 3 | 5,22 | | 4 | 5,42 |
| | 4 | 5,62 | 1998 | 1 | 5,56 |
| 1994 | 1 | 6,66 | | 2 | 5,14 |
| | 2 | 6,03 | | 3 | 4,06 |
| | 3 | 5,57 | | 4 | 5,15 |
| | 4 | 6,43 | 1999 | 1 | 5,03 |
| 1995 | 1 | 6,24 | | 2 | 4,82 |
| | 2 | 5,87 | | 3 | 4,21 |
| | 3 | 4,97 | | 4 | 4,07 |
| | 4 | 5,18 | 2000 | 1 | 4,19 |
| 1996 | 1 | 5,67 | | 2 | 4,14 |
| | 2 | 5,36 | | 3 | 3,20 |
| | 3 | 4,87 | | 4 | 3,52 |
| | 4 | 5,79 | 2001 | 1 | 3,73 |
| | | | | 2 | 3,70 |

Derivare le previsioni per il terzo e quarto trimestre del 2001 con un modello Holt-Winters di tipo additivo, con parametri $\alpha = 0{,}3$, $\beta = 0{,}8$ e $\gamma = 0{,}1$. Si applichi poi il modello di tipo moltiplicativo e si confrontino i risultati ottenuti.

□ Le serie del tasso di disoccupazione in Veneto è rappresentata nella Figura 4.2. La serie mostra un chiaro trend discendente ed una componente stagionale poco regolare nel periodo esaminato; il ricorso ad una procedura di tipo Holt-Winters stagionale risulta pertanto appropriato. Il modello stagionale additivo suppone che, nelle vicinanze di $n$, la serie storica $\{y_t\}_{t=1}^n$ sia (localmente) esprimibile come:

$$L_n + T_n(t-n) + S_n$$

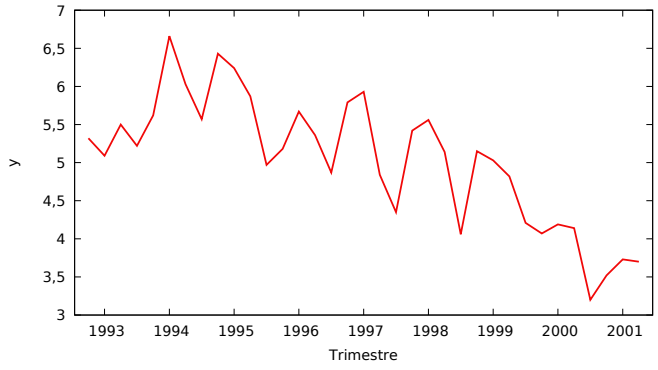
dove $S_n$ è un fattore stagionale di periodo pari a $s$, mentre $L_n$ e $T_n$ sono le componenti di livello e trend già introdotte nel precedente esercizio. La stima delle tre componenti si ottiene dalle seguente tre formule di aggiornamento:

$$\begin{cases} \hat{L}_t = \alpha(\hat{L}_{t-1} + \hat{T}_{t-1}) + (1-\alpha)(y_t - S_{t-s}) \\ \hat{T}_t = \beta\hat{T}_{t-1} + (1-\beta)(\hat{L}_t - \hat{L}_{t-1}) \\ \hat{S}_t = \gamma\hat{S}_{t-s} + (1-\gamma)(y_t - \hat{L}_t). \end{cases} \tag{4.10}$$

con $0 < \alpha, \beta, \gamma < 1$. La scelta del valore da assegnare ai parametri si basa sugli stessi ragionamenti visti finora; si noti che ciascuna componente è derivata come media ponderata della stima compiuta al passo precedente e di una funzione dell'osservazione $y_t$ nel periodo corrente. In questo esercizio i valori dei parametri sono noti e pari a $\alpha = 0{,}3$, $\beta = 0{,}8$ e $\gamma = 0{,}1$. Una volta ottenute le stime delle componenti al tempo $t$ la previsione del valore $y_{t+k}$ si ottiene dalle espressioni

$$\begin{aligned} \hat{y}_{t+k} &= \hat{L}_t + \hat{T}_t k + \hat{S}_{t+k-s}, \quad \text{per } 1 \le k \le s \\ &= \hat{L}_t + \hat{T}_t k + \hat{S}_{t+k-2s}, \quad \text{per } s+1 \le k \le 2s \\ &\vdots \end{aligned}$$

Figura 4.2
Tasso di disoccupazione in Veneto




Anche per questa variante è necessario impostare alcune condizioni iniziali per avviare la procedura ricorsiva definita dalle espressioni (4.10). Per rendere i riferimenti temporali più agevoli utilizziamo l'usuale indice temporale $t = 1, \cdots, 35$, dove $t = 1$ individua il quarto trimestre del 1992. Le seguenti

espressioni definiscono le condizioni iniziali:

$$\begin{aligned}
\hat{L}_3 &= \frac{1}{8}y_1 + \frac{1}{4}y_2 + \frac{1}{4}y_3 + \frac{1}{4}y_4 + \frac{1}{8}y_5 = 5{,}320 \\
\hat{L}_4 &= \frac{1}{8}y_2 + \frac{1}{4}y_3 + \frac{1}{4}y_4 + \frac{1}{4}y_5 + \frac{1}{8}y_6 = 5{,}554 \\
\hat{T}_4 &= \hat{L}_4 - \hat{L}_3 = 0{,}234 \\
\hat{S}_4 &= y_4 - \hat{L}_4 = -0{,}334 \\
\hat{S}_3 &= y_3 - \hat{L}_3 = 0{,}180 \\
\hat{S}_2 &= y_2 - \hat{L}_3 + \hat{T}_4 = 0{,}004 \\
\hat{S}_1 &= y_1 - \hat{L}_3 + 2\hat{T}_4 = 0{,}468
\end{aligned}$$

Una volta fissati i valori iniziali, si può iniziare la procedura di aggiornamento delle componenti a partire da $t = 5$ secondo le espressioni

$$\begin{cases}
\hat{L}_5 = 0{,}3(\hat{L}_4 + \hat{T}_4) + 0{,}7(y_5 - S_1) = 5{,}343 \\
\hat{T}_5 = 0{,}8\hat{T}_4 + 0{,}2(\hat{L}_5 - \hat{L}_4) = 0{,}145 \\
\hat{S}_5 = 0{,}1\hat{S}_1 + 0{,}9(y_5 - \hat{L}_5) = 0{,}296.
\end{cases}$$

Trovati questi valori la previsione per il valore $y_6$ è uguale a

$$\hat{y}_6 = \hat{L}_5 + \hat{T}_5 + \hat{S}_2 = 5{,}492.$$

La Tabella 4.1 mostra le stime ottenute per alcuni istanti temporali all'inizio ed alla fine del periodo osservato. Nell'ultima colonna è riportata la previsione del periodo corrispondente alla riga rispetto all'informazione posseduta fino al periodo precedente. L'esercizio richiede di calcolare le previsioni per il terzo e quarto trimestre del 2001, corrispondenti agli istanti temporali $t = 36, 37$. Le formule di previsione per tali valori sono:

$$\begin{aligned}
\hat{y}_{36} &= \hat{L}_{35} + \hat{T}_{35} + \hat{S}_{32} \\
&= 3{,}519 - 0{,}112 - 0{,}738 = 2{,}669 \\
\hat{y}_{37} &= \hat{L}_{35} + 2\hat{T}_{35} + \hat{S}_{33} \\
&= 3{,}519 - 0{,}224 - 0{,}087 = 3{,}207.
\end{aligned}$$

Confrontando ciascuna delle previsioni con i corrispondenti trimestri del 2000 si intuisce come il metodo estrapoli la tendenza al ribasso del tasso di disoccupazione evidente nell'ultima parte del campione: si passa da un tasso di disoccupazione del 3,20% ad uno del 2,67% previsto per il terzo trimestre e da un tasso del 3,52% al 3,21% per il quarto trimestre. L'aumento congiunturale tra i due trimestri previsti è invece attribuibile alla dinamica della componente stagionale, che evidenzia una forte riduzione della disoccupazione nel terzo trimestre dell'anno.

Un'altra versione del modello stagionale di tipo Holt-Winters è quello moltiplicativo, secondo il quale la stagionalità varia in proporzione ai mutamenti di lungo periodo della serie. In questo caso, in prossimità di $n$ la serie storica $\{y_t\}$ è localmente approssimata da

$$[L_n + T_n(t - n)]S_n.$$

Le formule di aggiornamento per le stime delle tre componenti si trasformano in

$$\begin{cases} \hat{L}_t = \alpha(\hat{L}_{t-1} + \hat{T}_{t-1}) + (1-\alpha)\dfrac{y_t}{\hat{S}_{t-s}} \\ \hat{T}_t = \beta\hat{T}_{t-1} + (1-\beta)(\hat{L}_t - \hat{L}_{t-1}) \\ \hat{S}_t = \gamma\hat{S}_{t-s} + (1-\gamma)\dfrac{y_t}{\hat{L}_t} \end{cases} \tag{4.11}$$

sempre con $0 < \alpha, \beta, \gamma < 1$.

Tabella 4.1
Tasso di disoccupazione in Veneto. Stime col metodo di Holt-Winters stagionale additivo

| $t$ | $y_t$ | $L_t$ | $T_t$ | $S_t$ | $y_{t+1}$ |
|---|---|---|---|---|---|
| 1 | 5,32 | | | 0,468 | |
| 2 | 5,09 | | | 0,004 | |
| 3 | 5,50 | 5,320 | | 0,180 | |
| 4 | 5,22 | 5,554 | 0,234 | -0,334 | |
| 5 | 5,62 | 5,343 | 0,145 | 0,296 | 6,255 |
| 6 | 6,66 | 6,306 | 0,308 | 0,319 | 5,492 |
| 7 | 6,03 | 6,079 | 0,201 | -0,026 | 6,794 |
| 8 | 5,57 | 6,017 | 0,149 | -0,436 | 5,947 |
| | | | ⋮ | | |
| 32 | 3,20 | 3,938 | -0,115 | -0,738 | 3,200 |
| 33 | 3,52 | 3,616 | -0,156 | -0,087 | 3,815 |
| 34 | 3,73 | 3,426 | -0,163 | 0,306 | 3,779 |
| 35 | 3,70 | 3,519 | -0,112 | 0,170 | 3,334 |

Di conseguenza, la formula di previsione diventa

$$\begin{aligned} \hat{y}_{t+k} &= (\hat{L}_t + \hat{T}_t k)\hat{S}_{t+k-s}, \quad \text{per } 1 \le k \le s \\ &= (\hat{L}_t + \hat{T}_t k)\hat{S}_{t+k-2s}, \quad \text{per } s+1 \le k \le 2s \end{aligned}$$

Le condizioni iniziali per il livello ed il trend della serie si ottengono come nel caso del modello con stagionalità additiva. Invece, i valori iniziali stagionali si possono ottenere dividendo le osservazioni per il trend lineare calcolato a partire da $\hat{L}_3$, $\hat{L}_4$ e $\hat{T}_4$:

$$\begin{aligned} \hat{S}_4 &= \frac{y_4}{\hat{L}_4} = 0{,}940 & \hat{S}_3 &= \frac{y_3}{\hat{L}_3} = 1{,}034 \\ \hat{S}_2 &= \frac{y_2}{\hat{L}_3 + \hat{T}_4} = 1{,}001 & \hat{S}_1 &= \frac{y_1}{\hat{L}_3 + 2\hat{T}_4} = 1{,}096 \end{aligned}$$

La Tabella 4.2 mostra i valori perequati relativi al modello stagionale moltiplicativo. Si vede chiaramente che la componente stagionale svolge un ruolo di fattore moltiplicativo del livello-trend della serie: un valore maggiore (minore) di 1 indica un picco stagionale superiore (inferiore) alla tendenza media. Dal confronto con le stime del modello additivo si nota come le serie perequate siano molto simili fra loro, il chè lascia intendere come per questa serie non sia necessario impiegare una componente stagionale di

Tabella 4.2
Tasso di disoccupazione in Veneto. Stime col metodo di Holt-Winters stagionale moltiplicativo

| $t$ | $y_t$ | $L_t$ | $T_t$ | $S_t$ | $y_{t+1}$ |
|---|---|---|---|---|---|
| 1 | 5,32 | | | 1,096 | |
| 2 | 5,09 | | | 1,001 | |
| 3 | 5,50 | 5,320 | | 1,034 | |
| 4 | 5,22 | 5,554 | 0,234 | 0,940 | |
| 5 | 5,62 | 5,325 | 0,141 | 1,056 | 6,345 |
| 6 | 6,66 | 6,298 | 0,308 | 1,051 | 5,470 |
| 7 | 6,03 | 6,065 | 0,199 | 0,996 | 6,829 |
| 8 | 5,57 | 6,028 | 0,152 | 0,927 | 5,888 |
| | | | ⋮ | | |
| 32 | 3,20 | 3,818 | -0,143 | 0,816 | 3,359 |
| 33 | 3,52 | 3,574 | -0,164 | 0,976 | 3,664 |
| 34 | 3,73 | 3,439 | -0,158 | 1,088 | 3,686 |
| 35 | 3,70 | 3,529 | -0,108 | 1,048 | 3,340 |

tipo moltiplicativo. Le previsioni per gli ultimi due trimestri del 2001 sono

$$\begin{aligned}\hat{y}_{36} &= (\hat{L}_{35} + \hat{T}_{35})\hat{S}_{32} \\ &= (3{,}529 - 0{,}108)0{,}816 = 2{,}793 \\ \hat{y}_{37} &= (\hat{L}_{35} + 2\hat{T}_{35})\hat{S}_{33} \\ &= (3{,}529 - 0{,}108)0{,}976 = 3{,}232\end{aligned}$$

leggermente più alte di quelle ottenute in precedenza. In particolare, nel terzo trimestre si ha una previsione più alta dello 0,12% (da 2,67% a 2,79%): tale aumento è in larga misura attribuibile alla stima della componente stagionale, considerato che l'incremento dovuto alla stima del livello-trend della serie è pari ad appena lo 0,01% (da 3,41% a 3,42%).

**4.8** La tabella sottostante contiene: (i) la serie delle vendite di un certo prodotto ($y_t$) fino al periodo 9, (ii) la serie perequata mediante lisciamento esponenziale ($LE_t$) con $\delta = 0{,}1$ fino al periodo 8 e (iii) la serie perequata con il metodo di Holt-Winters non stagionale ($HW_t$) con $\alpha = 0{,}1$ e $\beta = 0{,}2$, e le corrispondenti serie stimate del livello e del trend fino al periodo 8.

| $t$ | $y_t$ | $LE_t$ | $HW_t$ | $L_t$ | $T_t$ |
|---|---|---|---|---|---|
| | | $\delta = 0{,}1$ | $\alpha = 0{,}1$ e $\beta = 0{,}2$ | | |
| 1 | 105 | 105,00 | - | - | - |
| 2 | 142 | 105,00 | - | 142,00 | 37,00 |
| 3 | 123 | 138,30 | 179,00 | 128,60 | -3,32 |
| 4 | 160 | 124,53 | 125,28 | 156,53 | 21,68 |
| 5 | 195 | 156,45 | 178,21 | 193,32 | 33,77 |
| 6 | 213 | 191,15 | 227,09 | 214,41 | 23,62 |
| 7 | 250 | 210,81 | 238,03 | 248,80 | 32,24 |
| 8 | 346 | 246,08 | 281,04 | 339,50 | 79,01 |
| 9 | 474 | | | | |

(a) Calcolare i valori perequati per $t = 9$ e le previsioni per $t = 10$ con entrambi i metodi.

(b) Quale tra i due metodi fornisce le previsioni più attendibili?

□ (a) I valori perequati al tempo $t = 9$ nei due casi si ottengono banalmente dall'applicazione delle espressioni (4.2) e (4.9), rispettivamente. In questo esercizio, infatti, vengono già fornite le stime fino al tempo 8, per cui non è necessario iterare le formule di aggiornamento per i periodi passati. Sfruttando i dati nella tabella, per il lisciamento esponenziale si ha semplicemente

$$\begin{aligned} LE_9 &= \delta LE_8 + (1-\delta)y_8 \\ &= 0{,}1(246{,}08) + 0{,}9(346) = 336{,}01 \end{aligned}$$

dove, seguendo la notazione usata in questo esercizio, con $LE_t$ si indica la serie perequata secondo il lisciamento esponenziale al tempo $t$. Per il metodo Holt-Winters sappiamo che la previsione al tempo $t$ della serie è pari alla somma del livello e del trend stimati al tempo precedente $t-1$. Entrambe queste stime sono già disponibili nella tabella, quindi il valore perequato $HW_9$ è dato da

$$\begin{aligned} HW_9 &= \hat{L}_8 + \hat{T}_8 \\ &= 339{,}50 + 79{,}01 = 418{,}51. \end{aligned}$$

La previsione del valore corrispondente al periodo $t = 10$ secondo il metodo del lisciamento esponenziale si calcola ancora con la formula (4.2), sfruttando il valore perequato calcolato per il passo 9. Il valore previsto è infatti

$$LE_{10} = 0{,}1(336{,}01) + 0{,}9(474) = 460{,}20.$$

Per derivare la previsione secondo Holt-Winters si devono invece aggiornare le stime del livello e del trend al passo 9, cioè $\hat{L}_9$ e $\hat{T}_9$, tenendo in considerazione l'osservazione entrante $y_9$. Seguendo le formule di aggiornamento definite nella (4.9), le stime richieste si trovano con le seguenti espressioni:

$$\begin{aligned} \hat{L}_9 &= 0{,}1(339{,}50 + 79{,}01) + 0{,}9(474) = 468{,}45 \\ \hat{T}_9 &= 0{,}2(79{,}01) + 0{,}8(468{,}45 - 339{,}50) = 118{,}96. \end{aligned}$$

La previsione secondo il metodo Holt-Winters $HW_{10}$ è quindi data da

$$HW_{10} = \hat{L}_9 + \hat{T}_9 = 468{,}45 + 118{,}96 = 587{,}41.$$

(b) Per stabilire quale tra i due metodi fornisca previsioni più attendibili possiamo calcolare un indice sintetico sugli errori di previsione commessi. Introduciamo in questo esercizio un indice molto utilizzato in letteratura, la media quadratica degli errori di previsione $e_t$. Questo indice è definito come

$$MQE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \tag{4.12}$$

con $n$ il numero di osservazioni considerate. Questo indice ha il vantaggio di permettere una scomposizione dell'errore medio in tre differenti tipologie: l'errore sistematico, l'errore nella variabilità e l'errore nelle covarianze (si veda a tal proposito Di Fonzo e Lisi (2005), p.154-155).

Gli errori di previsione $e_t$ si ottengono dalla differenza dei valori stimati e quelli osservati. Partendo dal periodo 3, si ottiene per il lisciamento esponenziale la seguente serie di errori $e_t^{le}$

$$\begin{aligned} e_3^{le} &= LE_3 - y_3 = 138{,}30 - 123 = 15{,}30 \\ e_4^{le} &= LE_4 - y_4 = 124{,}53 - 160 = -35{,}47 \\ &\vdots \\ e_9^{le} &= LE_9 - y_9 = 336{,}01 - 474 = -137{,}99 \end{aligned}$$

e per il metodo Holt-Winters gli errori $e_t^{hw}$

$$\begin{aligned} e_3^{hw} &= HW_3 - y_3 = 179{,}00 - 123 = 56{,}00 \\ e_4^{hw} &= HW_4 - y_4 = 125{,}28 - 160 = -34{,}72 \\ &\vdots \\ e_9^{hw} &= HW_9 - y_9 = 418{,}51 - 474 = -55{,}49. \end{aligned}$$

Applicando la formula (4.12) alle due serie di errori per il periodo $t = 3, \cdots, 9$, si ottengono i seguenti indici:

$$\begin{aligned} MQE^{le} &= 69{,}71 \\ MQE^{hw} &= 41{,}86. \end{aligned}$$

Dal confronto degli indici risulta chiaro che il metodo Holt-Winters è più affidabile per perequare e prevedere questa serie storica: l'errore quadratico medio è circa il 40% inferiore a quello del lisciamento esponenziale semplice. In effetti, il valore delle osservazioni è crescente nel tempo: in questi casi è sempre meglio esplicitare una componente di trend che riesca a cogliere tale tendenza.

## Esercizi supplementari

**C4.1** Data la serie storica degli investimenti fissi netti italiani (miliardi di lire, valori a prezzi 1995)

| *Anno* | Investimenti |
|---|---|
| 1989 | 134.302 |
| 1990 | 140.674 |
| 1991 | 137.456 |
| 1992 | 126.308 |
| 1993 | 84.453 |
| 1994 | 80.443 |
| 1995 | 93.707 |
| 1996 | 100.023 |
| 1997 | 101.297 |
| 1998 | 109.663 |
| 1999 | 119.365 |
| 2000 | 134.503 |

calcolare le previsioni per il 2001 ed il 2002 utilizzando:

- l'interpolante lineare $y_t = 113.516{,}2 + 785{,}2182 \cdot t$, ottenuto per $t$ =-11,-9,-7,-5,-3,-1,1,3,5,7,9,11;
- il lisciamento esponenziale semplice con $\delta = 0{,}2$.

**C4.2** Per lisciare la serie $\{y_t\}_{t=1}^{10}$ si decide di applicare la media mobile simmetrica $M = \left\{[3]; \frac{1}{4}[1,\mathbf{2}]\right\}$. Si estrapoli il valore di $y_{11}$ mediante il lisciamento esponenziale semplice con $\delta = 0{,}3$ (per semplicità arrestato al quinto termine) e si calcoli il valore perequato $y_{10}^*$. Mostrare che questa procedura è equivalente ad applicare una media mobile asimmetrica ad $y_{10}$ e determinare tale media mobile.

# 5
# Processi stocastici

### Funzioni di autocovarianza e autocorrelazione

**5.1** Definire la funzione di autocorrelazione globale di un processo stocastico stazionario, riportare le sue proprietà e fornire l'espressione di una sua stima fatta in base alla serie storica osservata $\{y_t\}_{t=1}^n$.

□ La funzione di autocorrelazione è un indice normalizzato delle relazioni lineari esistenti tra coppie di variabili casuali componenti il processo stocastico prese a differenti *lag* temporali. Nel caso di un processo stazionario tale funzione dipende esclusivamente dallo sfasamento temporale esistente tra le due variabili e non dagli istanti a cui esse si riferiscono. La funzione di autocorrelazione globale $\rho_k$ si ottiene dalla associata funzione di autocovarianza $\gamma_k$. Infatti si ha

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad k = 0, \pm 1, \pm 2, \ldots$$

Al pari di $\gamma_k$, la funzione di autocorrelazione è una funzione pari di $k$ per cui il calcolo può essere limitato ai soli valori positivi di $k$.

La funzione di autocorrelazione soddisfa le seguenti proprietà:

$$\begin{aligned} \rho_0 &= 1 \\ \rho_k &= \rho_{-k} \\ |\rho_k| &\leq 1. \end{aligned}$$

L'ultima proprietà discende direttamente dalla disuguaglianza di Cauchy-Schwarz, secondo la quale

$$|E(Y_t Y_{t+k})| \leq E(Y_t^2).$$

La funzione di autocorrelazione caratterizza il comportamento di un processo stocastico, nel senso che ne definisce *a livello teorico* la struttura probabilistica temporale. Nelle applicazioni reali si dispone soltanto di serie storiche, intese come parti finite di realizzazioni di processi stocastici. Non è quindi possibile conoscere con certezza né il processo generatore dei dati né tantomeno la sua funzione di autocorrelazione. Sotto certe condizioni è possibile tuttavia stimare in modo consistente la funzione di autocorrelazione partendo dalla serie storica dei dati osservati. Una prima condizione è che il processo sia stazionario. Per la stima di una qualunque funzione-momento deve valere inoltre la condizione di *ergodicità*, che assicura la convergenza in media quadratica dello stimatore al momento stesso (si veda Di Fonzo e Lisi, 2005, pp. 168-169).

Le proprietà dello stimatore della funzione di autocorrelazione dipendono ovviamente da quelle degli stimatori del valore medio e della funzione di autocovarianza del processo. Data una serie storica $\{y_t\}_{t=1}^n$, l'espressione

$$\hat{\mu} = \frac{1}{n} \sum_{t=1}^{n} y_t$$

fornisce uno stimatore corretto e consistente del valore medio se vale la condizione

$$\lim_{r\to\infty} \rho_r = 0,$$

la quale pone un limite alla memoria del processo. Se così non fosse, la porzione di processo stocastico osservata non sarebbe sufficiente per poter inferire sul valore atteso del processo. Per l'autocovarianza lo stimatore più utilizzato è

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(y_t - \hat{\mu})(y_{t+k} - \hat{\mu}) \tag{5.1}$$

che, pur essendo distorto per piccoli campioni, garantisce stime soddisfacenti di $\gamma_k$ per $n$ sufficientemente elevato e $k$ molto minore di $n$. Lo stimatore consistente (ma anch'esso distorto) della funzione di autocorrelazione è infine

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \tag{5.2}$$

che ha una varianza approssimata da

$$Var(\hat{\rho}_k) \simeq \frac{1}{n}\sum_{i=-q}^{q}\rho_i^2;$$

questa espressione è nota come *approssimazione di Bartlett*, spesso utilizzata nelle applicazioni per verificare se i residui di un modello stimato siano incorrelati (nel qual caso $Var(\hat{\rho}_k) = 1/n$).

**5.2** I due processi $X_t$ e $Y_t$ possiedono media nulla e funzione di autocovarianza teorica al ritardo $k$ pari a

$$\gamma_k^x \neq 0 \text{ per } k = 0, 1, 2, \quad \gamma_k^x = 0 \text{ per } k > 2$$

$$\gamma_k^y \neq 0 \text{ per } k = 0, 1, 2, 3, \quad \gamma_k^y = 0 \text{ per } k > 3.$$

Si derivi la funzione di autocorrelazione teorica per il processo $Z_t = X_t + bY_t$ in funzione di $\gamma_k^x$ e $\gamma_k^y$ assumendo che $X_t$ e $Y_t$ siano indipendenti e $b$ sia una costante.

□ Poiché $E(Z_t) = E(X_t) + bE(Y_t) = 0$, la funzione di autocovarianza del processo $Z_t$ è data da

$$E(Z_t Z_{t-k}) = \gamma_k^z \quad \text{per } k \geq 0.$$

Considerata la definizione del processo fornita dall'esercizio, si può scrivere

$$E[(X_t + bY_t)(X_{t-k} + bY_{t-k})] = \gamma_k^z;$$

moltiplicando tutti i termini tra loro e poiché il valore atteso di una combinazione lineare di v.c. è pari alla combinazione lineare dei valori attesi, si ottiene

$$\gamma_k^z = E(X_t X_{t-k}) + b^2 E(Y_t Y_{t-k}) + bE(X_t Y_{t-k}) + bE(X_{t-k} Y_t).$$

Data l'indipendenza tra i due processi $X_t$ e $Y_t$, l'autocovarianza di $Z_t$ si semplifica in

$$\gamma_k^z = \gamma_k^x + b^2\gamma_k^y$$

ovvero è una combinazione lineare con pesi $(1, b^2)$ delle autocovarianze dei due processi di partenza. Dal comportamento di queste ultime si desume la seguente struttura per $\gamma_k^z$:

$$\gamma_k^z = \begin{cases} \gamma_k^x + b^2\gamma_k^y & \text{per } k = 0, 1, 2 \\ b^2\gamma_3^y & \text{per } k = 3 \\ 0 & \text{per } k > 3. \end{cases}$$

La funzione di autocorrelazione teorica $\rho_k^z$ si ottiene nel modo usuale, rapportando l'autocovarianza $\gamma_k^z$ a $\gamma_0^z$. Pertanto si ha

$$\rho_k^z = \begin{cases} 1 & \text{per } k = 0 \\ \dfrac{\gamma_k^x + b^2\gamma_k^y}{\gamma_0^x + b^2\gamma_0^y} & \text{per } k = 1, 2 \\ \dfrac{b^2\gamma_3^y}{\gamma_0^x + b^2\gamma_0^y} & \text{per } k = 3 \\ 0 & \text{per } k > 3. \end{cases}$$

**5.3** Dato il processo $Y_t = \mu + \varepsilon_t + \varepsilon_{t-1}$, $t = 1, \ldots, N$, con $\varepsilon_t \sim WN(0,\sigma_\varepsilon^2) \quad \forall t$,

1. si calcoli $Var(\bar{Y})$, dove con $\bar{Y}$ si intende la media campionaria di $Y$;
2. si confronti il risultato ottenuto al punto precedente con quello che si sarebbe ottenuto se $Y_t = \mu + \varepsilon_t, t = 1, \ldots, N$.

□

$$\begin{aligned} \bar{Y} &= \frac{1}{N}\sum_{t=1}^{N} Y_t = \frac{1}{N}\sum_{t=1}^{N}(\mu + \varepsilon_t + \varepsilon_{t-1}) \\ &= \frac{1}{N}\sum_{t=1}^{N}\mu + \frac{1}{N}\sum_{t=1}^{N}(\varepsilon_t + \varepsilon_{t-1}) = \mu + \frac{1}{N}\sum_{t=1}^{N}(\varepsilon_t + \varepsilon_{t-1}) \end{aligned}$$

e pertanto $E[\bar{Y}] = \mu$. Osserviamo anche che $E[Y_t] = \mu$, $Var(Y_t) = 2\sigma_\varepsilon^2$ e

$$\gamma_y(k) = \begin{cases} \sigma_\varepsilon^2 & k = 1 \\ 0 & k \geq 2 \end{cases}.$$

In altre parole, $Y_t$ è un particolare processo MA(1).

Per calcolare $Var(\bar{Y})$ procediamo nel modo seguente:

$$\begin{aligned} Var(\bar{Y}) &= Var\left(\frac{1}{N}\sum_{t=1}^{N} Y_t\right) = \frac{1}{N^2} Var\left(\sum_{t=1}^{N} Y_t\right) \\ &= \frac{1}{N^2}\left[Var(Y_1) + Var(Y_2) + \ldots + Var(Y_N) + \right. \\ &\quad \left. + 2Cov(Y_1, Y_2) + 2Cov(Y_2, Y_3) + \ldots + 2Cov(Y_{N-1}, Y_N)\right] \\ &= \frac{1}{N^2}\left[N(2\sigma_\varepsilon^2) + 2(N-1)\sigma_\varepsilon^2\right] \\ &= \frac{1}{N^2}(4N\sigma_\varepsilon^2 - 2\sigma_\varepsilon^2) = \frac{2(2N-1)}{N^2}\sigma_\varepsilon^2 \end{aligned}$$

Se $Y_t = \mu + \varepsilon_t$, si ha

$$\bar{Y} = \frac{1}{N}\sum_{t=1}^{N} Y_t = \frac{1}{N}\sum_{t=1}^{N}(\mu + \varepsilon_t) = \mu + \frac{1}{N}\sum_{t=1}^{N}\varepsilon_t,$$

con $E(\bar{Y}) = \mu$ e

$$\begin{aligned} Var(\bar{Y}) &= Var\left(\mu + \frac{1}{N}\sum_{t=1}^{N}\varepsilon_t\right) \\ &= Var\left(\frac{1}{N}\sum_{t=1}^{N}\varepsilon_t\right) \\ &= \frac{1}{N^2}Var\left(\sum_{t=1}^{N}\varepsilon_t\right) = \frac{1}{N^2}\sum_{t=1}^{N}Var\left(\varepsilon_t\right) \\ &= \frac{\sigma_\varepsilon^2}{N} \end{aligned} .$$

Per $N$ grande $\frac{2(2N-1)}{N^2} = \frac{4N-2}{N^2} \sim \frac{4}{N}$, e quindi nel primo caso la varianza del processo considerato è di circa 4 volte superiore al secondo caso (a causa delle autocorrelazioni indotte dalla struttura a media mobile del processo).

**5.4** Si consideri il processo

$$\begin{aligned} X_1 &= \varepsilon_1 \\ X_t &= \phi X_{t-1} + \varepsilon_t \quad t = 2,3,\ldots \end{aligned}$$

dove $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ e $|\phi| < 1$.

(a) Si trovino media e varianza del processo $\{X_t\}$ e si dica, giustificando la risposta, se il processo è stazionario.

(b) Si mostri che

$$corr(X_t, X_{t-h}) = \phi^h \left[\frac{Var(X_{t-h}}{Var(X_t)}\right]^{\frac{1}{2}}.$$

□ Operando ricorsivamente, si ricava quanto segue:

$$\begin{aligned} E[X_1] = E[\varepsilon_1] &= 0 \\ E[X_2] = E[\phi X_1 + \varepsilon_2] &= \phi E[X_1] + E[\varepsilon_2] = 0 \\ E[X_3] = E[\phi X_2 + \varepsilon_3] &= \phi E[X_2] + E[\varepsilon_3] = 0 \\ &\cdots \\ E[X_t] = E[\phi X_{t-1} + \varepsilon_t] &= \phi E[X_{t-1}] + E[\varepsilon_t] = 0 \end{aligned}$$

e pertanto $E[X_t] = 0, \forall t$.

Allo stesso modo, si ha

$$
\begin{aligned}
&Var[X_1] = Var[\varepsilon_1] = \sigma_\varepsilon^2 \\
&Var[X_2] = \phi^2 Var[X_1] + Var[\varepsilon_2] = \phi^2\sigma_\varepsilon^2 + \sigma_\varepsilon^2 = \sigma_\varepsilon^2(1+\phi^2) \\
&Var[X_3] = \phi^2 Var[X_2] + Var[\varepsilon_3] = \phi^2\sigma_\varepsilon^2(1+\phi^2) + \sigma_\varepsilon^2 = \sigma_\varepsilon^2(1+\phi^2+\phi^4) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \dots \\
&Var[X_t] = \sigma_\varepsilon^2\left(1 + \sum_{i=1}^{t-1}\phi^{2i}\right)
\end{aligned}
$$

La varianza di $X_t$ dipende da $t$, e dunque il processo non è stazionario,

□ Per il calcolo della covarianza procediamo nel modo seguente:

$$
\begin{aligned}
Cov[X_t, X_{t-1}] &= E[X_t X_{t-1}] \\
&= E[(\phi X_{t-1} + \varepsilon_t)X_{t-1}] = \phi E[X_{t-1}^2 + E[\varepsilon_t X_{t-1}] \\
&= \phi Var[X_{t-1}]
\end{aligned}
$$

$$
\begin{aligned}
Cov[X_t, X_{t-2}] &= E[X_t X_{t-2}] \\
&= \phi E[X_{t-1}X_{t-2}] + E[\varepsilon_t X_{t-2}] \\
&= \phi Cov[X_{t-1}X_{t-2}] = \phi E[(\phi X_{t-2} + \varepsilon_{t-1})X_{t-2}]\,. \\
&= \phi^2 Var[X_{t-2}] \\
&\qquad\qquad \dots \\
Cov[X_t, X_{t-h}] &= \phi^h Var[X_{t-h}]
\end{aligned}
$$

Pertanto si ricava:

$$
\begin{aligned}
Corr[X_t, X_{t-h}] &= \frac{Cov[X_t, X_{t-h}]}{\sqrt{Var[X_t]Var[X_{t-h}}} \\
&= \phi^h \frac{Var[X_{t-h}]}{\sqrt{Var[X_t]Var[X_{t-h}}}\,. \\
&= \phi^h \sqrt{\frac{Var[X_{t-h}]}{Var[X_t]}}
\end{aligned}
$$

**5.5** Sia $X_t$ un processo a media mobile del tipo:

$$X_t = \varepsilon_{t-1} + 2\varepsilon_t + \varepsilon_{t+1}, \qquad \text{con } \varepsilon_t \sim IID(0,\sigma_\varepsilon^2).$$

1. Si calcoli la funzione di autocovarianza di $X_t$.
2. Si calcoli la funzione di autocorrelazione di $X_t$.

□ $X_t$ è una m.m. di ordine 3 del white noise $\varepsilon_t$, con pesi (1, 2, 1), positivi ma che non sommano a 1.

1. Poiché $E[X_t] = E[\varepsilon_{t-1} + 2\varepsilon_t + \varepsilon_{t+1}] = 0$, la funzione di autocovarianza di $X_t$ si calcola nel modo seguente:

$$\begin{aligned}
\gamma_0 = E[X_t^2] &= E\left[\varepsilon_{t-1}^2 + 4\varepsilon_t^2 + \varepsilon_{t+1}^2 + 4\varepsilon_{t-1}\varepsilon_t + 2\varepsilon_{t-1}\varepsilon_{t+1} + 4\varepsilon_t\varepsilon_{t+1}\right] \\
&= \sigma_\varepsilon^2 + 4\sigma_\varepsilon^2 + \sigma_\varepsilon^2 = 6\sigma_\varepsilon^2 \\
\gamma_1 = E[X_t X_{t+1}] &= E\left[(\varepsilon_{t-1} + 2\varepsilon_t + \varepsilon_{t+1})(\varepsilon_t + 2\varepsilon_{t+1} + \varepsilon_{t+2})\right] \\
&= E\left[\varepsilon_{t-1}\varepsilon_t + 2\varepsilon_{t-1}\varepsilon_{t+1} + \varepsilon_{t-1}\varepsilon_{t+2} + 2\varepsilon_t^2 + 4\varepsilon_t\varepsilon_{t+1} + 2\varepsilon_t\varepsilon_{t+2} + \varepsilon_{t+1}\varepsilon_t + 2\varepsilon_{t+1}^2 + \varepsilon_{t+1}\varepsilon_{t+2}\right] \\
&= 2\sigma_\varepsilon^2 + 2\sigma_\varepsilon^2 = 4\sigma_\varepsilon^2 \\
\gamma_2 = E[X_t X_{t+2}] &= E\left[(\varepsilon_{t-1} + 2\varepsilon_t + \varepsilon_{t+1})(\varepsilon_{t+1} + 2\varepsilon_{t+2} + \varepsilon_{t+3})\right] \\
&= E\left[\varepsilon_{t-1}\varepsilon_{t+1} + 2\varepsilon_t\varepsilon_{t+1} + \varepsilon_{t+1}^2 + 2\varepsilon_{t-1}\varepsilon_{t+2} + 4\varepsilon_t\varepsilon_{t+2} + 2\varepsilon_{t+1}\varepsilon_{t+2} + \varepsilon_{t-1}\varepsilon_{t+3} + \right. \\
&\quad \left. 2\varepsilon_t\varepsilon_{t+3} + \varepsilon_{t+1}\varepsilon_{t+3}\right] \\
&= \sigma_\varepsilon^2 \\
\gamma_k &= 0 \qquad k > 2
\end{aligned}$$

2. La funzione di autocorrelazione di $X_t$ si ottiene dividendo l'autocovarianza per $\gamma_0$:

$$\rho_k = \begin{cases} 1 & \text{per } k = 0 \\ \dfrac{4\sigma_\varepsilon^2}{6\sigma_\varepsilon^2} = \dfrac{2}{3} & \text{per } k = 1 \\ \dfrac{\sigma_\varepsilon^2}{6\sigma_\varepsilon^2} = \dfrac{1}{6} & \text{per } k = 2 \\ 0 & \text{per } k > 2 \end{cases}$$

**5.6** Data una serie $y_t$, $t = 1,2,\ldots,6$, realizzazione di un processo stocastico stazionario $Y_t$, si provi che se le autocovarianze del processo sono tutte positive, allora la varianza dello stimatore della media del processo, $Var(\hat{\mu})$, sarà maggiore rispetto al caso in cui tutte le autocovarianze siano nulle.

□ Essendo il processo $Y_t$ stazionario, la media campionaria è un "buon" stimatore della media del processo, per cui:

$$\hat{\mu} = \bar{Y} = \frac{1}{6}\sum_{t=1}^{6} Y_t.$$

Di conseguenza:

$$\begin{aligned}
Var(\hat{\mu}) &= \frac{1}{36} Var\left(\sum_{t=1}^{6} Y_t\right) \\
&= \frac{1}{36}\left[Var(Y_1) + Var(Y_2) + \ldots Var(Y_6) + 2Cov(Y_1,Y_2) + 2Cov(Y_1,Y_3) + \ldots + 2Cov(Y_5,Y_6)\right].
\end{aligned}$$

Ora, essendo $Y_t$ stazionario, si ha che $Var(Y_1) = \ldots = Var(Y_6) = \sigma_Y^2$, inoltre per ipotesi tutte le covarianze sono positive, e quindi:

$$Var(\hat{\mu}) = \frac{\sigma_Y^2}{6} + \frac{1}{36}B,$$

dove $B = 2\left[Cov(Y_1,Y_2) + \ldots + Cov(Y_5,Y_6)\right] > 0$. Se le autocovarianze fossero tutte nulle, allora $Var(\hat{\mu}) = \dfrac{\sigma_Y^2}{6} < \dfrac{\sigma_Y^2}{6} + \dfrac{1}{36}B$, e quindi la tesi è dimostrata.

## Stazionarietà

**5.7** Sia $X_t$ un processo stazionario di media $\mu$, varianza $\sigma_X^2$ e covarianza $\gamma_x(k)$, per $k \geq 1$. Si definisca il processo $Y_t$ come segue:

$$Y_t = \begin{cases} X_t & \text{per } t \text{ dispari} \\ X_t + 3 & \text{per } t \text{ pari} \end{cases}$$

Si mostri che:

1. il processo $Y_t$ è non stazionario, tuttavia
2. la sua funzione di autocovarianza dipende solo dal *lag* $k$ e non da $t$.

□

$$E(Y_t) = \begin{cases} E(X_t) = \mu & \text{per } t \text{ dispari} \\ E(X_t + 3) = \mu + 3 & \text{per } t \text{ pari} \end{cases}$$

Poiché il valore atteso di $Y_t$ varia al variare di $t$, il processo non è stazionario (c.v.d.).

□

$$Var(Y_t) = \begin{cases} Var(X_t) = \sigma^2 & \text{per } t \text{ dispari} \\ Var(X_t + 3) = \sigma^2 & \text{per } t \text{ pari} \end{cases}$$

Per $t$ dispari e $k$ dispari ($k = 1,3,5,\ldots$) si ha:

$$\begin{aligned} Cov(Y_t,Y_{t-k}) &= E\left[(X_t - \mu)(X_{t-k} + 3 - \mu - 3)\right] \\ &= \gamma_X(k) \end{aligned}.$$

Analogamente, se $t$ è pari e $k$ dispari, $Cov(Y_t,Y_{t-k}) = E\left[(X_t + 3 - \mu - 3)(X_{t-k} - \mu)\right] = \gamma_X(k)$. Infine, se $t$ è dispari e $k$ pari, $Cov(Y_t,Y_{t-k}) = E\left[(X_t + 3 - \mu - 3)(X_{t-k} - 3 - \mu + 3)\right] = \gamma_X(k)$. Pertanto, $\gamma_Y(k) = Cov(Y_t,Y_{t-k}) = \gamma_X(k)$, cioè dipende solo dal *lag* $k$ e non da $t$ (c.v.d.).

**5.8** Sia dato il processo

$$Y_t = (-1)^t(1 - \theta B)\varepsilon_t$$

con $\varepsilon_t \sim WN\left(0, \sigma_\varepsilon^2\right)$, $|\theta| < 1$. Dire, giustificando la risposta, se il processo $Y_t$ è debolmente stazionario.

□ Il processo $Y_t$ può essere riscritto nel modo seguente:

$$Y_t = \begin{cases} \varepsilon_t - \theta\varepsilon_{t-1} & t = 0 \text{ e } t \text{ pari} \\ -\varepsilon_t + \theta\varepsilon_{t-1} & t \text{ dispari} \end{cases}$$

È immediato verificare che, indipendentemente da $t$, $E(Y_t) = 0$ e $Var(Y_t) = \left(1 + \theta^2\right)\sigma_\varepsilon^2$. Quanto all'autocovarianza, si osservi che per $k > 1$ il valore atteso di $(Y_tY_{t+k})$ è sempre pari a zero, in quanto combinazione lineare (con pesi diversi a seconda che $t$ e/o $k$ siano pari o dispari) di quantità ($\varepsilon_t\varepsilon_{t+k}$, $\varepsilon_{t-1}\varepsilon_{t+k}$, $\varepsilon_t\varepsilon_{t+k-1}$, $\varepsilon_{t-1}\varepsilon_{t+k-1}$) il cui valore atteso è in ogni caso nullo.

Per $k = 1$, qualunque sia $t$, si ha invece:

$$\begin{aligned} E\left(Y_tY_{t+1}\right) &= -E\left[(\varepsilon_t - \theta\varepsilon_{t-1})(\varepsilon_{t+1} - \theta\varepsilon_t)\right] \\ &= -E\left[\varepsilon_t\varepsilon_{t+1} - \theta\varepsilon_{t-1}\varepsilon_{t+1} - \theta\varepsilon_t^2 + \theta^2\varepsilon_{t-1}\varepsilon_t\right] \\ &= \theta\sigma_\varepsilon^2 \end{aligned}$$

Il processo $Y_t$ è dunque (debolmente) stazionario.

**5.9** Dato il processo $X_t = \beta_1 + \beta_2 t + \varepsilon_t$, con $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ e $\beta_1$, $\beta_2$ costanti note:

(a) Si dica, giustificando la risposta, se $X_t$ è stazionario.

(b) Si consideri la media mobile

$$m_t = \frac{1}{2q+1}\sum_{j=-q}^{q} X_{t-j}$$

e si mostri che la media di $m_t$ è pari a $\beta_1 + \beta_2 t$.

□ Il processo $X_t$ non è stazionario in quanto contiene un trend deterministico. Infatti il suo valore atteso dipende da $t$:

$$E[X_t] = \beta_1 + \beta_2 t.$$

□ Per il calcolo del valore atteso di $m_t$ si ha:

$$E[m_t] = E\left[\frac{1}{2q+1}\sum_{j=-q}^{q} X_{t-j}\right] = \frac{1}{2q+1}E\left[\sum_{j=-q}^{q} X_{t-j}\right].$$

Poiché $E[X_{t-j}] = \beta_1 + \beta_2(t-j)$, si ha:

$$\begin{aligned} E[m_t] &= \frac{1}{2q+1}E\left[\sum_{j=-q}^{q}(\beta_1 + \beta_2(t-j)\right] \\ &= \frac{1}{2q+1}\left(\sum_{j=-q}^{q}\beta_1 + \sum_{j=-q}^{q}\beta_2 t - \sum_{j=-q}^{q}\beta_2 j\right) \\ &= \frac{1}{2q+1}\left(\beta_1(2q+1) + \beta_2 t(2q+1) + \beta_2 \cdot 0\right) \\ &= \beta_1 + \beta_2 t \end{aligned}$$

**5.10** Dato il processo $Y_t = a + bt + ct^2 + \varepsilon_t$, con $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$, si consideri $X_t = (1-B)^2 Y_t$.

1. Si stabilisca a che classe di processi appartiene $X_t$.
2. Si dica se $X_t$ è un processo stazionario e/o invertibile.

□ 1. Sviluppiamo l'espressione per $X_t$:

$$\begin{aligned} X_t &= (1-B)^2 Y_t = (1-2B+B^2)Y_t \\ &= Y_t - 2Y_{t-1} + Y_{t-2} \\ &= a + bt + ct^2 + \varepsilon_t - 2\left[a + b(t-1) + c(t-1)^2 + \varepsilon_{t-1}\right] + a + b(t-2) + c(t-2)^2 + \varepsilon_{t-2} \\ &= 2c + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} \end{aligned}$$

Quindi $X_t \sim MA(2)$ con media pari a $E[X_t] = \mu = 2c$.

2. Trattandosi di un processo $MA$, $X_t$ è certamente stazionario. Per valutarne l'invertibilità troviamo le radici dell'equazione

$$1 - 2B + B^2 = (1-B)^2 = 0 \quad \rightarrow \quad B_{1,2} = 1.$$

Il processo ha dunque due radici unitarie, e pertanto non è invertibile.

**5.11** Si consideri il processo $Y_t = C + Dt + X_t$, dove $X_t = X_{t-1} + \varepsilon_t$, con $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ e $X_0 = K$, dove $K$ è una costante nota.

1. Si supponga dapprima che $C$ e $D$ siano delle costanti note.
   (a) Si dica, giustificando la risposta, se il processo $Y_t$ è stazionario.
   (b) Si dica, giustificando la risposta, se il processo $(1 - B)Y_t$ è stazionario.
2. Si supponga ora che $C \sim \mathcal{N}(\mu_C, \sigma^2_C)$ e $D \sim \mathcal{N}(\mu_D, \sigma^2_D)$ indipendenti fra loro e con $X_t$.

□ 1a) $Y_t$ è non stazionario perché contiene un trend lineare $C + Dt$ e anche un trend stocastico (si può far vedere che $Y_t$ è un *random walk* con *drift*).

1b) $(1 - B)Y_t$ è invece stazionario. Infatti:

$$\begin{aligned}(1 - B)Y_t &= Y_t - Y_{t-1} \\ &= C + Dt + X_t - C - D(t - 1) - X_{t-1} \\ &= D + X_t - X_{t-1} = D + \varepsilon_t\end{aligned}$$

2a) Poiché, com'è immediato verificare, $E(X_t) = E\left(K + \sum_{j=1}^{t} \varepsilon_j\right) = K$, si ha: $E(Y_t) = E(C + Dt + X_t) = \mu_c + \mu_D t + K$. Quindi $E(Y_t)$ dipende da $t$, e dunque $Y_t$ non è stazionario.

2b) Come prima, $(1 - B)Y_t = D + \varepsilon_t$, solo che questa volta $D \sim \mathcal{N}(\mu_D, \sigma^2_D)$. Chiamato $W_t = (1 - B)Y_t$, si ha:

$$E(W_t) = E[D + \varepsilon_t] = E(D) + E(\varepsilon_t) = \mu_D \quad \forall t$$

$$Var(W_t) = Var[D + \varepsilon_t] = Var(D) + Var(\varepsilon_t) = \sigma^2_D + \sigma^2_\varepsilon \quad \forall t$$

$$\begin{aligned}Cov(W_t, W_{t+k}) &= E[(D + \varepsilon_t - \mu_D)(D + \varepsilon_{t+k} - \mu_D)] \\ &= E[D^2 + D\varepsilon_{t+k} - D\mu_D + \varepsilon_t D + \varepsilon_t\varepsilon_{t+k} - \varepsilon_t\mu_D - \mu_D D - \mu_D\varepsilon_{t+k} + \mu^2_D] \\ &= E(D^2) - \mu_D E(D) - \mu_D E(D) + \mu^2_D \\ &= E(D^2) - \mu^2_D - \mu^2_D + \mu^2_D = Var(D) = \sigma^2_D \quad \forall t\end{aligned}$$

quindi il processo è stazionario.

**5.12** Data una serie stagionale di osservazioni mensili, si assuma che l'effetto stagionale $S_t$ sia costante, cioè $S_t = S_{t-12}\ \forall t$, che il trend $T_t$ sia lineare e che $Y_t$ sia un processo stazionario a media nulla. Si consideri il modello (misto) $X_t = T_t S_t + Y_t$.

1. Si applichi l'operatore $(1 - B^{12})$ a $X_t$ e si dica, giustificando la risposta, se il processo risultante $W_t = (1 - B^{12})X_t$ è stazionario in media.
2. In caso negativo, si proponga un opportuno operatore differenza che risolva il problema.

□ Sviluppiamo anzitutto $W_t$:

$$W_t = (1 - B^{12})X_t = T_t S_t + Y_t - T_{t-12}S_{t-12} - Y_{t-12} = S_t(T_t - T_{t-12}) + Y_t - Y_{t-12}.$$

1. Poiché $T_t = a + bt$, si ha $W_t = S_t(a + bt - a - bt + 12 \cdot b) + Y_t - Y_{t-12}$, ossia:

$$W_t = 12 \cdot bS_t + Y_t - Y_{t-12}.$$

Pertanto, poiché $E(Y_t) = E(Y_{t-12}) = 0$, $E(W_t) = 12 \cdot bS_t$. La media di $W_t$ dipende quindi da $t$, sicché $W_t$ non è stazionario in media.

2. Applicando a $W_t$ l'operatore differenza stagionale $(1 - B^{12})$ si ottiene:

$$(1 - B^{12})W_t = 12 \cdot bS_t - 12 \cdot bS_{t-12} + Y_t - Y_{t-12} - Y_{t-12} + Y_{t-24} = Y_t - 2Y_{t-12} + Y_{t-24},$$

in quanto $S_t = S_{t-12}$. Stante la stazionarietà di $Y_t$, che ha media nulla, si ricava

$$E\left[(1 - B^{12})W_t\right] = 0$$

e dunque questo processo è stazionario in media.

## Processi ARMA

**5.13** Sia dato il processo stocastico

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1},$$

dove $|\phi_1| < 1$, $|\theta_1| < 1$ e $\varepsilon_t$ è un processo white noise di varianza $\sigma^2_\varepsilon$.

(a) Si identifichi $Y_t$ all'interno della classe dei processi ARMA.

(b) Si spieghi perché $Y_t$ è stazionario.

(c) Si calcoli $E(Y_t)$, $E(Y_t\varepsilon_t)$, $Var(Y_t) = \gamma_0$ e, infine, $\gamma_1$ e $\gamma_2$, ossia le autocovarianze di ordine 1 e 2.

□ (a) Un processo stocastico nella classe ARMA (*AutoRegressive Moving Average*) si rappresenta mediante legami di tipo lineare con il proprio passato (componente autoregressiva o AR) e con i valori correnti e passati di una componente stocastica di tipo *white noise* (componente media mobile o MA). Un processo ARMA($p$, $q$), talvolta detto processo misto, è definito dall'espressione

$$\begin{aligned} Y_t = \phi_0 + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} \\ +\varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \end{aligned}$$

che, utilizzando l'operatore ritardo $B$, si può scrivere nella forma compatta

$$\phi(B)Y_t = \phi_0 + \theta(B)\varepsilon_t$$

con

$$\begin{aligned} \phi(B) &= 1 - \phi_1 B - \cdots - \phi_p B^p \\ \theta(B) &= 1 - \theta_1 B - \cdots - \theta_q B^q. \end{aligned}$$

I valori $p$ e $q$ rappresentano l'*ordine* del processo ARMA, il primo riferito alla componente AR ed il secondo alla componente MA. Il processo proposto in questo esercizio è quindi un ARMA(1,1), il più semplice processo misto, senza il termine costante (ossia $\phi_0 = 0$).

Le proprietà stocastiche di un processo ARMA dipendono dalle proprietà algebriche degli operatori polinomiali $\phi(B)$ e $\theta(B)$: le condizioni poste sui parametri $|\phi_1| < 1$ e $|\theta_1| < 1$ assicurano la stazionarietà e l'invertibilità del processo $Y_t$. Il processo si dice quindi *ammissibile*: tale caratteristica permette di identificare univocamente i parametri del modello a partire dalla funzione di autocovarianza del processo ARMA.

(b) Un processo stocastico si definisce stazionario (in senso debole) se valgono le seguenti proprietà (Di Fonzo e Lisi, 2005, p. 165):

$$\begin{array}{ll} E(Y_t) = \mu & \forall t \\ E(Y_t - \mu)^2 = \sigma^2 < \infty & \forall t \\ E(Y_t - \mu)(Y_{t-k} - \mu) = \gamma_k & \forall t, \forall k. \end{array}$$

In sostanza, si pongono le condizioni che la media e la varianza del processo siano costanti nel tempo e che la funzione di autocovarianza, ossia la misura della covarianza fra le variabili casuali $Y_t$ e $Y_{t-k}$, non muti nel tempo ma dipenda esclusivamente dal ritardo temporale $k$.

Per un processo ARMA la stazionarietà si può verificare analizzando le proprietà del polinomio autoregressivo $\phi(B)$ (un processo MA($q$) è infatti sempre stazionario). Si dimostra che un processo ARMA ($p$, $q$) è stazionario se tutte le $p$ radici dell'equazione caratteristica $\phi(B) = 0$ sono, in modulo, maggiori di uno (Hamilton, 1994). In questo caso, l'equazione caratteristica è data da

$$1 - \phi_1 B = 0$$

la cui unica soluzione è $B = 1/\phi_1$. Dunque il processo è stazionario per $|\phi_1| < 1$, da cui si ottiene $|B| > 1$.

(c) Il valore atteso di un processo ARMA(1,1) è

$$\begin{aligned} E(Y_t) &= E(\phi_0 + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) \\ &= \phi_0 + \phi_1 E(Y_{t-1}) \end{aligned}$$

dove i termini $E(\varepsilon_t)$ e $E(\varepsilon_{t-1})$ si annullano considerando le proprietà di un processo *white-noise*. Ora, ipotizzando la stazionarietà del processo si può assumere $E(Y_t) = E(Y_{t-1}) = \mu$. Quindi,

$$\begin{aligned} E(Y_t) = \mu &= \phi_0 + \phi_1 \mu \\ \mu &= \frac{\phi_0}{1 - \phi_1}. \end{aligned}$$

Si noti che un processo autoregressivo stazionario ha una media sempre superiore al termine costante $\phi_0$: più alto è il valore del parametro autoregressivo $\phi_1$ maggiore sarà tale distanza. Per il processo in esame si ha $E(Y_t) = 0$, in quanto il termine costante è escluso dalla specificazione.

Il valore atteso del prodotto $(Y_t \varepsilon_t)$ esprime la covarianza fra il processo ARMA(1,1) ed il disturbo che guida il processo stesso. Si dimostra facilmente che

$$\begin{aligned} E(Y_t \varepsilon_t) &= E[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1})\varepsilon_t] \\ &= E(\varepsilon_t^2) = \sigma_\varepsilon^2 \end{aligned}$$

ovvero è pari alla varianza del disturbo.

La varianza di un processo ARMA si ottiene calcolando il momento centrale del secondo ordine. Nell'ambito delle serie storiche la varianza si indica spesso con $\gamma_0$, poiché rappresenta l'autocovarianza al ritardo 0. Nel caso in esame, considerando che $E(Y_t) = 0$ , si ha

$$\begin{aligned} E(Y_t^2) &= E[Y_t(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1})] \\ &= E(Y_t \cdot \phi_1 Y_{t-1}) + E(Y_t \varepsilon_t) - \theta_1 E(Y_t \varepsilon_{t-1}) \\ &= \phi_1 E(Y_t Y_{t-1}) + \sigma_\varepsilon^2 - \theta_1 E(Y_t \varepsilon_{t-1}). \end{aligned}$$

Sostituendo ancora, nel primo e nell'ultimo termine, l'espressione che identifica il processo $Y_t$ e sviluppando i prodotti si giunge all'espressione

$$\begin{aligned} E(Y_t^2) &= \phi_1 E[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) Y_{t-1}] + \sigma_\varepsilon^2 \\ &\quad -\theta_1 E[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) \varepsilon_{t-1}] \\ E(Y_t^2) &= \phi_1^2 E(Y_{t-1}^2) - \theta_1 \phi_1 \sigma_\varepsilon^2 + \sigma_\varepsilon^2 - \theta_1 \phi_1 \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2. \end{aligned}$$

Per un processo stazionario, la varianza è identica ad ogni istante temporale, quindi $E(Y_t^2) = E(Y_{t+k}^2) = \gamma_0$ per ogni $k$. È quindi possibile derivare una forma esplicita per la varianza di un processo ARMA(1,1):

$$\begin{aligned} \gamma_0 - \phi_1^2 \gamma_0 &= \sigma_\varepsilon^2 + \theta_1^2 \sigma_\varepsilon^2 - 2\theta_1 \phi_1 \sigma_\varepsilon^2 \\ (1 - \phi_1^2)\gamma_0 &= (1 + \theta_1^2 - 2\theta_1 \phi_1)\sigma_\varepsilon^2 \\ \gamma_0 &= \frac{(1 + \theta_1^2 - 2\theta_1 \phi_1)}{(1 - \phi_1^2)} \sigma_\varepsilon^2. \end{aligned}$$

Le autocovarianze di ordine 1 e 2, vale a dire $\gamma_1$ e $\gamma_2$, possono essere calcolate con lo stesso procedimento utilizzato per il calcolo della varianza $\gamma_0$. Per $\gamma_1$ si ha:

$$\begin{aligned} \gamma_1 = E(Y_t Y_{t-1}) &= E[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) Y_{t-1}] \\ &= \phi_1 E(Y_{t-1}^2) + E(Y_{t-1} \varepsilon_t) - \theta_1 E(Y_{t-1} \varepsilon_{t-1}). \end{aligned}$$

Semplificando con i risultati già ottenuti

$$\begin{aligned} \gamma_1 &= \phi_1 \gamma_0 - \theta_1 \sigma_\varepsilon^2 \\ &= \phi_1 \left[ \frac{(1 + \theta_1^2 - 2\theta_1 \phi_1)}{(1 - \phi_1^2)} \sigma_\varepsilon^2 \right] - \theta_1 \sigma_\varepsilon^2. \end{aligned}$$

Mettendo in evidenza il termine $\sigma_\varepsilon^2$, si giunge all'espressione per l'autocovarianza di ordine 1:

$$\begin{aligned} \gamma_1 &= \frac{(\phi_1 + \phi_1 \theta_1^2 - \theta_1 \phi_1^2 - \theta_1)}{(1 - \phi_1^2)} \sigma_\varepsilon^2 \\ &= \frac{(\phi_1 - \theta_1)(1 - \phi_1 \theta_1)}{(1 - \phi_1^2)} \sigma_\varepsilon^2. \end{aligned}$$

Infine, per $\gamma_2$ si ha

$$\begin{aligned} \gamma_2 = E(Y_t Y_{t-2}) &= E[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}) Y_{t-2}] \\ &= \phi_1 E(Y_{t-1} Y_{t-2}) = \phi_1 \gamma_1, \end{aligned}$$

che equivale all'autocovarianza di un processo AR(1). In effetti, per ritardi temporali superiori a $q$ la funzione di autocovarianza di un processo ARMA($p$, $q$) si comporta esattamente come quella di un processo AR di ordine $p$, ovvero vale la relazione

$$\gamma_k = \phi_1 \gamma_{k-1} = \phi^{k-1} \gamma_1, \quad k \geq p.$$

**5.14** Per ciascuna delle seguenti affermazioni dire se essa è vera, falsa o incerta. Commentare brevemente le soluzioni.

(a) Un processo stocastico MA di ordine finito è sempre stazionario.

(b) Un processo stocastico stazionario in senso forte è automaticamente stazionario in senso debole.

(c) Un processo del tipo $Y_t = \alpha Y_{t-4} + \varepsilon_t$, dove $\varepsilon_t$ è un *white noise*, è un processo AR(1).

□ (a) Vero. Secondo il teorema di Wold un processo stazionario può essere espresso come combinazione lineare di un numero infinito di processi *white noise*, cioè

$$Y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots = \psi(B)\varepsilon_t$$

dove la sommatoria infinita dei coefficienti al quadrato è finita, ovvero

$$\sum_j \psi_j^2 < \infty.$$

Quindi, un processo MA di ordine finito è sempre stazionario poichè la serie

$$\psi(B) = \theta(B) = 1 - \sum_j \theta_j B^j$$

è finita.

(b) Incerto. Un processo stocastico $Y_t$ è stazionario in senso forte se tutte le funzioni di ripartizione finite rimangono invariate per traslazioni dell'asse temporale, mentre è detto stazionario in senso debole quando sono invarianti rispetto al tempo *esclusivamente* i momenti del primo e del secondo ordine. Tuttavia, la stazionarietà in senso forte non implica l'esistenza dei momenti del processo, compresi i primi due. Pertanto, affinché un processo stazionario in senso forte sia anche un processo stazionario in senso debole è necessario che i momenti del primo e del secondo ordine siano **finiti**.

(c) Incerto. Il processo $Y_t$ è un AR(1) stagionale, noto in letteratura come SAR(1), con periodicità pari a $4$. Tale processo è adatto per cogliere comportamenti stagionali per serie trimestrali: ogni osservazione viene messa in relazione a quella corrispondente al trimestre dell'anno precedente. Si ricorda che le funzioni teoriche di autocorrelazione globale e parziale di un processo stocastico SAR(1) con periodo $s$ presentano la stessa dinamica di quella di un processo AR(1) replicata ogni $s$ passi: esiste quindi una forte affinità tra i due processi. L'analogia vale anche per $\alpha = 1$, nel qual caso la componente stagionale si dice non stazionaria.

**5.15** Sia dato il processo stocastico

$$Y_t = Y_{t-1} + \varepsilon_t + 0{,}8\varepsilon_{t-1},$$

dove $\varepsilon_t$ è un processo white noise di varianza $\sigma_\varepsilon^2$.

(a) Si classifichi $Y_t$ all'interno della classe dei processi ARMA.

(b) Si identifichi, sempre nell'ambito dei processi ARMA, il processo $(1 - B)Y_t$.

(c) Si dimostri che $Y_t$ non è stazionario.

□ (a) Il processo $Y_t$ è un ARMA(1,1), con $\phi_1 = 1$ e $\theta_1 = -0{,}8$. La radice della parte AR è unitaria, per cui il processo non è stazionario; esso risulta comunque invertibile, in quanto la condizione $|\theta_1| < 1$ è rispettata. Come vedremo più avanti, in questi casi è più utile una rappresentazione di tipo ARIMA, che estende quella ARMA distinguendo esplicitamente tra radici autoregressive stazionarie e non stazionarie (o unitarie). Nella fattispecie il processo corrisponde ad un ARIMA(0,1,1), in cui con il termine centrale si indicano per l'appunto il numero di radici unitarie in esso contenute.

(b) L'applicazione dell'operatore differenza $(1 - B)$ provoca la rimozione della radice unitaria dal processo $Y_t$. Infatti,

$$\begin{aligned}(1 - B)Y_t &= Y_t - Y_{t-1} \\ &= [Y_{t-1} + \varepsilon_t + 0{,}8\varepsilon_{t-1}] - Y_{t-1} \\ &= \varepsilon_t + 0{,}8\varepsilon_{t-1}\end{aligned}$$

che è un processo MA(1) con $\theta_1 = -0{,}8$. Questo è un esempio di processo non stazionario che può essere reso stazionario tramite un'opportuna trasformazione. Generalizzando, l'applicazione dell'operatore $(1 - B)^d$ ad un processo stocastico comporta la rimozione di $d$ radici unitarie.

(c) La condizione di stazionarietà $|\phi_1| < 1$ non è rispettata, ne consegue che il processo $Y_t$ non è stazionario. Analizziamo ora in dettaglio che tipo di non stazionarietà caratterizza $Y_t$. Riscriviamo il processo $Y_t$ attraverso sostituzioni ripetute dei propri valori passati:

$$\begin{aligned}Y_t &= Y_{t-1} + \varepsilon_t + 0{,}8\varepsilon_{t-1} \\ &= Y_{t-2} + \varepsilon_t + 1{,}8\varepsilon_{t-1} + 0{,}8\varepsilon_{t-2} \\ &= Y_{t-3} + \varepsilon_t + 1{,}8\varepsilon_{t-1} + 1{,}8\varepsilon_{t-2} + 0{,}8\varepsilon_{t-3} \\ &= \cdots \\ &= Y_0 + \varepsilon_t + 1{,}8\varepsilon_{t-1} + \cdots + 1{,}8\varepsilon_2 + 0{,}8\varepsilon_1.\end{aligned}$$

L'ultima espressione indica che $Y_t$ dipende dalla *condizione iniziale* del processo $Y_0$, per semplicità riferita al tempo 0, e dalla *storia passata* dei disturbi $\varepsilon_t$. Assumendo $Y_0$ costante ($Y_0 = y_0$), e prendendo il valore atteso di entrambi i membri si ottiene

$$E(Y_t) = y_0$$

dal momento che il valore atteso per i termini $\varepsilon_t$ è sempre nullo. La media del processo è quindi costante per ogni tempo, sebbene il valore $y_0$ sia ignoto.
Vediamo ora qual è il comportamento nel tempo della varianza di $Y_t$. Ipotizzando $y_0 = 0$, la varianza si ottiene facilmente dallo sviluppo dell'espressione

$$\begin{aligned}Var(Y_t) = E(Y_t^2) &= E[(\varepsilon_t + 1{,}8\varepsilon_{t-1} + \cdots + 1{,}8\varepsilon_2 + 1{,}8\varepsilon_1)^2] \\ &= E(\varepsilon_t^2) + 3{,}24E(\varepsilon_{t-1}^2) + \cdots \\ &\quad +3{,}24E(\varepsilon_2^2) + 3{,}24E(\varepsilon_1^2) \\ &= \sigma_\varepsilon^2[1 + 3{,}24(t-1)].\end{aligned}$$

La varianza non è quindi costante ma cresce linearmente con l'aumentare dell'orizzonte temporale. L'accumulo dei disturbi $\varepsilon_t$, in assenza di un fattore di smorzamento $\phi_1$ inferiore (in valore assoluto) ad 1, determina una crescita costante dell'incertezza, la quale induce il comportamento non stazionario del processo $Y_t$.

**5.16** Dati i due processi AR(1) debolmente stazionari

$$\begin{aligned}Y_t &= \phi_1 Y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim WN(0{,}\sigma_\varepsilon^2) \\ X_t &= \phi_1 X_{t-1} + \nu_t, \qquad \nu_t \sim WN(0{,}\sigma_\nu^2)\end{aligned}$$

con $\varepsilon_t$ indipendente da $\nu_t$, si consideri il processo $Z_t = Y_t + X_t$ risultante dalla somma dei due.

(a) Dire che tipo di processo è $Z_t$ (giustificare la risposta).

(b) Calcolare $E(Z_t)$, $Var(Z_t)$ e $Cov(Z_t Z_{t-1})$.

□ (a) Il processo $Z_t$ è il risultato della somma di due processi AR(1) debolmente stazionari, con lo stesso parametro autoregressivo. Scrivendo il processo per esteso otteniamo

$$\begin{aligned} Z_t &= Y_t + X_t \\ Z_t &= \phi_1 Y_{t-1} + \phi_1 X_{t-1} + \varepsilon_t + \nu_t \\ Z_t &= \phi_1 (Y_{t-1} + X_{t-1}) + \varepsilon_t + \nu_t \\ Z_t &= \phi_1 Z_{t-1} + \xi_t \end{aligned}$$

dove $\xi_t = \varepsilon_t + \nu_t$. Quest'ultimo processo conserva le proprietà di un processo *white noise*: infatti si ha

$$\begin{aligned} E(\xi_t) &= E(\varepsilon_t + \nu_t) = 0 \\ Var(\xi_t) &= E(\xi_t^2) = E[(\varepsilon_t + \nu_t)^2] \\ &= E(\varepsilon_t^2) + E(\nu_t^2) = \sigma_\varepsilon^2 + \sigma_\nu^2 \\ E(\xi_t \xi_{t-k}) &= E[(\varepsilon_t + \nu_t)(\varepsilon_{t-k} + \nu_{t-k})] = 0 \end{aligned}$$

considerato che l'indipendenza dei due disturbi originari comporta l'annullamento dei termini

$$E(\varepsilon_r \nu_s) = 0 \quad \forall r,s.$$

La stazionarietà debole di $Y_t$ e $X_t$ pone un vincolo al parametro autoregressivo $\phi_1$ comune ai due processi, ovvero deve valere $|\phi_1| < 1$. Quindi, il processo $Z_t$ è anch'esso un AR(1) debolmente stazionario.

(b) La media di un processo AR(1) è data da

$$\frac{\phi_0}{1 - \phi_1}$$

dove con $\phi_0$ si indica il termine costante. Entrambi i processi $Y_t$ e $X_t$ non presentano nella propria specifica tale termine, pertanto $\phi_0 = 0$. Se ne deduce che $E(Y_t) = 0$ e $E(X_t) = 0$. Per il processo somma si ha ovviamente

$$E(Z_t) = E(Y_t + X_t) = E(Y_t) + E(X_t) = 0.$$

Dalle proprietà di un processo AR(1) otteniamo le seguenti espressioni per la varianza di $Z_t$

$$\begin{aligned} E(Z_t^2) = \gamma_0 &= \frac{\sigma_\xi^2}{1 - \phi_1^2} = \frac{\sigma_\varepsilon^2 + \sigma_\nu^2}{1 - \phi_1^2} \\ &= \frac{\sigma_\varepsilon^2}{1 - \phi_1^2} + \frac{\sigma_\nu^2}{1 - \phi_1^2} = E(Y_t^2) + E(X_t^2) \end{aligned}$$

e per la sua autocovarianza al *lag* 1

$$\begin{aligned} E(Z_t Z_{t-1}) = \gamma_1 &= \phi_1 \gamma_0 \\ \gamma_1 &= \phi_1 E(Y_t^2) + \phi_1 E(X_t^2). \end{aligned}$$

Quindi, la risultante della somma di due processi AR(1) tra loro indipendenti è ancora un processo AR(1) che possiede una variabilità pari alla somma di quella dei processi originari. Si noti che la funzione di autocorrelazione è tuttavia identica nei tre casi: i processi condividono infatti la stessa struttura di dipendenza temporale stabilita dalla legge esponenziale

$$\rho_k = \phi_1^k \qquad \text{per } k \geq 0.$$

**5.17** Approssimando i calcoli alla seconda cifra decimale, indicare l'ordine $q$ e scrivere in forma esplicita il processo MA($q$) atto a rappresentare il seguente processo ARMA(1,1):

$$Y_t = 0{,}1Y_{t-1} + \varepsilon_t - 0{,}2\varepsilon_{t-1}$$

□ Il processo può essere scritto nel modo seguente:

$$(1 - 0{,}1B)Y_t = (1 - 0{,}2B)\varepsilon_t,$$

ovvero, per $B \neq 10$,

$$Y_t = \frac{(1 - 0{,}2B)}{(1 - 0{,}1B)}\varepsilon_t.$$

Ricordando che, per $|a| < 1$,

$$\sum_{i=0}^{\infty} a^i = 1 + a + a^2 + a^3 + \cdots = \frac{1}{1-a},$$

possiamo scrivere

$$\frac{1}{1 - 0{,}1B} = 1 + 0{,}1B + 0{,}01B^2 + 0{,}001B^3 + \cdots .$$

Approssimando i calcoli alla seconda cifra decimale possiamo dunque scrivere

$$Y_t = (1 + 0{,}1B + 0{,}01B^2)(1 - 0{,}2B)\varepsilon_t.$$

Sviluppando si ha

$$Y_t = (1 + 0{,}1B + 0{,}01B^2 - 0{,}2B - 0{,}02B^2 - 0{,}002B^3)\varepsilon_t.$$

Raccogliendo e approssimando alla seconda cifra decimale si ha:

$$Y_t = \varepsilon_t - 0{,}1\varepsilon_{t-1} - 0{,}01\varepsilon_{t-2}.$$

Pertanto, al livello di approssimazione dato, $Y_t$ può essere rappresentato da un processo MA(2) con coefficienti $\theta_1 = 0{,}1$ e $\theta_2 = 0{,}01$.

**5.18** Approssimando i calcoli alla terza cifra decimale, indicare l'ordine $q$ e scrivere in forma esplicita il processo MA($q$) atto a rappresentare un processo AR(1) con parametro $\phi_1 = -0{,}2$.

□ Operando sostituzioni successive per il processo AR(1)

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t$$

si ottiene:

$$\begin{aligned} Y_t &= \phi_1(\phi_1 Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &= \phi_1^2 Y_{t-2} + \varepsilon_t + \phi_1 \varepsilon_{t-1} \\ &= \phi_1^2(\phi_1 y_{t-3} + \varepsilon_{t-2}) + \varepsilon_t + \phi_1 \varepsilon_{t-1} \\ &= \phi_1^3 Y_{t-3} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} \\ &\vdots\,\vdots \\ &= \phi_1^m Y_{t-m} + \sum_{j=0}^{m-1} \phi_1^j \varepsilon_{t-j}. \end{aligned}$$

Poiché $(-0{,}2)^4 = 0{,}0016$ e $(-0{,}2)^5 = -0{,}00032$, con approssimazione pari alla terza cifra decimale l'ordine $q$ del modello MA atto a rappresentare il processo AR(1) in questione è 4. La forma esplicita di tale processo è la seguente

$$Y_t = \varepsilon_t - 0{,}2\varepsilon_{t-1} + 0{,}04\varepsilon_{t-2} - 0{,}008\varepsilon_{t-3} + 0{,}002\varepsilon_{t-4}.$$

**5.19** Si dica se il processo ARMA(2,2)

$$Y_t = 0{,}75Y_{t-1} - 0{,}125Y_{t-2} + \varepsilon_t - 0{,}8\varepsilon_{t-1} + 0{,}15\varepsilon_{t-2}$$

è sovraparametrizzato e, in caso affermativo, lo si semplifichi.

□ Il processo ARMA(2,2) può essere scritto nella forma compatta

$$(1 - 0{,}75B + 0{,}125B^2)Y_t = (1 - 0{,}8B + 0{,}15B^2)\varepsilon_t.$$

Per verificare se il processo è sovraparametrizzato bisogna risolvere le due equazioni caratteristiche associate ai polinomi AR e MA, cioè

$$\begin{aligned} \phi(B) &= 1 - 0{,}75B + 0{,}125B^2 = 0 \\ \theta(B) &= 1 - 0{,}8B + 0{,}15B^2 = 0 \end{aligned}$$

e confrontare le rispettive soluzioni. Risolviamo dapprima l'equazione AR:

$$B_{1/2} = \frac{0{,}75 \pm \sqrt{0{,}5625 - 0{,}5}}{0{,}25} = \frac{0{,}75 \pm 0{,}25}{0{,}25} = \begin{cases} 4 \\ 2 \end{cases}.$$

La parte AR può essere quindi espressa attraverso il prodotto dei due binomi

$$\phi(B) = (1 - 0{,}25B)(1 - 0{,}5B).$$

L'equazione caratteristica del polinomio MA ha invece soluzioni

$$B_{1/2} = \frac{0{,}8 \pm \sqrt{0{,}64 - 0{,}6}}{0{,}3} = \frac{0{,}8 \pm 0{,}2}{0{,}3} = \begin{cases} \frac{10}{3} \\ 2 \end{cases}$$

e pertanto esso può essere scomposto nel prodotto

$$\theta(B) = (1 - \frac{3}{10}B)(1 - 0{,}5B).$$

Esiste quindi una radice comune ai due polinomi ($B = 2$), pertanto il processo ARMA(2,2) originario può essere semplificato nel processo ARMA(1,1) seguente:

$$(1 - 0{,}25B)Y_t = (1 - 0{,}3B)\varepsilon_t.$$

.

## Esercizi supplementari

**C5.1** Dati i due processi MA(1) invertibili

$$\begin{aligned} Y_t &= \varepsilon_t - \theta_1\varepsilon_{t-1}, \qquad \varepsilon_t \sim WN(0{,}\sigma^2_\varepsilon) \\ X_t &= \nu_t - \theta_1\nu_{t-1}, \qquad \nu_t \sim WN(0{,}\sigma^2_\nu) \end{aligned}$$

con $\varepsilon_t$ indipendente da $\nu_t$, si consideri il processo $Z_t = Y_t + X_t$ risultante dalla somma dei due.

(a) Dire che tipo di processo è $Z_t$ (giustificare la risposta).

(b) Calcolare $E(Z_t)$, $Var(Z_t)$ e $Cov(Z_tZ_{t-1})$.

**C5.2** Si consideri un processo stocastico, $X_t$, con funzione di autocovarianza pari a:

$$\gamma_k = \begin{cases} 2 \quad \text{per} \quad k = 0 \\ 0{,}5 \quad \text{per} \quad k = \pm 1 \\ 0 \quad \text{altrimenti} \end{cases}$$

(a) Dire, **giustificando la risposta**, con che modello della classe ARIMA questa funzione di autocovarianza è coerente.

(b) Si calcolino i valori dei parametri del modello, stazionario e invertibile, indicato al punto precedente.

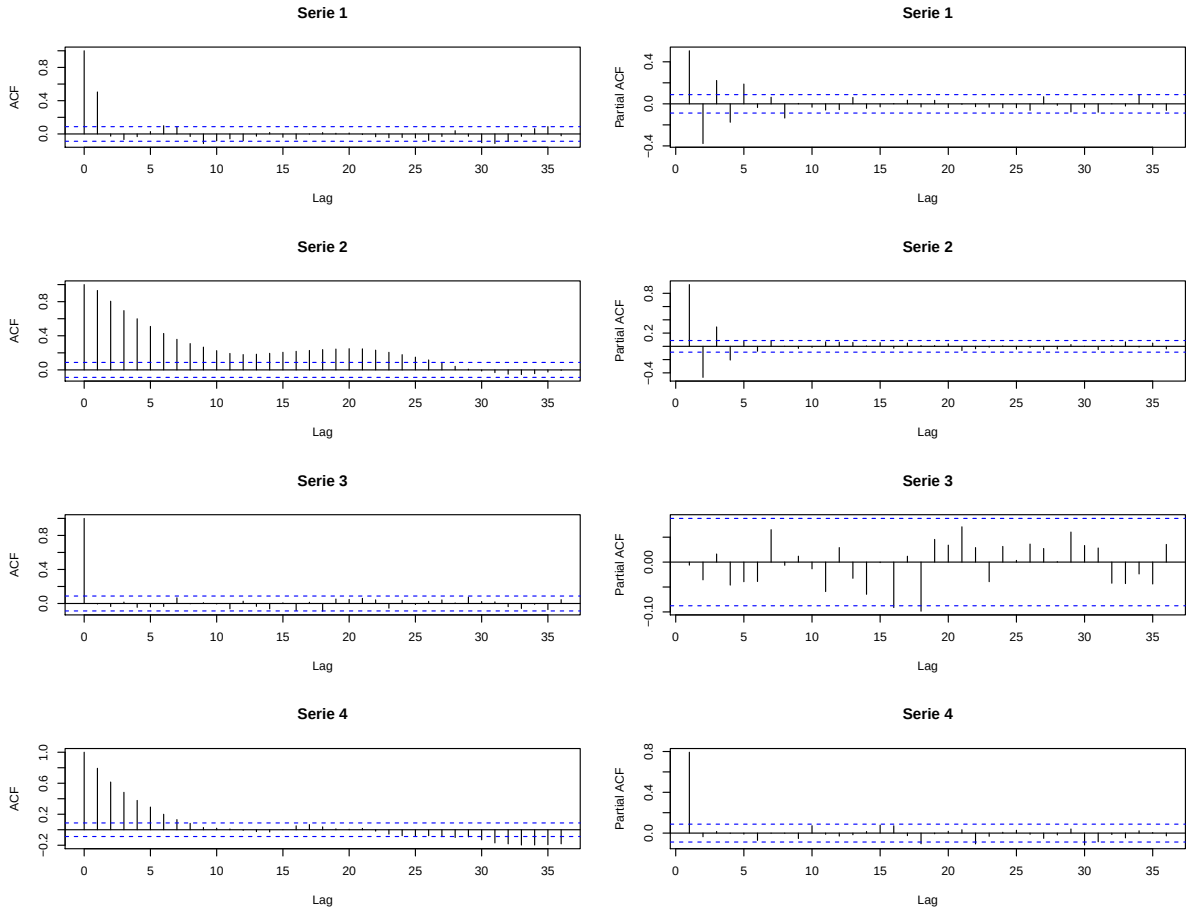
**C5.3** Assegnare, **giustificando la risposta**, le funzioni di autocorrelazione riportate in figura ai seguenti processi:

(a) $X_t = \varepsilon_t$

(b) $X_t = 0{,}8X_{t-1} + \varepsilon_t$

(c) $X_t = \varepsilon_t + 0{,}8\varepsilon_{t-1}$

(d) $X_t = 0{,}9X_{t-1} + \varepsilon_t + 0{,}8\varepsilon_{t-1}$




**C5.4** Dati i due processi:

$$X_t = \varepsilon_t - \frac{7}{6}\varepsilon_{t-1} + \frac{1}{3}\varepsilon_{t-2} \qquad \text{e} \qquad Y_t = \frac{5}{12}Y_{t-1} + \frac{1}{6}Y_{t-2} + X_t$$

dove $\varepsilon_t \sim WN(0,1)$.

(a) si mostri che $Y_t$ può essere scritto come un $ARMA(p,q)$ e si determinino gli ordini $p$ e $q$.

(b) Dire, giustificando la risposta, se il processo $Y_t$ è stazionario e invertibile.

(c) Calcolare la varianza di $Y_t$.

**C5.5** Si consideri il processo

$$X_t = W_t + \frac{1}{2}W_{t-1} \qquad \text{con} \qquad W_t \sim IID\ U(0,2)$$

(a) Calcolare il valore atteso e la varianza di $X_t$.

(b) Calcolare la funzione di autocorrelazione di $X_t$.

(c) Dire, **giustificando la risposta**, se il processo è stazionario.

(N.B. Si ricordi che se $X \sim U(a,b)$ allora $f(x) = \frac{1}{b-a}$ per $x \in [a,b]$, 0 altrimenti)

# 6
# Processi ARIMA

**6.1** Si consideri un modello ARIMA(1,1,1) con $\phi = 0{,}3$ e $\theta = 0{,}5$. Scriverlo in forma di modello ARMA($p$, $q$) e verificare che le radici dell'equazione caratteristica non soddisfano le condizioni di stazionarietà.

□ L'espressione usuale di un modello ARIMA(1,1,1) è

$$(1 - 0{,}3B)(1 - B)Y_t = (1 - 0{,}5B)\varepsilon_t.$$

Sviluppando il prodotto dei due polinomi in $B$ a sinistra del segno di uguaglianza, si ottiene

$$(1 - 1{,}3B + 0{,}3B^2)Y_t = (1 - 0{,}5B)\varepsilon_t,$$

che si presenta come un modello ARMA(2,1). Per come è stato ottenuto, sappiamo però che le due radici dell'equazione caratteristica della parte AR,

$$(1 - 1{,}3B + 0{,}3B^2) = 0$$

sono

$$B_1 = \frac{1}{0{,}3} = \frac{10}{3} \quad \text{e} \quad B_2 = 1.$$

Solo una radice, dunque, è in modulo maggiore di uno, mentre l'altra è una radice unitaria, per cui, com'è ovvio, il modello non è stazionario.

**6.2** Sia $\{Y_t\}_{t=1}^n$ una serie storica generata da un processo stazionario tale che $E(Y_t) = 0$, $Var(Y_t) = 0{,}81$, $Cov(Y_t,Y_s) = 0 \ \forall t \neq s$. Si consideri il processo $X_t$ ottenuto applicando la media mobile

$$X_t = \frac{1}{3}(Y_{t-1} + Y_t + Y_{t+1}).$$

(a) Calcolare il valore atteso e la funzione di autocorrelazione del processo $X_t$. Commentare.

(b) Basandosi solamente sulla funzione di autocorrelazione, quale potrebbe essere il modello della classe ARIMA più adatto a descrivere $X_t$?

□ (a) Il valore atteso del processo $X_t$ è anch'esso nullo, secondo l'espressione

$$\begin{aligned} E(X_t) &= \frac{1}{3}E(Y_{t-1} + Y_t + Y_{t+1}) \\ &= \frac{1}{3}E(Y_{t-1}) + \frac{1}{3}E(Y_t) + \frac{1}{3}E(Y_{t+1}) \\ &= 0. \end{aligned}$$

Per il calcolo della funzione di autocorrelazione si determina in primo luogo la varianza del processo $X_t$. Considerando l'ipotesi di incorrelazione del processo $Y_t$ a periodi diversi, si può scrivere

$$E(X_t^2) = \frac{1}{9}E(Y_{t-1}^2) + \frac{1}{9}E(Y_t^2) + \frac{1}{9}E(Y_{t+1}^2)$$

che, una volta sostituito il valore della varianza di $Y_t$, equivale a

$$E(X_t^2) = 0{,}27.$$

Analogamente, l'autocovarianza del primo ordine $\gamma_1$ è pari a

$$E(X_t X_{t-1}) = \frac{1}{9}E(Y_{t-1}^2) + \frac{1}{9}E(Y_t^2) = 0{,}18$$

mentre quella del secondo ordine $\gamma_2$ è

$$E(X_t X_{t-2}) = \frac{1}{9}E(Y_{t-1}) = 0{,}09.$$

Dal ritardo 3 in poi l'autocovarianza è nulla, ovvero $\gamma_k = 0$ per $k > 2$. La funzione di autocorrelazione $\rho_k$ risulta quindi pari a

$$\rho_k = \begin{cases} 0{,}666 & \text{per } k = 1 \\ 0{,}333 & \text{per } k = 2 \\ 0 & \text{per } k \geq 3. \end{cases}$$

Rispetto al processo originario $Y_t$ il processo $X_t$ possiede una varianza inferiore ed include una correlazione seriale spuria, generata dall'applicazione della media mobile di tre termini ad un processo incorrelato (effetto di Slutsky-Yule, già introdotto nel capitolo sulle medie mobili).

(b) Una funzione di autocorrelazione globale che si annulla dal ritardo 3 in poi è tipica di un processo MA(2).

**6.3** Sia $x_t$, $t = 1{,}2, \ldots, 240$, una serie storica che viene modellata come un $ARIMA(p{,}d{,}q)$. Dopo aver differenziato la serie, si ottengono le seguenti funzioni di autocorrelazioni empiriche:

| Lag | ACF | PACF |
|---|---|---|
| 1 | -0,2946 | ? |
| 2 | -0,0868 | -0,1901 |
| 3 | -0,0698 | -0,1775 |
| 4 | 0,0869 | -0,0165 |
| 5 | 0,0624 | 0,0673 |
| 6 | -0,0952 | -0,0482 |

1. Completare la tabella, fornendo il valore della PACF al ritardo 1.
2. Suggerire, motivando la risposta, i valori di $p$, $d$ e $q$ che dovrebbero meglio modellare i dati.
3. Scrivere la rappresentazione del modello scelto.

□ Aver differenziato la serie equivale ad avere fissato (almeno per il momento) $d = 1$.

1. La PACF al ritardo 1 è per definizione uguale alla ACF allo stesso ritardo, e pertanto il valore mancante è pari a -0,2946.
2. Le bande di confidenza per ACF e PACF sono approssimativamente pari a $\pm\frac{2}{\sqrt{240}} = \pm 0{,}1291$. Ne deriva che l'ACF al ritardo 1 e le PACF ai primi 3 ritardi sono significativamente diverse da 0, mentre tutte le altre no. Questo andamento è tipico di una serie storica generata da un processo MA(1), e quindi si può ritenere che il processo generatore della serie originale sia un $ARIMA(0,1,1)$.
3. La rappresentazione del modello $ARIMA(0,1,1)$ è la seguente:

$$(1-B)X_t = (1-\phi B)\varepsilon_t \qquad \text{con } \varepsilon_t \sim WN(0,\sigma^2_\varepsilon).$$

**6.4** Si consideri il processo

$$Y_t = Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim WN(0,1), \quad Y_0 = 0.$$

(a) Calcolare il valore atteso e la varianza del processo.

(b) Calcolare la funzione di autocorrelazione del processo.

□ (a) Il processo $Y_t$ è non stazionario e prende il nome di *random walk*. Per il calcolo del valore atteso e della varianza è utile operare sostituzioni ripetute di $Y_t$ ai vari ritardi:

$$\begin{aligned} Y_t &= Y_{t-1} + \varepsilon_t \\ &= Y_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ &= Y_{t-3} + \varepsilon_{t-2} + \varepsilon_{t-1} + \varepsilon_t \\ &= \cdots \\ &= Y_0 + \sum_{i=1}^{t} \varepsilon_i, \end{aligned} \tag{6.1}$$

da cui, tenuto conto che $Y_0 = 0$, si ricava che il processo può essere riscritto nella forma $Y_t = \sum_{i=1}^{t} \varepsilon_i$. Prendendo il valore atteso di ambo i membri si ottiene

$$E(Y_t) = \sum_{i=1}^{t} E(\varepsilon_i) = 0.$$

La varianza del processo è pari a

$$Var(Y_t) = E(Y_t^2) = \sum_{i=1}^{t} Var(\varepsilon_i^2) = t,$$

ovvero cresce con il passare del tempo $t$. Il processo quindi risulta non stazionario in varianza.

(b) Per calcolare la funzione di autocovarianza conviene riferirsi di nuovo all'espressione (6.1):

$$E(Y_t Y_{t+k}) = E\left(\sum_{i=1}^{t} \varepsilon_i \sum_{j=1}^{t+k} \varepsilon_j\right)$$

Considerato che il processo $\varepsilon_t$ è un *white noise*, tutti i prodotti incrociati - che cioè interessano indici $i$ e $j$ diversi - si annullano. Gli unici termini non nulli sono gli errori al quadrato comuni nei due istanti temporali. L'espressione può essere pertanto semplificata in

$$E(Y_t Y_{t+k}) = \sum_{i=1}^{t} E(\varepsilon_i^2) = t.$$

La funzione di autocorrelazione del processo *random walk* è quindi pari a

$$\rho_k = \frac{E(Y_t Y_{t+k})}{\sqrt{Var(Y_t) \cdot Var(Y_{t+k})}} = \frac{t}{\sqrt{t \cdot (t+k)}} = \sqrt{\frac{t}{t+k}}.$$

Si noti come l'autocorrelazione dipenda in questo caso dall'indice temporale $t$ e che al crescere dello stesso il suo valore tenda ad 1 per tutti i ritardi temporali. Per tale motivo il processo si dice avere memoria infinita.

### Stazionarietà ed invertibilità

**6.5** Si verifichi la stazionarietà del processo

$$Y_t = t^\alpha (1 - \theta B)\varepsilon_t$$

con

$$\varepsilon_t \sim WN(0,\sigma^2),\ |\theta| < 1,\ \alpha > 0.$$

□ Calcoliamo anzitutto la media del processo:

$$E(Y_t) = t^\alpha E(\varepsilon_t - \theta\varepsilon_{t-1}) = 0 \quad \forall\alpha.$$

La varianza è invece pari a

$$Var(Y_t) = t^{2\alpha} E[(\varepsilon_t - \theta\varepsilon_{t-1})^2] = t^{2\alpha}\sigma^2(1+\theta^2).$$

Perché il processo sia (debolmente) stazionario, è necessario che la varianza non dipenda da $t$. Per $\alpha > 0$ ciò non accade, e dunque il processo è non stazionario.

**6.6** Determinare l'insieme dei valori $\alpha$ per cui risulta stazionario il processo

$$Z_t = Z_{t-1} + \alpha Z_{t-2} + \eta_t,\ \eta_t \sim WN(0;\ 1{,}4^2)$$

per $\alpha \in R$.

□ $Z_t$ è un processo AR(2), con parametro del primo ordine fissato ($\phi_1 = 1$) e parametro del secondo ordine libero ($\phi_2 = \alpha$). Il processo AR(2) in esame risulta stazionario se le radici dell'equazione caratteristica

$$-\alpha B^2 - B + 1 = 0$$

sono in modulo maggiori di uno. Risolvendo l'equazione si ottengono le due soluzioni

$$B_{1/2} = \frac{1 \pm \sqrt{1+4\alpha}}{-2\alpha} = \begin{cases} \dfrac{-1-\sqrt{1+4\alpha}}{2\alpha} \\ \\ \dfrac{-1+\sqrt{1+4\alpha}}{2\alpha} \end{cases} .$$

Per verificare la proprietà di stazionarietà bisogna valutare le condizioni $|B_1| > 1$ e $|B_2| > 1$ in funzione del parametro $\alpha$. Notiamo anzitutto che le radici sono *reali* per $1 + 4\alpha \geq 0$, ossia per $\alpha \geq -\frac{1}{4}$, e *complesse* per $\alpha < -\frac{1}{4}$. Concentriamo l'attenzione prima sulle soluzioni reali dell'equazione, ossia valutiamo $B_{1/2}$ per $\alpha \geq -\frac{1}{4}$. Osserviamo che per $\alpha > 0$ si ha $B_2 > 0$, per cui per tale radice (e per $\alpha > 0$) iniziamo col determinare i valori di $\alpha$ per i quali sia $B_2 > 1$, ossia:

$$\begin{aligned} -1 + \sqrt{1+4\alpha} &> 2\alpha \\ \sqrt{1+4\alpha} &> 2\alpha + 1 \\ 1 + 4\alpha &> 4\alpha^2 + 4\alpha + 1 \\ 4\alpha^2 &< 0 \end{aligned}$$

Visto che l'ultima espressione non può mai essere soddisfatta, ne consegue che per $\alpha > 0$ si ha sempre $0 < B_2 < 1$, e dunque almeno una delle due radici è all'interno del cerchio di raggio unitario. In altre parole, per $\alpha > 0$ il processo non è stazionario. Passiamo ora a considerare il caso $-\frac{1}{4} \leq \alpha < 0$. È immediato verificare che, in questo caso, $0 < B_2 \leq B_1$, per cui è sufficiente determinare per quali valori di $\alpha$ si ha $B_2 > 1$, ossia:

$$\begin{aligned} -1 + \sqrt{1+4\alpha} &< 2\alpha \\ \sqrt{1+4\alpha} &< 2\alpha + 1\,. \\ 4\alpha^2 &> 0 \end{aligned}$$

Ora, poiché per $\alpha \geq -\frac{1}{4}$ (e in verità per $\alpha$ qualsiasi) si ha sempre $4\alpha^2 > 0$, la diseguaglianza $B_2 > 1$ è sempre vera e, a maggior ragione, $B_1 > 1$. In altre parole, per $-\frac{1}{4} \leq \alpha < 0$ il processo è stazionario.

Studiamo ora le radici per $\alpha < -\frac{1}{4}$, ovvero nel campo immaginario. La condizione di stazionarietà di un processo con radici complesse richiede che il loro modulo sia superiore ad uno. Il modulo $R$ è definito nel seguente modo (Hamilton, 1990):

$$R = \sqrt{a^2 + b^2} \tag{6.2}$$

dove $a$ e $b$ rappresentano rispettivamente la parte reale e la parte immaginaria della radice complessa. Nel nostro caso tali valori sono pari a

$$\begin{aligned} a &= -\frac{1}{2\alpha} \\ b &= \frac{\sqrt{-1-4\alpha}}{2\alpha}. \end{aligned}$$

L'espressione (6.2) assume quindi la forma

$$\begin{aligned} R &= \sqrt{\frac{1}{4\alpha^2} + \frac{-1-4\alpha}{4\alpha^2}} \\ &= \sqrt{-\frac{1}{\alpha}}. \end{aligned}$$

Affinché il processo risulti stazionario con radici complesse il modulo $R$ deve essere superiore ad 1, ossia

$$R = \sqrt{-\frac{1}{\alpha}} > 1,$$

relazione che è soddisfatta per $-1 < \alpha < 0$.

Si ricorda tuttavia che la regione di valori di $\alpha$ per avere radici complesse non supera il valore $-1/4$, pertanto la condizione precedente si restringe a

$$-1 < \alpha < -1/4.$$

Il processo $Z_t$ è pertanto stazionario per $-1 < \alpha < 0$, e le radici del polinomio caratteristico associato sono complesse per $-1 < \alpha < -\frac{1}{4}$ e reali per $-\frac{1}{4} \leq \alpha < 0$.

**6.7** Dire (giustificando la risposta) se il processo

$$Y_t = \varepsilon_t - 0{,}4\varepsilon_{t-1} - 0{,}21\varepsilon_{t-2},\ \varepsilon_t \sim WN(0, \sigma^2),$$

è stazionario ed invertibile.

□ Si tratta di un processo MA(2) che, essendo combinazione lineare di componenti *white noise*, è sempre stazionario. Quanto all'invertibilità, è necessario calcolare le radici dell'equazione caratteristica

$$-0{,}21B^2 - 0{,}4B + 1 = 0.$$

Il processo è invertibile se le due radici della precedente equazione sono, in modulo, maggiori di uno. Risolvendo l'equazione si trova:

$$B_{1/2} = \frac{0{,}4 \pm \sqrt{0{,}16 + 0{,}84}}{-0{,}42},$$

ossia

$$B_1 = \frac{0{,}4 + \sqrt{0{,}16 + 0{,}84}}{-0{,}42} = -\frac{1{,}4}{0{,}42} = -3{,}33$$
$$B_2 = \frac{0{,}4 - \sqrt{0{,}16 + 0{,}84}}{-0{,}42} = \frac{0{,}6}{0{,}42} = 1{,}43.$$

Pertanto, il processo è invertibile.

**6.8** Si considerino i processi

$$(1 - 0{,}8B + 0{,}12B^2)X_t = (1 - 0{,}2B)\varepsilon_t$$
$$(1 - 0{,}7B)Y_t = (1 - 0{,}4B - 0{,}21B^2)\varepsilon_t$$

con $\varepsilon_t \sim WN(0, \sigma^2)$. Si dica se sono stazionari e/o invertibili.

□ Il processo $X_t$ è un ARMA(2,1), mentre il processo $Y_t$ è un ARMA(1,2). Un processo ARMA è stazionario quando le radici del polinomio caratteristico della parte autoregressiva sono in modulo maggiori di uno; si dice invece invertibile quando le radici del polinomio caratteristico della componente a media mobile sono in modulo maggiori di uno. Verifichiamo nell'ordine tali condizioni per il processo $X_t$. L'equazione caratteristica del polinomio autoregressivo è

$$1 - 0{,}8B + 0{,}12B^2 = 0$$

le cui soluzioni sono

$$B_{1/2} = \frac{0{,}8 \pm \sqrt{0{,}16}}{0{,}24} = \begin{cases} 5 \\ 1{,}666 \end{cases}.$$

Il processo $X_t$ è dunque stazionario. La radice del polinomio a media mobile è pari a $B_1 = 1/0{,}2 = 5$, pertanto il processo è anche invertibile. Si noti, peraltro, che il polinomio autoregressivo e quello a media mobile presentano una radice comune ($B_1 = 5$), e dunque il processo valido per $X_t$ è di fatto un AR(1) stazionario. Analizziamo ora le proprietà del processo $Y_t$. La parte autoregressiva è del primo ordine, sicché la condizione di stazionarietà è immediatamente verificabile: la radice è pari a $B_1 = 1/0{,}7 = 1{,}429$, pertanto il processo $Y_t$ è stazionario. La verifica dell'invertibilità richiede invece nuovamente di risolvere l'equazione caratteristica di secondo grado associata al polinomio MA

$$1 - 0{,}4B - 0{,}21B^2 = 0$$

che ha soluzioni

$$B_{1/2} = \frac{0{,}4 \pm \sqrt{1}}{-0{,}42} = \begin{cases} -3{,}333 \\ 1{,}429 \end{cases} .$$

Essendo entrambe le radici maggiori di 1 in valore assoluto, si può concludere che il processo $Y_t$ è stazionario e invertibile. Anche in questo caso, però, si può operare una semplificazione visto che c'è una radice comune ($B = 1{,}429$). Il processo $Y_t$ è dunque, di fatto, un MA(1) stazionario.

**6.9** Discutere stazionarietà ed invertibilità del processo

$$Y_t = 0{,}3Y_{t-1} + \varepsilon_t - 0{,}09\varepsilon_{t-2},$$

con $\varepsilon_t \sim WN(0,\sigma^2)$, e calcolare la funzione di autocorrelazione globale.

□ Il processo può essere scritto nella forma

$$(1 - 0{,}3B)Y_t = (1 - 0{,}09B^2)\varepsilon_t$$

ossia, poiché $1 - 0{,}09B^2 = (1 - 0{,}3B)(1 + 0{,}3B)$,

$$(1 - 0{,}3B)y_t = (1 - 0{,}3B)(1 + 0{,}3B)\varepsilon_t.$$

Semplificando si ottiene il processo MA(1) stazionario ed invertibile

$$Y_t = \varepsilon_t + 0{,}3\varepsilon_{t-1},$$

la cui funzione di autocorrelazione globale è pari a

$$\rho_k = \begin{cases} -\dfrac{\theta_1}{1 + \theta_1^2} & \text{per } k = 1 \\ 0 & \text{per } k > 1 \end{cases}$$

con $\theta_1 = -0{,}3$. Si ha dunque:

$$\rho_k = \begin{cases} 0{,}27523 & \text{per } k = 1 \\ 0 & \text{per } k > 1. \end{cases}$$

**6.10** Si consideri il processo

$$Y_t = 0{,}2Y_{t-1} + 0{,}15Y_{t-2} + \varepsilon_t + 0{,}3\varepsilon_{t-1}, \quad \varepsilon_t \sim WN(0{,}1).$$

(a) Verificare se il processo è stazionario e/o invertibile.

(b) Calcolare i pesi della rappresentazione AR($\infty$) e MA($\infty$) per $j = 1{,}2{,}3{,}4{,}5$ e fornire l'espressione per il calcolo dei generici pesi $\psi_j$ e $\pi_j$.

(c) Calcolare il valore atteso e la funzione di autocorrelazione globale del processo.

□ (a) Il processo può essere scritto in forma compatta

$$\phi(B)Y_t = \theta(B)\varepsilon_t,$$

con $\phi(B) = 1 - 0{,}2B - 0{,}15B^2$ e $\theta(B) = 1 + 0{,}3B$. Cerchiamo anzitutto le radici del polinomio autoregressivo, ossia le soluzioni dell'equazione $\phi(B) = 0$:

$$1 - 0{,}2B - 0{,}15B^2 = 0 \Leftrightarrow B_{1/2} = \frac{0{,}2 \pm \sqrt{0{,}04 + 0{,}6}}{-0{,}3}$$
$$B_{1,2} = \frac{0{,}2 \pm 0{,}8}{-0{,}3} = \begin{cases} -3{,}33 \\ 2 \end{cases}.$$

Poiché $\phi(B) = 1 - 0{,}2B - 0{,}15B^2 = (1 + 0{,}3B)(1 - 0{,}5B)$, il processo si semplifica nel modo seguente:

$$(1 - 0{,}5B)Y_t = \varepsilon_t \Rightarrow Y_t = 0{,}5y_{t-1} + \varepsilon_t.$$

Il processo $Y_t$ è dunque un AR(1) stazionario ($\phi = 0{,}5$), che è per definizione invertibile.

(b) Un processo ARMA stazionario ed invertibile può essere scritto nelle forme equivalenti

$$\pi(B)Y_t = \varepsilon_t$$
$$\pi(B) = \frac{\phi(B)}{\theta(B)}$$

e

$$Y_t = \psi(B)\varepsilon_t$$
$$\psi(B) = \frac{\theta(B)}{\phi(B)}$$

Nel primo caso si tratta di un processo AR($\infty$), il secondo è un processo MA($\infty$). Per derivare il valore dei coefficienti bisogna impostare le identità

$$\pi(B) = (1 - \pi_1 B - \pi_2 B^2 - \ldots) = \phi(B)\theta^{-1}(B)$$

e

$$\psi(B) = (1 - \psi_1 B - \psi_2 B^2 - \ldots) = \phi^{-1}(B)\theta(B),$$

e uguagliare i coefficienti delle stesse potenze di $B$ a sinistra e a destra del segno di uguale. Nel caso specifico di un AR(1) i pesi della rappresentazione AR($\infty$) sono $\pi_1 = \phi = 0{,}5$ e $\pi_j = 0$ per $j > 1$. Sfruttando poi la nota relazione

$$(1 - \phi_1 B)^{-1} = (1 + \phi B + \phi^2 B^2 + \phi^3 B^3 + \ldots)$$

i pesi della rappresentazione MA($\infty$) risultano pari a

$$\psi_j = -\phi_1^j,\ j > 0.$$

Sostituendo a $\phi$ il valore 0,5 si ha

$$\psi_1 = -0{,}5,\ \psi_2 = -0{,}25,\ \psi_3 = -0{,}125,$$
$$\psi_4 = -0{,}0625,\ \psi_5 = -0{,}03125.$$

(c) Il valore atteso di un processo AR(1) senza costante è nullo. Infatti si ha

$$E(Y_t) = E(0.5Y_{t-1} + \varepsilon_t) = 0{,}5E(Y_{t-1}) + E(\varepsilon_t);$$

data la stazionarietà del processo si può considerare $E(Y_t) = E(Y_{t-1})$, pertanto

$$E(Y_t) = 0{,}5E(Y_t) \Leftrightarrow E(Y_t) = 0.$$

Per derivare la funzione di autocorrelazione globale è necessario derivare la funzione di autocovarianza. Questa è data da

$$\begin{aligned}\gamma_k &= E[Y_{t-k}Y_t] = E[Y_{t-k}(\phi Y_{t-1} + \varepsilon_t)] \\ &= \phi E[Y_{t-k}Y_{t-1}] + E[Y_{t-k}\varepsilon_t] \\ &= \phi\gamma_{k-1} + E[Y_{t-k}\varepsilon_t].\end{aligned}$$

Poiché

$$E[Y_{t-k}\varepsilon_t] = \begin{cases}\sigma^2 & k = 0 \\ 0 & k > 0\end{cases}$$

risulta essere $\gamma_k = \phi\gamma_{k-1}$. Per ripetizione successive si ottiene la ricorsione $\gamma_k = \phi^k\gamma_0$. Pertanto, la funzione di autocorrelazione globale di un processo AR(1) è pari a

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k \quad k \geq 0.$$

In questo caso abbiamo quindi

$$\rho_k = 0{,}5^k \quad k \geq 0.$$

**6.11** Dato il processo

$$Y_t = 2Y_{t-1} + \varepsilon_t - 4\varepsilon_{t-2}$$

con $\varepsilon_t \sim WN(0, \sigma^2)$, se ne discutano stazionarietà ed invertibilità.

□ Il processo può essere scritto in forma compatta

$$(1 - 2B)Y_t = (1 - 4B^2)\varepsilon_t.$$

Cerchiamo anzitutto le radici dei due polinomi, ossia le soluzioni delle equazioni $\phi(B) = 0$ e $\theta(B) = 0$:

$$\begin{aligned}1 - 2B = 0 &\Rightarrow B = 0{,}5 \\ 1 - 4B^2 = 0 &\Rightarrow B_{1/2} = \begin{cases}-0{,}5 \\ 0{,}5.\end{cases}\end{aligned}$$

I polinomi $\phi(B)$ e $\theta(B)$ hanno dunque radici (in valore assoluto) minori di uno, il chè farebbe pensare ad un processo non stazionario e non invertibile. Tuttavia, una delle radici in questione è comune (e pari a 0,5), per cui è possibile semplificare il processo, riducendone gli ordini. Infatti, poiché

$$\theta(B) = 1 - 4B^2 = (1 - 2B)(1 + 2B),$$

il processo si semplifica nel modo seguente:

$$Y_t = (1 + 2B)\varepsilon_t \Rightarrow Y_t = \varepsilon_t + 2\varepsilon_{t-1}.$$

Il processo è dunque un MA(1) non invertibile senza alcun problema di stazionarietà.

**6.12** Dato il processo

$$Y_t = 2{,}5Y_{t-1} - Y_{t-2} + \varepsilon_t - 2{,}25\varepsilon_{t-1} + 0{,}5\varepsilon_{t-2}$$

con $\varepsilon_t \sim WN(0, \sigma^2)$, se ne discutano stazionarietà ed invertibilità.

□ Il processo si scrive in forma compatta

$$(1 - 2{,}5B + B^2)Y_t = (1 - 2{,}25B + 0{,}5B^2)\varepsilon_t.$$

Risolvendo le equazioni caratteristiche di entrambi i polinomi, ovvero $\phi(B) = 0$ e $\theta(B) = 0$, si ha:

$$B^2 - 2{,}5B + 1 = 0 \Rightarrow B_{1/2} = \frac{2{,}5 \pm \sqrt{6{,}25 - 4}}{2} =$$

$$B_{1/2} = \frac{2{,}5 \pm 1{,}5}{2} = \begin{cases} 2 \\ 0{,}5 \end{cases}$$

$$0{,}5B^2 - 2{,}25B + 1 = 0 \Rightarrow B_{1/2} = 2{,}25 \pm \sqrt{5{,}0625 - 2} =$$

$$B_{1/2} = 2{,}25 \pm 1{,}75 = \begin{cases} 4 \\ 0{,}5. \end{cases}$$

I polinomi $\phi(B)$ e $\theta(B)$ hanno dunque una radice (in valore assoluto) minore di uno. Anche in questo caso si è portati a pensare ad un processo non stazionario e non invertibile. Tuttavia, è possibile semplificare il processo a causa di una radice comune tra i polinomi, e cioè 0,5. Infatti, poiché

$$\phi(B) = (0{,}5 - B)(2 - B)$$
$$\theta(B) = 0{,}5(0{,}5 - B)(4 - B)$$

il processo $Y_t$ si semplifica nel modo seguente:

$$(2 - B)Y_t = 0{,}5(4 - B)\varepsilon_t \Rightarrow 2Y_t - Y_{t-1} = 2\varepsilon_t - 0{,}5\varepsilon_{t-1}$$

ossia

$$Y_t - 0{,}5Y_{t-1} = \varepsilon_t - 0{,}25\varepsilon_{t-1}$$

che è un ARMA(1,1) stazionario ed invertibile. La radice rimossa determinava l'assenza apparente di queste condizioni nel processo iniziale.

**6.13** Dato il processo

$$Y_t = 0{,}5Y_{t-1} - 2Y_{t-2} + Y_{t-3} + \varepsilon_t + 2\varepsilon_{t-2}$$

con $\varepsilon_t \sim WN(0, \sigma^2)$, se ne discutano stazionarietà ed invertibilità.

□ Scriviamo il processo in forma compatta:

$$(1 - 0{,}5B + 2B^2 - B^3)Y_t = (1 + 2B^2)\varepsilon_t.$$

Poiché, come è immediato verificare,

$$1 - 0{,}5B + 2B^2 - B^3 = (1 - 0{,}5B)(1 + 2B^2),$$

il processo si semplifica nel modo seguente:

$$(1 - 0{,}5B)Y_t = \varepsilon_t$$

ossia $Y_t = 0{,}5Y_{t-1} + \varepsilon_t$. Il processo è un AR(1) con parametro inferiore ad uno in valore assoluto e, dunque, stazionario ed invertibile.

**6.14** Dato il processo

$$Y_t = 1{,}4Y_{t-1} - 0{,}4Y_{t-2} + \varepsilon_t - 0{,}1\varepsilon_{t-1} - 0{,}12\varepsilon_{t-2},$$

(a) dire, giustificando la risposta, se è stazionario;

(b) si calcoli $\rho_2$ per il processo $Z_t = Y_t - Y_{t-1}$.

□ (a) Il processo $Y_t$ si può scrivere nella forma compatta

$$(1 - 1{,}4B + 0{,}4B^2)Y_t = (1 - 0{,}1B - 0{,}12B^2)\varepsilon_t.$$

Cerchiamo le radici delle equazioni caratteristiche $\phi(B) = 0$ e $\theta(B) = 0$. Nel primo caso si ha

$$\begin{aligned} 1 - 1{,}4B + 0{,}4B^2 = 0 \Rightarrow B_{1/2} &= \frac{1{,}4 \pm \sqrt{1{,}96 - 1{,}6}}{0{,}8} \\ &= \frac{1{,}4 \pm 0{,}6}{0{,}8} = \begin{cases} \frac{10}{4} \\ 1 \end{cases} \end{aligned}$$

mentre nel secondo

$$\begin{aligned} 1 - 0{,}1B - 0{,}12B^2 = 0 \Rightarrow B_{1/2} &= \frac{0{,}1 \pm \sqrt{0{,}01 + 0{,}48}}{-0{,}24} \\ &= -\frac{0{,}1 \pm 0{,}7}{0{,}24} = \begin{cases} -\frac{10}{3} \\ \frac{10}{4}. \end{cases} \end{aligned}$$

Il processo può dunque essere scritto nella forma seguente:

$$(1 - 0{,}4B)(1 - B)Y_t = (1 - 0{,}4B)(1 + 0{,}3B)\varepsilon_t.$$

Dividendo a destra e a sinistra per $(1 - 0{,}4B)$ il processo viene semplificato in

$$(1 - B)Y_t = (1 + 0{,}3B)\varepsilon_t.$$

Il processo $Y_t$ è dunque un processo ARIMA(0,1,1), che ovviamente non è stazionario.

(b) Il processo differenziato $Z_t = Y_t - Y_{t-1}$ è invece un MA(1). Infatti,

$$\begin{aligned} Y_t - Y_{t-1} &= \varepsilon_t + 0{,}3B\varepsilon_t \\ Z_t &= (1 + 0{,}3B)\varepsilon_t. \end{aligned}$$

L'autocorrelazione globale di un MA(1), come è noto, è diversa da zero solo al ritardo 1: pertanto, $\rho_2 = 0$.

**6.15** Si dica, giustificando la risposta, se il processo

$$(1 - 0{,}5B)(1 - 0{,}7B)(1 - 0{,}2B)Y_t = \varepsilon_t$$

con $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ , è stazionario e invertibile e lo si riscriva utilizzando l'usuale rappresentazione $\phi(B)Y_t = \theta(B)\varepsilon_t$.

□ Il processo $Y_t$ è un $AR(3)$, quindi certamente invertibile. Esso è anche stazionario, dal momento che è evidente che tutte le radici del polinomio

$$(1-0{,}5B)(1-0{,}7B)(1-0{,}2B)$$

sono in modulo maggiori di uno.

Essendo

$$\begin{aligned}\phi(B) &= (1-0{,}5B)(1-0{,}7B)(1-0{,}2B)\\ &= (1-0{,}7B-0{,}5B+0{,}35B^2)(1-0{,}2B)\\ &= (1-1{,}2B+0{,}35B^2)(1-0{,}2B)\\ &= 1-1{,}2B+0{,}35B^2-0{,}2B+0{,}24B^2-0{,}07B^3\\ &= 1-1{,}4B+0{,}59B^2-0{,}07B^3\end{aligned}$$

il processo può essere scritto come

$$(1-1{,}4B+0{,}59B^2-0{,}07B^3)Y_t=\varepsilon_t$$

o, equivalentemente,

$$Y_t = 1{,}4Y_{t-1}-0{,}59Y_{t-2}+0{,}07Y_{t-3}+\varepsilon_t.$$

**6.16** Dati i seguenti modelli:

(a) $X_t = 0{,}8X_{t-1}-0{,}15X_{t-2}+\varepsilon_t-0{,}3\varepsilon_{t-1}$

(b) $X_t = X_{t-1}-0{,}5X_{t-2}+\varepsilon_t-\varepsilon_{t-1}$

(c) $X_t = 0{,}5X_{t-1}+\varepsilon_t-1{,}2\varepsilon_{t-1}+0{,}35\varepsilon_{t-2}$

dove $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$:

1. esprimerli utilizzando l'operatore ritardo $B$;
2. classificarli come un $ARIMA(p,d,q)$ (ossia, trovare $p$, $d$ e $q$);
3. verificare se tali modelli sono rappresentati nella forma più parsimoniosa;
4. determinare, per ciascuno, se il modello è stazionario e/o invertibile.

N.B. Si ricordi che il modulo di un numero complesso $a \pm ib$ è $|a+ib| = \sqrt{a^2+b^2}$.

□ Vale quanto segue:

(1.a) $(1-0{,}8B+0{,}15B^2)X_t = (1-0{,}3B)\varepsilon_t$.

(1.b) $(1-B-0{,}5B^2)X_t = (1-B)\varepsilon_t$.

(1.c) $(1-0{,}5B)X_t = (1-1{,}2B+0{,}35B^2)\varepsilon_t$.

(2.a ) Apparentemente $X_t$ è un $ARMA(2,1)$. Risolvendo l'equazione caratteristica

$$0{,}15B^2-0{,}8B+1=0 \qquad \rightarrow \qquad B_{1/2} = \begin{cases} \dfrac{0{,}8-\sqrt{0{,}64-0{,}6}}{0{,}3} = 2 \\ \dfrac{0{,}8+\sqrt{0{,}64+0{,}6}}{0{,}3} = \dfrac{10}{3} \end{cases}$$

e pertanto il polinomio $AR$ può essere fattorizzato in $(1-0{,}5B)(1-0{,}3B)$.

(2.b ) Apparentemente $X_t$ è un $ARMA(2,1)$ non invertibile ($\theta = 1$). Risolvendo l'equazione caratteristica

$$0{,}5B^2 + B - 1 = 0 \qquad \rightarrow \qquad B_{1/2} = \begin{cases} \dfrac{-1-\sqrt{1-2}}{1} = -1-\sqrt{-1} \\ \dfrac{-1+\sqrt{1-2}}{1} = -1+\sqrt{-1} \end{cases}$$

Per entrambe le radici dell'equazione il modulo è pari a $\sqrt{2} > 1$. Sicchè il modello è stazionario ma non è invertibile.

(2.c ) Apparentemente $X_t$ è un $ARMA(1,2)$. La radice dell'equazione caratteristica della parte $AR$ è pari a 2, mentre per la parte $MA$ si ha:

$$0{,}35B^2 - 1{,}2B + 11 = 0 \qquad \rightarrow \qquad B_{1/2} = \begin{cases} \dfrac{1{,}2-\sqrt{1{,}44-1{,}4}}{0{,}7} = \dfrac{10}{7} \\ \dfrac{1{,}2+\sqrt{1{,}44-1{,}4}}{0{,}7} = 2 \end{cases}$$

e pertanto il polinomio $MA$ può essere fattorizzato in $(1 - 0{,}5B)(1 - 0{,}7B)$.

(3.a) Semplificando il fattore comune $(1 - 0{,}3B)$, il modello diventa $(1 - 0{,}5B)X_t = \varepsilon_t$, che è un $AR(1)$.

(3.b) Questo modello non può essere riportato ad una forma più parsimoniosa.

(3.c) Semplificando il fattore comune $(1 - 0{,}5B)$, il modello diventa $X_t = (1 - 0{,}7B)\varepsilon_t$, che è un $MA(1)$.

(4.a) Il modello è un $AR(1)$ stazionario.

(4.b) Il modello è un $ARMA(2,1)$ stazionario e non invertibile

(4.c) Il modello è un $MA(1)$ invertibile.

## Autocovarianze ed autocorrelazioni

**6.17** Derivare la funzione di autocovarianza per un processo MA(1).

□ Un processo MA(1) assume la forma

$$Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} = (1 - \theta_1 B)\varepsilon_t.$$

In assenza di costante il processo ha media nulla, pertanto la varianza del processo, indicata con $\gamma_0$, è

$$\gamma_0 = Var(Y_t) = E(Y_t^2).$$

Sviluppando l'ultimo termine otteniamo

$$\begin{aligned} E(Y_t^2) &= E(Y_t Y_t) \\ &= E[Y_t(\varepsilon_t - \theta_1\varepsilon_{t-1})] \\ &= E(Y_t\varepsilon_t) - \theta_1 E(y_t\varepsilon_{t-1}). \end{aligned}$$

Procedendo ancora con la sostituzione di $Y_t$ si ha

$$\begin{aligned}\gamma_0 &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1})\varepsilon_t] - \theta_1 E[(\varepsilon_t - \theta_1\varepsilon_{t-1})\varepsilon_{t-1}] \\ &= E(\varepsilon_t^2) + \theta_1^2 E(\varepsilon_{t-1}^2) \\ &= \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 \\ &= (1+\theta_1^2)\sigma_\varepsilon^2. \qquad (6.3)\end{aligned}$$

Con lo stesso procedimento riusciamo a calcolarci l'autocovarianza di ordine $k > 0$, indicata con $\gamma_k$. Infatti,

$$\begin{aligned}\gamma_k = E(Y_{t-k}Y_t) &= E[Y_{t-k}(\varepsilon_t - \theta_1\varepsilon_{t-1})] \\ &= E(Y_{t-k}\varepsilon_t) - \theta_1 E(Y_{t-k}\varepsilon_{t-1}).\end{aligned}$$

Ora, per $k = 1$ abbiamo

$$\begin{aligned}E(Y_{t-1}\varepsilon_t) &= E[(\varepsilon_{t-1} - \theta_1\varepsilon_{t-2})\varepsilon_t] = 0 \\ E(Y_{t-1}\varepsilon_{t-1}) &= E[(\varepsilon_{t-1} - \theta_1\varepsilon_{t-2})\varepsilon_{t-1}] = \sigma_\varepsilon^2.\end{aligned}$$

e dunque

$$\gamma_1 = -\theta_1\sigma_\varepsilon^2.$$

Per $k = 2$ si ha invece

$$\begin{aligned}E(Y_{t-2}\varepsilon_t) &= E[(\varepsilon_{t-2} - \theta_1\varepsilon_{t-3})\varepsilon_t] = 0 \\ E(Y_{t-2}\varepsilon_{t-1}) &= E[(\varepsilon_{t-2} - \theta_1\varepsilon_{t-3})\varepsilon_{t-1}] = 0\end{aligned}$$

quindi $\gamma_2 = 0$. Per $k > 2$ i valori attesi continuano ad essere nulli. Riassumendo, la funzione di autocovarianza di un processo MA(1) è pari a

$$\gamma_k = \begin{cases}(1+\theta_1^2)\sigma_\varepsilon^2 & \text{per } k = 0 \\ -\theta_1\sigma_\varepsilon^2 & \text{per } k = 1 \\ 0 & \text{per } k \geq 2.\end{cases} \qquad (6.4)$$

**6.18** Derivare la funzione di autocovarianza per un processo AR(1).

□ Un processo AR(1) senza costante assume la forma

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t,$$

con $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Procedendo con lo stesso metodo utilizzato nel precedente esercizio, scriviamo la formula generale per l'autocovarianza $\gamma_k$ ed operiamo sostituzione degli elementi:

$$\begin{aligned}\gamma_k = E(Y_{t-k}Y_t) &= E[Y_{t-k}(\phi_1 Y_{t-1} + \varepsilon_t)] \\ &= \phi_1 E(Y_{t-k}Y_{t-1}) + E(Y_{t-k}\varepsilon_t) \\ &= \phi_1\gamma_{k-1} + E(Y_{t-k}\varepsilon_t).\end{aligned}$$

Poiché

$$E(Y_{t-k}\varepsilon_t) = \begin{cases}\sigma_\varepsilon^2 & \text{per } k = 0 \\ 0 & \text{per } k > 0,\end{cases}$$

risulta essere $\gamma_k = \phi_1 \gamma_{k-1}$. Per sostituzioni successive si ricava la relazione

$$\gamma_k = \phi_1^k \gamma_0$$

che definisce in maniera ricorsiva la funzione di autocovarianza. Resta da determinare la varianza del processo $\gamma_0$. Essa può esprimersi come

$$\begin{aligned} E(Y_t^2) &= E(Y_t Y_t) \\ &= E[(\phi_1 Y_{t-1} + \varepsilon_t)(\phi_1 Y_{t-1} + \varepsilon_t)] \\ &= \phi_1^2 E(Y_{t-1}^2) + E(\varepsilon_t^2) \end{aligned} \tag{6.5}$$

essendo il termine $E(Y_{t-1}\varepsilon_t)$ nullo. Considerando l'ipotesi di stazionarietà del processo AR(1), è lecito ipotizzare $E(Y_t^2) = E(Y_{t-1}^2) = \gamma_0$. Pertanto, la (6.5) si modifica nella

$$\begin{aligned} \gamma_0 &= \phi_1^2 \gamma_0 + \sigma_\varepsilon^2 \\ (1 - \phi_1^2)\gamma_0 &= \sigma_\varepsilon^2 \\ \gamma_0 &= \frac{\sigma_\varepsilon^2}{1 - \phi_1^2}. \end{aligned}$$

Riassumendo, la funzione di autocovarianza di un processo AR(1) è

$$\gamma_k = \begin{cases} \dfrac{\sigma_\varepsilon^2}{1 - \phi_1^2} & \text{per } k = 0 \\ \phi_1^k \gamma_0 & \text{per } k > 0. \end{cases}$$

**6.19** Derivare la funzione di autocovarianza per un processo ARMA(1,1).

□ Un processo ARMA(1,1) è dato da

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

con $\varepsilon_t \sim WN(0,\ \sigma_\varepsilon^2)$. Assumendo per comodità $\phi_0 = 0$ (cioè che il processo abbia media nulla) ed indicando $Var(Y_t) = \gamma_0$, si ha:

$$\begin{aligned} \gamma_0 = E(Y_t^2) &= E[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1})^2] \\ &= E[\phi_1^2 y_{t-1}^2 + \varepsilon_t^2 + \theta_1^2 \varepsilon_{t-1}^2 + 2\phi_1 Y_{t-1}\varepsilon_t - 2\theta_1 \phi_1 Y_{t-1}\varepsilon_{t-1} \\ &\quad -2\theta_1 \varepsilon_t \varepsilon_{t-1}]. \end{aligned}$$

Nello sviluppare questa espressione conviene tenere presente che:

- $E(Y_{t-1}^2) = E(Y_t^2) = \gamma_0$ perché il processo è stazionario;
- $E(\varepsilon_{t-1}^2) = E(\varepsilon_t^2) = \sigma_\varepsilon^2$;
- $E(Y_{t-1}\varepsilon_t) = 0$ perché $Y_{t-1}$ dipende da $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$;
- $E(Y_{t-1}\varepsilon_{t-1}) = E[(\phi_1 Y_{t-2} + \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2})\varepsilon_{t-1}] = \sigma_\varepsilon^2$.

Pertanto:

$$\begin{aligned}\gamma_0 &= \phi_1^2\gamma_0 + \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 - 2\theta_1\phi_1\sigma_\varepsilon^2 \\ &\Rightarrow (1-\phi_1^2)\gamma_0 = (1+\theta_1^2-2\theta_1\phi_1)\sigma_\varepsilon^2 \\ &\Rightarrow \gamma_0 = \frac{(1+\theta_1^2-2\theta_1\phi_1)}{1-\phi_1^2}\sigma_\varepsilon^2. \qquad (6.6)\end{aligned}$$

Calcoliamo ora la funzione di autocovarianza $\gamma_k$ per $k > 0$. Per $k = 1$ si ha:

$$\begin{aligned}\gamma_1 &= E[Y_tY_{t-1}] = E[Y_{t-1}(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1})] \\ &= E[\phi_1 Y_{t-1}^2 + Y_{t-1}\varepsilon_t - \theta_1 Y_{t-1}\varepsilon_{t-1}] \\ &= \phi_1\gamma_0 - \theta_1\sigma_\varepsilon^2 \\ &= \left(\phi_1\frac{1+\theta_1^2-2\theta_1\phi_1}{1-\phi_1^2} - \theta_1\right)\sigma_\varepsilon^2 \\ &= \frac{(\phi_1-\theta_1)(1-\phi_1\theta_1)}{1-\phi_1^2}\sigma_\varepsilon^2 \qquad (6.7)\end{aligned}$$

Per $k > 1$, invece

$$\begin{aligned}\gamma_k &= E[Y_tY_{t-k}] = E[Y_{t-k}(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1})] \\ &= E[\phi_1 Y_{t-1}Y_{t-k} + Y_{t-k}\varepsilon_t - \theta_1 Y_{t-k}\varepsilon_{t-1}] \\ &= \phi_1\gamma_{k-1} = \phi_1^{k-1}\gamma_1.\end{aligned}$$

**6.20** Derivare la funzione di autocovarianza per un processo MA(2).

□ Un processo MA(2) assume la forma

$$Y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}.$$

Il processo ha media nulla, pertanto la funzione di autocovarianza si scrive come al solito

$$\begin{aligned}\gamma_k &= E(Y_tY_{t-k}) \\ &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2})Y_{t-k}].\end{aligned}$$

La funzione di autocovarianza si esplicita sostituendo ad $Y_{t-k}$ la forma MA(2) ai vari periodi di ritardo. Iniziamo con $k = 0$, ovvero con la varianza del processo:

$$\begin{aligned}\gamma_0 &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2})(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2})] \\ &= E(\varepsilon_t^2) + \theta_1^2E(\varepsilon_{t-1}^2) + \theta_2^2E(\varepsilon_{t-2}^2) \\ &= \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 + \theta_2^2\sigma_\varepsilon^2 \\ &= \sigma_\varepsilon^2(1+\theta_1^2+\theta_2^2). \qquad (6.8)\end{aligned}$$

Per $k = 1$ abbiamo

$$\begin{aligned}\gamma_1 &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2})(\varepsilon_{t-1} - \theta_1\varepsilon_{t-2} - \theta_2\varepsilon_{t-3})] \\ &= -\theta_1E(\varepsilon_{t-1}^2) + \theta_1\theta_2E(\varepsilon_{t-2}^2) \\ &= -\theta_1\sigma_\varepsilon^2 + \theta_1\theta_2\sigma_\varepsilon^2 \\ &= \sigma_\varepsilon^2(-\theta_1 + \theta_1\theta_2). \qquad (6.9)\end{aligned}$$

Per $k = 2$

$$\begin{aligned}\gamma_2 &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2})(\varepsilon_{t-2} - \theta_1\varepsilon_{t-3} - \theta_2\varepsilon_{t-4})] \\ &= -\theta_2 E(\varepsilon_{t-2}^2) \\ &= -\theta_2\sigma_\varepsilon^2. \qquad (6.10)\end{aligned}$$

A partire dal ritardo 3 l'autocovarianza si annulla: la memoria di un processo MA(2) è limitata a 2 soli ritardi. Riassumendo, la funzione di autocovarianza è

$$\gamma_k = \begin{cases} \sigma_\varepsilon^2(1 + \theta_1^2 + \theta_2^2) & \text{per } k = 0 \\ \sigma_\varepsilon^2(-\theta_1 + \theta_1\theta_2) & \text{per } k = 1 \\ -\theta_2\sigma_\varepsilon^2 & \text{per } k = 2 \\ 0 & \text{per } k > 2. \end{cases} \qquad (6.11)$$

**6.21** Derivare la funzione di autocovarianza per un processo AR(2).

□ Un processo AR(2) si esprime come

$$Y_t = \phi_1 Y_{t-1}\phi_2 Y_{t-2} + \varepsilon_t \qquad (6.12)$$

con $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$. Per derivare la funzione di autocovarianza di questo processo utilizziamo una logica differente rispetto a quella adottata finora. Assunta nulla la costante, moltiplichiamo entrambi i membri dell'espressione (6.12) per $Y_{t-k}$ e prendiamone i valori attesi:

$$E(Y_t Y_{t-k}) = \phi_1 E(Y_{t-1}Y_{t-k}) + \phi_2 E(Y_{t-2}Y_{t-k}) + E(\varepsilon_t Y_{t-k})$$

Per $k = 0$ possiamo scrivere

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma_\varepsilon^2 \qquad (6.13)$$

mentre per $k > 0$ vale la relazione ricorsiva

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2}. \qquad (6.14)$$

Dividendo ambo i membri della (6.13) per $\gamma_0$ si ricava che la varianza del processo può essere espressa nel modo seguente:

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \phi_1\rho_1 - \phi_2\rho_2},$$

ovvero in funzione dei parametri del processo e delle prime due autocorrelazioni. Questa espressione della varianza richiede quindi preliminarmente il calcolo della funzione di autocorrelazione, che può essere calcolata indipendentemente dalla varianza sfruttando le relazioni (6.14) per $k > 0$. Infatti, dividendo ancora ambo i membri di questa espressione per $\gamma_0$ otteniamo

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2}.$$

Scrivendo per esteso la precedente espressione per $k = 1$ e $k = 2$ otteniamo il sistema di equazioni lineari

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_1 \\ \rho_2 &= \phi_1\rho_1 + \phi_2\end{aligned}$$

tenuto conto che $\rho_0 = 1$ e $\rho_i = \rho_{-i}$. Le equazioni precedenti sono note come *equazioni di Yule-Walker* e consentono di calcolare le autocorrelazioni di un processo autoregressivo a partire dal valore dei parametri (e viceversa).

Risolvendo il precedente sistema si ottiene:

$$(1-\phi_2)\,\rho_1 = \phi_1 \quad \rightarrow \quad \rho_1 = \frac{\phi_1}{1-\phi_2},$$

$$\rho_2 = \frac{\phi_2^2}{1-\phi_2} + \phi_2 \quad \rightarrow \quad \rho_2 = \frac{\phi_1^2 + \phi_2 - \phi_2^2}{1-\phi_2}.$$

e

$$\rho_3 = \phi_1\rho_2 + \phi_2\rho_1 \rightarrow \frac{\phi_1^3 + \phi_1\phi_2 - \phi_1\phi_2^2}{1-\phi_2} + \frac{\phi_2\phi_1}{1-\phi_2} = \frac{\phi_1^3 + 2\phi_1\phi_2 - \phi_1\phi_2^2}{1-\phi_2}$$

Una volta note le autocorrelazioni $\rho_1$ e $\rho_2$ è quindi possibile calcolare la varianza $\gamma_0$ dalla (6.13). Le autocovarianze possono essere infine desunte moltiplicando le autocorrelazioni per $\gamma_0$.

**6.22** Derivare la funzione di autocorrelazione globale per il processo

$$Y_t = \varepsilon_t + \delta\varepsilon_{t-2},$$

con $\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$ e $|\delta| < 1$.

□ Il processo è a media nulla, pertanto la varianza può essere espressa come

$$\begin{aligned}\gamma_0 &= E(Y_t^2)\\ &= E[(\varepsilon_t + \delta\varepsilon_{t-2})^2]\\ &= E(\varepsilon_t^2) + \delta^2 E(\varepsilon_{t-2}^2) + 2\delta E(\varepsilon_t\varepsilon_{t-2}).\end{aligned}$$

L'ultimo termine si annulla poiché il processo $\varepsilon_t$ è un *white noise*, pertanto risolvendo si ha

$$\begin{aligned}\gamma_0 &= \sigma_\varepsilon^2 + \delta^2\sigma_\varepsilon^2\\ &= \sigma_\varepsilon^2(1+\delta^2).\end{aligned}$$

La prima autocovarianza è invece nulla, in base alle seguenti relazioni

$$\begin{aligned}\gamma_1 &= E(Y_tY_{t-1})\\ &= E(\varepsilon_tY_{t-1}) + \delta E(\varepsilon_{t-2}Y_{t-1}).\end{aligned}$$

Poiché $Y_{t-1}$ è incorrelato sia a $\varepsilon_t$ sia a $\varepsilon_{t-2}$ l'autocovarianza al primo ritardo è nulla. Verifichiamo ora l'autocovarianza al secondo ritardo $\gamma_2$.

$$\begin{aligned}\gamma_2 &= E(Y_tY_{t-2})\\ &= E(\varepsilon_tY_{t-2}) + \delta E(\varepsilon_{t-2}Y_{t-2})\\ &= \delta E(\varepsilon_{t-2}^2) = \delta\sigma_\varepsilon^2\end{aligned}$$

Procedendo per $k = 3$ otteniamo

$$\begin{aligned}\gamma_3 &= E(Y_tY_{t-3})\\ &= E(\varepsilon_tY_{t-3}) + \delta E(\varepsilon_{t-2}Y_{t-3})\end{aligned}$$

cioè nuovamente un espressione nulla. Per il processo $Y_t$ la funzione di autocovarianza $\gamma_k$ si annulla per $k \geq 3$. Riassumendo, la funzione di autocorrelazione del processo $Y_t$ risulta pari a

$$\rho_k = \begin{cases} 0 & \text{per } k = 1. \\ \dfrac{\delta}{1+\delta^2} & \text{per } k = 2 \\ 0 & \text{per } k \geq 3. \end{cases}$$

**6.23** Derivare la funzione di autocorrelazione globale per il processo

$$Y_t = \delta Y_{t-2} + \varepsilon_t,$$

con $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$ e $|\delta| < 1$.

□ La varianza del processo equivale a

$$\begin{aligned} \gamma_0 &= E(Y_t^2) \\ &= \delta^2 E(Y_{t-2}^2) + E(\varepsilon_t^2) + 2\delta E(Y_{t-2}\varepsilon_t) \\ &= \delta^2 \gamma_0 + \sigma^2_\varepsilon \end{aligned}$$

ovvero

$$\gamma_0 = \frac{\sigma^2_\varepsilon}{1-\delta^2}$$

cioè la stessa forma di un processo AR(1). Calcoliamo ora l'autocovarianza di ordine 1, cioè $\gamma_1$:

$$\begin{aligned} \gamma_1 &= E(Y_{t-1}Y_t) \\ &= \delta E(Y_{t-1}Y_{t-2}) + E(Y_{t-1}\varepsilon_t) \end{aligned}$$

che, poiché l'ultimo termine si annulla, diventa

$$\gamma_1 = \delta\gamma_1. \tag{6.15}$$

L'unico valore possibile di $\gamma_1$ per cui la relazione (6.15) è soddisfatta è 0, pertanto $\gamma_1 = 0$. Verifichiamo ora cosa succede per $\gamma_2$:

$$\begin{aligned} \gamma_2 &= E(Y_{t-2}Y_t) \\ &= \delta E(Y_{t-2}^2) + E(Y_{t-2}\varepsilon_t) \end{aligned}$$

che è uguale a

$$\gamma_2 = \delta\gamma_0 = \frac{\delta}{1-\delta^2}\sigma^2_\varepsilon.$$

Per $k = 3$ abbiamo

$$\begin{aligned} \gamma_3 &= E(Y_{t-3}Y_t) \\ &= \delta E(Y_{t-3}Y_{t-2}) + E(Y_{t-3}\varepsilon_t) \end{aligned}$$

per cui

$$\gamma_3 = \delta\gamma_1$$

da cui $\gamma_3 = 0$. L'autocovarianza di ordine 4 è

$$\begin{aligned}\gamma_4 &= E(Y_{t-4}Y_t) \\ &= \delta E(Y_{t-4}Y_{t-2}) + E(Y_{t-4}\varepsilon_t)\end{aligned}$$

cioè

$$\begin{aligned}\gamma_4 &= \delta\gamma_2 \\ &= \frac{\delta^2}{1-\delta^2}\sigma_\varepsilon^2.\end{aligned}$$

In generale si ha $\gamma_k = \delta\gamma_{k-2}$ per $k \geq 2$ e pertanto la funzione di autocorrelazione del processo vale

$$\rho_k = \begin{cases} 0 & \text{per } k = 1 \\ \delta\rho_{k-2} & \text{per } k \geq 2. \end{cases}$$

**6.24** Calcolare l'autocorrelazione parziale di ordine 3 per il processo

$$Y_t = 0{,}512Y_{t-3} + \varepsilon_t - 0{,}8\varepsilon_{t-1},$$

con $\varepsilon_t \sim WN(0,\sigma_\varepsilon^2)$.

□ Il processo può essere scritto nella forma compatta

$$(1 - 0{,}512B^3)Y_t = (1 - 0{,}8B)\varepsilon_t.$$

Poiché $0{,}512 = 0{,}8^3$, si ha

$$\begin{aligned}\phi(B) &= 1 - 0{,}512B^3 = 1 - 0{,}8^3B^3 \\ &= (1 - 0{,}8B)(1 + 0{,}8B + 0{,}64B^2),\end{aligned}$$

e pertanto il processo si semplifica nel modo seguente:

$$(1 + 0{,}8B + 0{,}64B^2)Y_t = \varepsilon_t.$$

Si tratta di un processo AR(2), la cui funzione di autocorrelazione parziale è nulla per $k \geq 3$. Pertanto, $P_3 = 0$.

**6.25** Si consideri il processo MA(1)

$$Y_t = (1 - \theta B)\varepsilon_t$$

con $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$ e $|\theta| < 1$.

(a) Calcolare e confrontare la varianza del processo $Y_t$ e quella del processo $(1 - B)Y_t$. Commentare.

(b) Calcolare la funzione di autocorrelazione per il processo $Y_t$ e per il processo $(1 - B)Y_t$.

□ (a) La varianza di un processo MA(1) è pari a

$$\gamma_0' = \sigma_\varepsilon^2(1+\theta^2),$$

come dimostrato nella (6.3). Calcoliamo invece la varianza del processo $(1-B)Y_t$. Il processo può essere scritto nella forma

$$\begin{aligned}(1-B)Y_t &= Y_t - Y_{t-1} \\ &= \varepsilon_t - \theta\varepsilon_{t-1} - \varepsilon_{t-1} + \theta\varepsilon_{t-2} \\ &= \varepsilon_t - (1+\theta)\varepsilon_{t-1} + \theta\varepsilon_{t-2}.\end{aligned}$$

Esso risulta quindi essere un processo MA(2) con parametri vincolati $\theta_1 = 1+\theta$ e $\theta_2 = -\theta$. Applicando la formula (6.8), la varianza del processo $(1-B)Y_t$ assume la forma

$$\begin{aligned}\gamma_0'' &= \sigma_\varepsilon^2(1+(1+\theta)^2+(-\theta)^2) \\ &= \sigma_\varepsilon^2(1+1+\theta^2+2\theta+\theta^2) \\ &= 2\sigma_\varepsilon^2(1+\theta+\theta^2).\end{aligned}$$

L'applicazione dell'operatore $(1-B)$ amplifica la varianza del processo MA(1) originario di una quantità pari a $\sigma_\varepsilon^2(1+\theta)^2$. L'utilizzo dell'operatore differenza per un processo stocastico stazionario comporta inevitabilmente un aumento della variabilità, specie nel caso in cui la componente di disturbo sia correlata.

(b) La funzione di autocorrelazione per un processo MA(1) si può derivare abilmente partendo dalla funzione di autocovarianza riportata nella (6.4). Essa equivale a

$$\rho_k' = \begin{cases} -\dfrac{\theta}{1+\theta^2} & \text{per } k=1 \\ 0 & \text{per } k \geq 2. \end{cases}$$

Invece, come detto il processo $(1-B)Y_t$ è riconducibile ad un MA(2) con coefficienti vincolati a $\theta$. L'autocovarianza di un MA(2) è stata già esplicitata nella (6.11), qui di seguito forniamo la forma della funzione di autocorrelazione:

$$\rho_k'' = \begin{cases} \dfrac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & \text{per } k=1 \\ \dfrac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{per } k=2 \\ 0 & \text{per } k>2. \end{cases}$$

Sostituendo a $\theta_1$ e a $\theta_2$ le corrispondenti funzioni di $\theta$, otteniamo per $\rho_1''$ la seguente espressione

$$\begin{aligned}\rho_1'' &= \frac{-1(1+\theta)-(1+\theta)\theta}{2(1+\theta+\theta^2)} \\ &= \frac{-\theta^2-2\theta-1}{2(1+\theta+\theta^2)} \\ &= -\frac{(1+\theta)^2}{2(1+\theta+\theta^2)}.\end{aligned}$$

Per $k = 2$ abbiamo più semplicemente

$$\rho''_2 = \frac{\theta}{2(1+\theta+\theta^2)}.$$

**6.26** Individuare un processo stazionario, specificandone l'ordine ed il valore dei parametri, con la seguente funzione di autocorrelazione parziale

$$P_0 = 1, \quad P_1 = 0{,}4, \quad P_k = 0 \quad \text{per} \quad k \geq 2.$$

□ Il processo AR(1)

$$(1 - \phi B)Y_t = \varepsilon_t$$

possiede la seguente funzione di autocorrelazione parziale

$$P_k = \begin{cases} \phi & \text{per } k = 1 \\ 0 & \text{per } k > 1. \end{cases}$$

La soluzione dell'esercizio è quindi

$$(1 - 0{,}4B)Y_t = \varepsilon_t$$

che è un processo stazionario poiché il coefficiente autoregressivo $\phi$ è inferiore ad 1.

**6.27** Individuare un processo invertibile, specificandone l'ordine ed il valore dei parametri, con la seguente funzione di autocorrelazione globale

$$\rho_0 = 1, \quad \rho_1 = 0{,}4, \quad \rho_k = 0 \text{ per } k \geq 2.$$

□ Un processo dotato di una simile funzione di autocorrelazione globale è il processo MA(1)

$$Y_t = (1 - \theta B)\varepsilon_t,$$

la cui invertibilità è garantita dalla condizione $|\theta| < 1$. Come sappiamo, per tale processo si ha

$$\begin{aligned} \gamma_0 &= \sigma^2_\varepsilon(1+\theta^2) \\ \gamma_1 &= -\sigma^2_\varepsilon\theta \\ \gamma_k &= 0 \quad \text{per } k \geq 2. \end{aligned}$$

L'autocorrelazione del primo ordine è quindi pari a

$$\rho_1 = -\frac{\theta}{1+\theta^2}.$$

Il valore del parametro $\theta$ tale che $\rho_1 = 0{,}4$ si può ricavare risolvendo l'equazione

$$-\frac{\theta}{1+\theta^2} = 0{,}4 \quad \Leftrightarrow \quad 0{,}4\theta^2 + \theta + 0{,}4 = 0.$$

Le soluzioni dell'espressione precedente sono pari a

$$\theta_{1/2} = \frac{-1 \pm \sqrt{1 - 0{,}64}}{0{,}8} = \begin{cases} -2 \\ -0{,}5. \end{cases}$$

Il valore del parametro del processo MA(1) invertibile con $\rho_1 = 0{,}4$ è quindi $\theta = -0{,}5$. Il processo corrispondente è

$$Y_t = \varepsilon_t + 0{,}5\varepsilon_{t-1} = (1 + 0{,}5B)\varepsilon_t.$$

**6.28** Dati i seguenti processi

$$\begin{aligned} Y_t &= 0{,}4Y_{t-1} + 0{,}05Y_{t-2} + \varepsilon_t + 0{,}1\varepsilon_{t-1}, \quad \varepsilon_t \sim WN(0{,}9) \\ z_t &= 0{,}5z_{t-1} + \nu_t, \quad \nu_t \sim WN(0{,}1) \end{aligned}$$

con $E(\varepsilon_r \nu_s) = 0 \; \forall r,s$, calcolare l'autocorrelazione parziale di ordine 2 del processo $w_t = Y_t + z_t$.

□ Scriviamo il processo $Y_t$ nella forma compatta

$$(1 - 0{,}4B - 0{,}05B^2)Y_t = (1 + 0{,}1B)\varepsilon_t.$$

Le radici dell'equazione caratteristica

$$1 - 0{,}4B - 0{,}05B^2 = 0$$

sono pari a

$$B_{1/2} = \frac{0{,}4 \pm \sqrt{0{,}16 + 0{,}2}}{-0{,}1} = \begin{cases} -10 \\ 2. \end{cases}$$

Il polinomio $\phi(B)$ può essere quindi scomposto in

$$\phi(B) = (1 - 0{,}5B)(1 + 0{,}1B),$$

pertanto il processo $Y_t$ si semplifica nel modo seguente:

$$(1 - 0{,}5B)Y_t = \varepsilon_t.$$

Il processo $w_t = Y_t + z_t$ assume quindi la forma semplificata

$$\begin{aligned} w_t &= 0{,}5Y_{t-1} + 0{,}5z_{t-1} + \varepsilon_t + \nu_t \\ &= 0{,}5(Y_{t-1} + z_{t-1}) + \xi_t \\ &= 0{,}5w_{t-1} + \xi_t \end{aligned}$$

con $\xi_t = \varepsilon_t + \nu_t \sim WN(0, \sigma_\xi^2)$ e $\sigma_\xi^2 = \sigma_\varepsilon^2 + \sigma_\nu^2$ (infatti $\varepsilon_t$ e $\nu_t$ sono incorrelati tra loro). Il processo $w_t$ è dunque un AR(1) stazionario, la cui funzione di autocorrelazione parziale è nulla per ritardi superiori ad 1. Pertanto, $P_2 = 0$.

**6.29** Dato il processo $Y_t = 0{,}2Y_{t-1} + \varepsilon_t$, dove $\varepsilon_t$ è un *white noise* di varianza unitaria, descrivere le caratteristiche del processo $z_t = Y_t + Y_{t-1}$.

□ Il processo $Y_t$ è un AR(1) stazionario, con varianza pari a

$$\gamma_0 = \frac{1}{1 - 0{,}2^2} = 1.0417$$

ed autocovarianza

$$\gamma_k = 0{,}2\gamma_{k-1} \quad \text{per } k > 0.$$

Il processo $z_t$ può essere esplicitato nel seguente modo:

$$\begin{aligned} z_t &= Y_t + Y_{t-1} = 0{,}2Y_{t-1} + \varepsilon_t + 0{,}2Y_{t-2} + \varepsilon_{t-1} \\ &= 0{,}2(Y_{t-1} + Y_{t-2}) + (1 + B)\varepsilon_t \\ &= 0{,}2z_{t-1} + (1 + B)\varepsilon_t. \end{aligned}$$

Il processo risultante è quindi un ARMA(1,1) stazionario ma non invertibile, poiché la radice del polinomio MA è, in valore assoluto, unitaria. Verifichiamo ora cosa succede alla funzione di autocovarianza. Secondo la (6.6), la varianza del processo ARMA(1,1) $z_t$ è pari a

$$\gamma_0 = \frac{2{,}4}{1 - 0{,}2^2} = 2{,}5\ ,$$

quindi due volte e mezzo superiore a quella del processo originario $Y_t$. Utilizzando l'espressione nella (6.7), la prima autocovarianza risulta essere

$$\gamma_1 = \frac{1{,}44}{0{,}96} = 1{,}5.$$

Per $k > 1$ la funzione di autocovarianza $\gamma_k$ ha la stessa forma dell'AR(1) originario, cioè

$$\gamma_k = 0{,}2\gamma_{k-1}.$$

**6.30** Dato il processo $Y_t = \varepsilon_t - 0{,}4\varepsilon_{t-1}$, dove $\varepsilon_t$ è un *white noise* di varianza unitaria, descrivere le caratteristiche del processo $Z_t = Y_t + Y_{t-1}$.

□ Il processo $Y_t$ è un MA(1) invertibile, con varianza pari a

$$\gamma_0 = (1 + 0{,}4^2) = 1{,}16$$

ed autocovarianza del primo ordine uguale a

$$\gamma_1 = -0{,}4,$$

con $\gamma_k = 0$ per $k > 1$. Il processo $Z_t$ può essere scritto come

$$\begin{aligned} Z_t &= Y_t + Y_{t+1} = \varepsilon_t - 0{,}4\varepsilon_{t-1} + \varepsilon_{t-1} - 0{,}4\varepsilon_{t-2} \\ &= \varepsilon_t + 0{,}6\varepsilon_{t-1} - 0{,}4\varepsilon_{t-2}. \end{aligned}$$

L'applicazione dell'operatore $(1 + B)$ ha trasformato quindi il processo MA(1) originario in un processo MA(2). Per verificare la proprietà di invertibilità del nuovo processo bisogna risolvere l'equazione caratteristica di secondo grado

$$1 + 0{,}6B - 0{,}4B^2 = 0,$$

la quale ha soluzioni

$$B_1 = -1 \text{ e } B_2 = 2{,}5.$$

Una radice è pertanto unitaria, per cui il processo MA(2) $Z_t$ non è invertibile. Sfruttando le espressioni (6.8-6.10) si perviene alla funzione di autocovarianza del nuovo processo

$$\gamma_k = \begin{cases} (1 + 0{,}6^2 + 0{,}4^2) = 1{,}52 & \text{per } k = 0 \\ (0{,}6 - 0{,}6 \times 0{,}4) = 0{,}36 & \text{per } k = 1 \\ -0{,}4 & \text{per } k = 2 \\ 0 & \text{per } k > 2 \end{cases}.$$

## Stima

**6.31** Date due serie storiche $\{x_t\}_{t=1}^T$ e $\{y_t\}_{t=1}^T$ osservate, per un campione di ampiezza $T = 100$, sia nota la seguente matrice dei prodotti incrociati:

| | $y_t$ | $x_t$ | $y_{t-1}$ | $x_{t-1}$ |
|---|---|---|---|---|
| $y_t$ | 100 | 80 | 48 | 24 |
| $x_t$ | 80 | 100 | 0 | 0 |
| $Y_{t-1}$ | 48 | 0 | 100 | 80 |
| $x_{t-1}$ | 24 | 0 | 80 | 100 |

cosicché, ad esempio, $\sum_{t=2}^T y_t x_{t-1} = 24$. Inoltre, è noto che $\sum_{t=1}^T y_t = \sum_{t=1}^T x_t = 0$.

(a) Calcolare l'autocorrelazione campionaria di ordine 1 della serie $y_t$.

(b) Stimare i parametri di un modello AR(1) per $y_t$ (compresa la varianza del *white noise*).

□ La stima dell'autocorrelazione è data da (Di Fonzo e Lisi, 2006, p. 169)

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0},$$

con la varianza e l'autocovarianza campionarie pari rispettivamente a

$$\hat{\gamma}_0 = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{\mu})^2$$

e

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1+k}^{T} (y_t - \hat{\mu})(Y_{t-k} - \hat{\mu}).$$

La condizione $\sum_{t=1}^T y_t = 0$ implica una media campionaria $\hat{\mu}$ nulla, pertanto i dati forniti nella tabella sono sufficienti a calcolare le informazioni richieste.

(a) La varianza campionaria della serie $y_t$ è pari a

$$\hat{\gamma}_0 = \frac{1}{100} \sum_{t=1}^{100} y_t^2 = 1$$

La serie ha dunque varianza unitaria; ne deriva che l'autocorrelazione stimata del primo ordine coincide con l'autocovarianza $\hat{\gamma}_1$. Quest'ultima può essere calcolata attraverso l'espressione

$$\hat{\gamma}_1 = \frac{1}{100}\sum_{t=2}^{100} y_t y_{t-1} = \frac{48}{100} = 0{,}48.$$

Pertanto $\hat{\rho}_1 = 0{,}48$.

(b) Il modello AR(1) per $Y_t$ ha la forma

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

con $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$. I parametri ignoti da stimare sono $\phi$ e $\sigma^2_\varepsilon$, che è la varianza dell'errore $\varepsilon_t$. Nel caso di modelli AR, le stime attraverso il metodo dei minimi quadrati sono efficienti e coincidono con quelle di massima verosimiglianza. Il coefficiente $\phi$ è ottenuto mediante la formula

$$\hat{\phi} = \frac{\sum_{t=2}^{T} y_t y_{t-1}}{\sum_{t=1}^{T} y_t^2} = \frac{48}{100} = 0{,}48.$$

Esiste quindi una corrispondenza esatta tra coefficiente autoregressivo e autocorrelazione del primo ordine. Per quanto riguarda $\sigma^2_\varepsilon$, un buon stimatore è dato da

$$\hat{\sigma}^2_\varepsilon = \frac{1}{T-p}\sum_{t=p+1}^{T} (Y_t - \hat{\phi} Y_{t-1})^2.$$

Espandendo il quadrato del binomio nella sommatoria si ottiene

$$\hat{\sigma}^2_\varepsilon = \frac{1}{T-p}\left[\sum_{t=p+1}^{T} Y_t^2 - 2\hat{\phi}\sum_{t=p+1}^{T} Y_t Y_{t-1} + \hat{\phi}^2 \sum_{t=p+1}^{T} Y_{t-1}^2\right].$$

Sostituendo a questa espressione i valori noti si ottiene la varianza stimata di $\varepsilon$:

$$\begin{aligned}\hat{\sigma}^2_\varepsilon &= \frac{1}{99}[100 - (0{,}96 \times 48) + (0{,}2304 \times 100)] \\ &= \frac{76{,}96}{99} = 0{,}777.\end{aligned}$$

**6.32** Sapendo che $\hat{\rho}_1 = 0{,}03$, $\hat{\rho}_2 = 0{,}08$ e $\hat{\rho}_3 = 0{,}13$, fornire una stima di prima approssimazione dei parametri di un processo AR(3).

□ Per ricavare una stima di prima approssimazione dei parametri del processo AR(3) (assunto, senza perdere di generalità, di media nulla)

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \varepsilon_t$$

possiamo sfruttare le equazioni di *Yule-Walker*:

$$\begin{array}{rcl} \rho_1 &=& \phi_1 + \rho_1\phi_2 + \rho_2\phi_3 \\ \rho_2 &=& \rho_1\phi_1 + \phi_2 + \rho_1\phi_3 \\ \rho_3 &=& \rho_2\phi_1 + \rho_1\phi_2 + \phi_3 \end{array}.$$

Sostituendo $\hat{\rho}_k$ a $\rho_k$, per $k = 1{,}2{,}3$, si ha il seguente sistema di tre equazioni nelle tre incognite $\phi_1, \phi_2, \phi_3$

$$\begin{array}{rcl} 0{,}03 &=& \phi_1 + 0{,}03\phi_2 + 0{,}08\phi_3 \\ 0{,}08 &=& 0{,}03\phi_1 + \phi_2 + 0{,}03\phi_3 \\ 0{,}13 &=& 0{,}08\phi_1 + 0{,}03\phi_2 + \phi_3 \end{array}$$

che può essere risolto sfruttando la nota regola di Cramer, ovvero calcolando il rapporto tra i determinanti delle seguenti matrici

$$\hat{\phi}_1 = \frac{\begin{vmatrix} 0{,}03 & 0{,}03 & 0{,}08 \\ 0{,}08 & 1 & 0{,}03 \\ 0{,}13 & 0{,}03 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}03 & 0{,}08 \\ 0{,}03 & 1 & 0{,}03 \\ 0{,}08 & 0{,}03 & 1 \end{vmatrix}} \qquad \hat{\phi}_2 = \frac{\begin{vmatrix} 1 & 0{,}03 & 0{,}08 \\ 0{,}03 & 0{,}08 & 0{,}03 \\ 0{,}08 & 0{,}13 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}03 & 0{,}08 \\ 0{,}03 & 1 & 0{,}03 \\ 0{,}08 & 0{,}03 & 1 \end{vmatrix}}$$

$$\hat{\phi}_3 = \frac{\begin{vmatrix} 1 & 0{,}03 & 0{,}03 \\ 0{,}03 & 1 & 0{,}08 \\ 0{,}08 & 0{,}03 & 0{,}13 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}03 & 0{,}08 \\ 0{,}03 & 1 & 0{,}03 \\ 0{,}08 & 0{,}03 & 1 \end{vmatrix}}.$$

Il determinante della matrice a denominatore, comune nei tre casi, equivale a

$$\begin{vmatrix} 1 & 0{,}03 & 0{,}08 \\ 0{,}03 & 1 & 0{,}03 \\ 0{,}08 & 0{,}03 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0{,}03 \\ 0{,}03 & 1 \end{vmatrix} - 0{,}03 \begin{vmatrix} 0{,}03 & 0{,}03 \\ 0{,}08 & 1 \end{vmatrix} + \\ +0{,}08 \begin{vmatrix} 0{,}03 & 1 \\ 0{,}08 & 0{,}03 \end{vmatrix}$$

$$\begin{aligned} &= (1 - 0{,}0009) - 0{,}03(0{,}03 - 0{,}0024) + \\ &\quad +0{,}08(0{,}0009 - 0{,}08) \\ &= 0{,}9991 - 0{,}000828 - 0{,}006328 = 0{,}991944. \end{aligned}$$

Risolvendo nella stessa maniera i determinanti della matrici a numeratore si ottengono le seguenti stime dei coefficienti autoregressivi (approssimate alla quarta cifra decimale):

$$\begin{aligned} \hat{\phi}_1 &= \frac{0{,}017482}{0{,}991944} = 0{,}0176 \\ \hat{\phi}_2 &= \frac{0{,}075072}{0{,}991944} = 0{,}0757 \\ \hat{\phi}_3 &= \frac{0{,}125302}{0{,}991944} = 0{,}1263. \end{aligned}$$

**6.33** Le prime due autocorrelazioni stimate per la serie storica $\{Y_t\}_{t=1}^n$ sono, rispettivamente, $\hat{\rho}_1 = 0{,}7$ e $\hat{\rho}_2 = 0{,}4$. Ricavare una stima dei parametri del processo

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t.$$

□ Le equazioni di *Yule-Walker* per un processo AR(2) sono date da:

$$\begin{array}{rcl} \rho_1 & = & \phi_1 + \rho_1\phi_2 \\ \rho_2 & = & \rho_1\phi_1 + \phi_2 \end{array}.$$

Inserendo le autocorrelazioni stimate in tale sistema e risolvendo per $\phi_1$ e $\phi_2$ si ottengono le stime (di prima approssimazione) $\hat{\phi}_1$ e $\hat{\phi}_2$:

$$\begin{array}{rcl} 0{,}7 & = & \phi_1 + 0{,}7\phi_2 \\ 0{,}4 & = & 0{,}7\phi_1 + \phi_2 \end{array}.$$

Applicando come nel precedente esercizio la regola di Cramer abbiamo:

$$\hat{\phi}_1 = \frac{\begin{vmatrix} 0{,}7 & 0{,}7 \\ 0{,}4 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}7 \\ 0{,}7 & 1 \end{vmatrix}} \qquad \hat{\phi}_2 = \frac{\begin{vmatrix} 1 & 0{,}7 \\ 0{,}7 & 0{,}4 \end{vmatrix}}{\begin{vmatrix} 1 & 0{,}7 \\ 0{,}7 & 1 \end{vmatrix}}.$$

Risolvendo i determinanti otteniamo le stime richieste:

$$\hat{\phi}_1 = \frac{0{,}42}{0{,}51} = 0{,}8235$$
$$\hat{\phi}_2 = \frac{-0{,}09}{0{,}51} = -0{,}1765.$$

## Esercizi supplementari

**C6.1** Descrivere gli obiettivi e gli strumenti statistici usati per il controllo diagnostico nella procedura di Box e Jenkins.

**C6.2** Si verifichi la stazionarietà del processo

$$Y_t = 0{,}5 + 0{,}45Y_{t-1} + 0{,}4Y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim WN(0{,}9).$$

**C6.3** Dire se il processo ARMA(2,1)

$$Y_t = -0{,}2Y_{t-1} + 0{,}15Y_{t-2} + \varepsilon_t - 0{,}3\varepsilon_{t-1}, \quad \varepsilon_t \sim WN(0{,}1),$$

è stazionario ed invertibile. Giustificare la risposta.

**C6.4** Si consideri il processo

$$Y_t = \varepsilon_t - 0{,}5\varepsilon_{t-1} - 0{,}2\varepsilon_{t-2},\ \varepsilon_t \sim WN(0{,}4).$$

(a) Verificare se il processo è stazionario ed invertibile.

(b) Calcolare la funzione di autocorrelazione globale del processo.

**C6.5** Dire se il modello ARMA(2,2)

$$Y_t = 0{,}75Y_{t-1} - 0{,}125Y_{t-2} + \varepsilon_t - 0{,}8\varepsilon_{t-1} + 0{,}15\varepsilon_{t-2}$$

è sovraparametrizzato e, in caso affermativo, lo si semplifichi.

**C6.6** Calcolare la funzione di autocorrelazione del processo vincolato

$$Y_t = \varepsilon_t - 0{,}5\varepsilon_{t-3}, \quad \varepsilon_t \sim WN(0{,}1).$$

**C6.7** Calcolare la funzione di autocorrelazione del processo vincolato

$$Y_t = 0{,}5Y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim WN(0{,}1).$$

**C6.8** Sia $X_t \sim SARIMA(1{,}0{,}0)(1{,}0{,}0)_4$ con $\mid \phi_1 \mid < 1$ e $\mid \Phi_1 \mid < 1$.
Si trovi la rappresentazione $AR(p)$ di $X_t$.

# 7
# Previsioni ARIMA

**7.1** Per una serie storica di 60 osservazioni è stato stimato un modello ARIMA(1,1,2). Esplicitare il previsore a $k$ passi, per $k = 1,2,3$.

□ Un modello ARIMA(1,1,2) si esplicita nella seguente forma compatta:

$$(1 - \phi_1 B)(1 - B)Y_t = (1 - \theta_1 B - \theta_2 B^2)\varepsilon_t.$$

Moltiplicando i due polinomi AR otteniamo un polinomio AR generalizzato con radici stazionarie e non stazionarie

$$\begin{aligned}(1 - \phi_1 B)(1 - B) &= 1 - \phi_1 B - B + \phi_1 B^2 \\ &= 1 - (1 + \phi_1)B + \phi_1 B^2.\end{aligned}$$

Pertanto, il modello per $Y_t$ può essere riscritto nella forma esplicita

$$Y_t = (1 + \phi_1)Y_{t-1} - \phi_1 Y_{t-2} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2}.$$

Il previsore per il periodo $n + k$ al tempo $n$ è dunque pari a:

$$\begin{aligned}\hat{Y}_{n+k|n} = (1 + \phi_1)E_n[Y_{n+k-1}] - \phi_1 E_n[Y_{n+k-2}] \\ +E_n[\varepsilon_{n+k}] - \theta_1 E_n[\varepsilon_{n+k-1}] - \theta_2 E_n[\varepsilon_{n+k-2}].\end{aligned}$$

Pertanto, con $n = 60$ e $k = 1,2,3$, abbiamo

$$\begin{aligned}\hat{y}_{61|60} &= (1 + \phi_1)y_{60} - \phi_1 y_{59} - \theta_1 e_{60} - \theta_2 e_{59} \\ \hat{y}_{62|60} &= (1 + \phi_1)\hat{y}_{61|60} - \phi_1 y_{60} - \theta_2 e_{60} \\ \hat{y}_{63|60} &= (1 + \phi_1)\hat{y}_{62|60} - \phi_1 \hat{y}_{61|60},\end{aligned}$$

dove $e_{59}$ e $e_{60}$ sono i disturbi stimati dal modello nei due periodi.

**7.2** Una serie storica a frequenza mensile è stata generata da un processo SARIMA$(0,1,1)\times(0,1,0)_{12}$. Esplicitare il previsore a $k$ passi, per $k = 1,2$.

□ Il processo indicato nell'esercizio ha la seguente espressione:

$$(1 - B)(1 - B^{12})Y_t = (1 - \theta_1 B)\varepsilon_t.$$

Poiché

$$(1 - B)(1 - B^{12}) = 1 - B - B^{12} + B^{13},$$

il processo può essere riscritto nel modo seguente:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \varepsilon_t - \theta_1 \varepsilon_{t-1}.$$

Il previsore è dunque pari a

$$\begin{aligned}\hat{Y}_{n+k|n} = E_n[Y_{n+k-1}] + E_n[Y_{n+k-12}] - E_n[Y_{n+k-13}] + \\ + E_n[\varepsilon_{n+k}] - \theta_1 E_n[\varepsilon_{n+k-1}].\end{aligned}$$

Per $k = 1{,}2$ l'espressione precedente dà origine alle previsioni

$$\begin{aligned}\hat{y}_{n+1|n} &= y_n + y_{n-11} - y_{n-12} - \theta_1 e_n \\ \hat{y}_{n+2|n} &= \hat{y}_{n+1|n} + y_{n-10} - y_{n-11},\end{aligned}$$

con $e_n$ disturbo stimato al tempo $n$.

**7.3** Derivare il previsore a $k$ passi, per $k = 1{,}2{,}3$, per un processo SARIMA(1,0,1)$\times$(1,0,0)$_4$.

□ Il processo ha la seguente forma compatta:

$$(1 - \phi_1 B)(1 - \Phi_1 B^4) Y_t = (1 - \theta_1 B)\varepsilon_t.$$

Il prodotto dei due polinomi autoregressivi dà luogo al seguente polinomio AR generalizzato

$$(1 - \phi_1 B)(1 - \Phi_1 B^4) = 1 - \phi_1 B - \Phi_1 B^4 + \phi_1 \Phi_1 B^5.$$

Il processo può essere quindi esplicitato nella seguente forma:

$$Y_t = \phi_1 Y_{t-1} + \Phi_1 Y_{t-4} - \phi_1 \Phi_1 Y_{t-5} + \varepsilon_t - \theta_1 \varepsilon_{t-1}.$$

Il previsore a $k$ passi ha quindi la forma

$$\begin{aligned}\hat{Y}_{n+k|n} = \phi_1 E_n[Y_{n+k-1}] + \Phi_1 E_n[Y_{n+k-4}] - \phi_1 \Phi_1 E_n[Y_{n+k-5}] + \\ + E[\varepsilon_{n+k}] - \theta_1 E[\varepsilon_{n+k-1}].\end{aligned}$$

Pertanto per $k = 1{,}2{,}3$ si ha

$$\begin{aligned}\hat{y}_{n+1|n} &= \phi_1 y_n + \Phi_1 y_{n-3} - \phi_1 \Phi_1 y_{n-4} - \theta_1 e_n \\ \hat{y}_{n+2|n} &= \phi_1 \hat{y}_{n+1|n} + \Phi_1 y_{n-2} - \phi_1 \Phi_1 y_{n-3} \\ \hat{y}_{n+3|n} &= \phi_1 \hat{y}_{n+2|n} + \Phi_1 y_{n-1} - \phi_1 \Phi_1 y_{n-2}.\end{aligned}$$

**7.4** Derivare il previsore a $k$ passi, per $k = 1{,}2{,}3$, per un processo SARIMA(0,1,1)$\times$(0,1,1)$_{12}$.

□ Il processo indicato nell'esercizio ́dato da:

$$(1 - B)(1 - B^{12}) Y_t = (1 - \theta_1 B)(1 - \Theta_1 B^{12})\varepsilon_t.$$

Tale processo è spesso indicato con il termine *airline*, perché utilizzato, in forma logaritmica, da Box e Jenkins (1970) per modellare una serie storica di passeggeri su linee aeree internazionali. Il processo è molto utilizzato perché ben si adatta a numerose serie storiche osservate che presentano trend e stagionalità piuttosto regolari nel tempo. Come nei precedenti esercizi, esplicitiamo il polinomio AR generalizzato come prodotto dei due polinomi non stazionari

$$(1-B)(1-B^{12}) = 1 - B - B^{12} + B^{13}.$$

Allo stesso modo possiamo derivare il polinomio MA generalizzato:

$$(1-\theta_1 B)(1-\Theta_1 B^{12}) = 1 - \theta_1 B - \Theta_1 B^{12} + \theta_1\Theta_1 B^{13}.$$

Il processo può quindi essere riscritto nel modo seguente:

$$Y_t = Y_{t-1} + Y_{t-12} - Y_{t-13} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \Theta_1\varepsilon_{t-12} + \theta_1\Theta_1\varepsilon_{t-13}.$$

Il previsore è dunque pari a:

$$\begin{aligned}\hat{Y}_{n+k|n} = {} & E_n[Y_{n+k-1}] + E_n[Y_{n+k-12}] - E_n[Y_{n+k-13}] + \\ & + E_n[\varepsilon_{n+k}] - \theta_1 E_n[\varepsilon_{n+k-1}] - \Theta_1 E_n[\varepsilon_{n+k-12}] + \\ & + \theta_1\Theta_1 E_n[\varepsilon_{n+k-13}].\end{aligned}$$

Pertanto, per $k = 1,2,3$ si ha:

$$\begin{aligned}\hat{y}_{n+1|n} &= y_n + y_{n-11} - y_{n-12} - \theta_1 e_n - \Theta_1 e_{n-11} + \theta_1\Theta_1 e_{n-12} \\ \hat{y}_{n+2|n} &= \hat{y}_{n+1|n} + y_{n-10} - y_{n-11} - \Theta_1 e_{n-10} + \theta_1\Theta_1 e_{n-11} \\ \hat{y}_{n+3|n} &= \hat{y}_{n+2|n} + y_{n-9} - y_{n-10} - \Theta_1 e_{n-9} + \theta_1\Theta_1 e_{n-10},\end{aligned}$$

dove $e_{n-j}$ sono i disturbi stimati al tempo $n-j$, con $j \geq 0$.

**7.5** Dato un modello SARIMA(1,0,0)$\times$(0,1,1)$_4$, calcolare la previsione con orizzonte temporale $k = 5$ sapendo che:

- il parametro AR $\phi_1$ è pari a -0,8;
- l'ultimo valore osservato della serie storica è $y_n = 3$;
- $\hat{y}_{n+1|n} = -2$ e $\hat{y}_{n+4|n} = 1$.

□ Il processo indicato si esprime nella forma:

$$(1-\phi_1 B)(1-B^4)Y_t = (1-\Theta_1 B^4)\varepsilon_t.$$

Poiché

$$(1-\phi_1 B)(1-B^4) = 1 - \phi_1 B - B^4 + \phi_1 B^5,$$

il processo può essere riscritto nel modo seguente:

$$Y_t = \phi_1 Y_{t-1} + Y_{t-4} - \phi_1 Y_{t-5} + \varepsilon_t - \Theta_1\varepsilon_{t-4}.$$

Il previsore di $Y_{n+5}$ è dunque pari a:

$$\begin{aligned}\hat{Y}_{n+5|n} &= E_n[Y_{n+5}] \\ &= \phi_1 E_n[Y_{n+4}] + E_n[Y_{n+1}] - \phi_1 E_n[Y_n] + \\ &\quad + E_n[\varepsilon_{n+5}] - \Theta_1 E_n[\varepsilon_{n+1}].\end{aligned}$$

Il calcolo dei valori attesi condizionati avviene secondo le seguenti espressioni:

$$E_n[Y_{n+j}] = \begin{cases} y_{n+j} & j \leq 0 \\ \hat{y}_{n+j|n} & j > 0 \end{cases}$$
$$E_n[\varepsilon_{n+j}] = \begin{cases} e_{n+j} & j \leq 0 \\ 0 & j > 0 \end{cases}.$$

Pertanto, la previsione $\hat{y}_{n+5|n}$ è pari a

$$\hat{y}_{n+5|n} = \phi_1 \hat{y}_{n+4|n} + \hat{y}_{n+1|n} - \phi_1 y_n.$$

Sostituendo i valori noti nella precedente espressione otteniamo il risultato richiesto:

$$\begin{aligned}\hat{y}_{n+5|n} &= -0{,}8 - 2 - (-0{,}8) \cdot 3 \\ &= -0{,}8 - 2 + 2{,}4 = -0{,}4.\end{aligned}$$

**7.6** Dato un modello SARIMA(0,0,1)×(1,1,0)$_4$, calcolare la previsione con orizzonte temporale $k = 8$ sapendo che

- il parametro AR stagionale $\Phi_1$ è pari a 0,3;
- l'ultimo valore osservato della serie storica è $y_n = 4$;
- $\hat{y}_{n+4|n} = 2$.

□ Il processo ha la forma compatta

$$(1 - \Phi_1 B^4)(1 - B^4) Y_t = (1 - \theta_1 B)\varepsilon_t.$$

Il polinomio AR generalizzato è pari a

$$(1 - \Phi_1 B^4)(1 - B^4) = 1 - \Phi_1 B^4 - B^4 + \Phi_1 B^8.$$

Sfruttando questa espressione, il processo originario si trasforma in

$$Y_t = (1 + \Phi_1) Y_{t-4} - \Phi_1 Y_{t-8} + \varepsilon_t - \theta_1 \varepsilon_{t-1}.$$

Il previsore di $Y_{n+8}$ è dunque pari a:

$$\hat{Y}_{n+8|n} = (1 + \Phi_1) E_n[Y_{n+4}] - \Phi_1 E_n[Y_n] + E_n[\varepsilon_{n+8}] - \theta_1 E_n[\varepsilon_{n+7}].$$

Pertanto, utilizzando le espressioni indicate nell'esercizio precedente, la previsione di $Y_{n+8}$ secondo questo processo è

$$\begin{aligned}\hat{y}_{n+8|n} &= (1 + \Phi_1) \hat{y}_{n+4|n} - \Phi_1 y_n \\ &= 1{,}3 \cdot 2 - 0{,}3 \cdot 4 = 2{,}6 - 1{,}2 = 1{,}4.\end{aligned}$$

**7.7** Sia dato il processo

$$Y_t = 2 + 0{,}8Y_{t-1} - 0{,}1Y_{t-2} + \varepsilon_t$$

con $\varepsilon_t \sim WN(0, \sigma^2_\varepsilon)$.

1. Dire se $Y_t$ è stazionario.
2. Date quattro osservazioni da tale processo:

$$y_1 = 4, y_2 = 3, y_3 = 1, y_4 = 2{,}5$$

si calcolino le previsioni per $Y_5$, $Y_6$ e $Y_7$.

□ $Y_t$ ha la forma di un processo $AR(2)$.

1. Le soluzioni dell'equazione caratteristica

$$1 - 0{,}8B + 0{,}1B^2 = 0$$

sono pari a (verificare) $B_1 = 1{,}55051$ e $B_2 = 6{,}44949$. Essendo entrambe le radici maggiori di 1, il processo è stazionario.

2. Applicando la regola del valore atteso condizionato per il calcolo delle previsioni, abbiamo:

$$\begin{aligned}\hat{Y}_{5|4} &= E_4[Y_5] = 2 + 0{,}8E_4[Y_4] - 0{,}1E_4[Y_3] + E_4[\varepsilon_5] \\ &= 2 + 0{,}8y_4 - 0{,}1y_3 \\ &= 2 + 0{,}8 \cdot 2{,}5 - 0{,}1 \cdot 1 = \mathbf{3{,}9}\end{aligned}$$

$$\begin{aligned}\hat{Y}_{6|4} &= E_4[Y_6] = 2 + 0{,}8E_4[Y_5] - 0{,}1E_4[Y_4] + E_4[\varepsilon_6] \\ &= 2 + 0{,}8\hat{y}_{5|4} - 0{,}1y_4 \\ &= 2 + 0{,}8 \cdot 3{,}9 - 0{,}1 \cdot 2{,}5 = \mathbf{4{,}87}\end{aligned}$$

$$\begin{aligned}\hat{Y}_{7|4} &= E_4[Y_7] = 2 + 0{,}8E_4[Y_6] - 0{,}1E_4[Y_5] + E_4[\varepsilon_7] \\ &= 2 + 0{,}8\hat{y}_{6|4} - 0{,}1\hat{y}_{5|4} \\ &= 2 + 0{,}8 \cdot 4{,}87 - 0{,}1 \cdot 3{,}9 = \mathbf{5{,}506}\end{aligned}$$

## Esercizi supplementari

**C7.1** Derivare le previsioni a $k$ passi, per $k = 1,2,3,4$, per un processo SARIMA$(0,1,1)\times(0,1,1)_3$.

**C7.2** Si consideri il processo

$$X_t = \varepsilon_t + 0{,}2\varepsilon_{t-1} \qquad \text{con} \qquad \varepsilon_t \sim IID\ N(0{,}0{,}000625)$$

per $t = 1,2,\ldots,100$. Si supponga anche che $\varepsilon_{100} = 0{,}01$.

(a) Si calcoli la previsione un passo in avanti per $X_t$ assieme all'errore standard del suo errore di previsione.

(b) Si calcoli la previsione due passi in avanti per $X_t$ assieme al suo errore di previsione.

**C7.3** Si consideri il processo:

$$Y_t = 0{,}5Y_{t-1} + \varepsilon_t - 1{,}2\varepsilon_{t-1} + 0{,}35\varepsilon_{t-2}$$

per $t = 1,2,\ldots,n$, dove $\varepsilon_t \sim WN(0,\sigma^2_\varepsilon)$.

(a) Dire, giustificando la risposta, se il processo è stazionario e invertibile.

(b) Dire a quale classe dei processi ARIMA appartiene.

(c) Calcolare i coefficienti di autocorrelazione globale, $\rho_k$, ai ritardi $k = 1,5,10$ e 100.

(d) Calcolare le previsioni 1, 5, 10 e 100 passi in avanti.

**C7.4** Si consideri il processo:

$$X_t = \begin{cases} \varepsilon_t & \text{per } t \text{ pari} \\ \frac{\varepsilon^2_{t-1}-1}{\sqrt{2}} & \text{per } t \text{ dispari} \end{cases} \tag{7.1}$$

dove $\epsilon_t \sim IID\mathcal{N}(0,1)$.

(a) Si dimostri che il processo $X_t \sim WN(0,1)$.

(b) Si trovi la previsione di $X_t$ un passo in avanti, ossia

$$\hat{X}_{n+1} = E[X_{n+1} \mid X_n, \cdots, X_1].$$

(Si ricordi che se $Z \sim \mathcal{N}(0,1)$ allora $E[Z^3] = 0$ e $E[Z^4] = 3$)

# 8
# Soluzioni agli esercizi supplementari

### Capitolo 1 - Concetti introduttivi

**C1.1** Vedi Di Fonzo e Lisi (2005, pp. 34-37).

### Capitolo 2 - Funzioni deterministiche del tempo

**C2.1** Vedi Di Fonzo e Lisi (2005, pp. 42-44).

**C2.2** $y_t = \alpha t + \beta t^2 + \gamma_1 d_{1t} + \gamma_2 d_{2t} + \varepsilon_t$, $t = 1, \ldots, n$, con $d_{1t}$ e $d_{2t}$ variabili *dummy* stagionali riferite al primo e secondo semestre e $\varepsilon_t$ un processo *white-noise*.

**C2.3** (a) No, sono necessari i valori osservati $y_t$.

(b) $E(y_t) = 15{,}5$, $Var(y_t) = 16$.

(c) No, è necessario l'operatore $(1 - B)^2$ per rimuovere un trend quadratico.

**C2.4**

$$f(t) = T_t + S_t = \begin{array}{r} \alpha + \beta t + \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t} + \gamma_4 d_{4t} \\ = \beta t + \gamma_1^* d_{1t} + \gamma_2^* d_{2t} + \gamma_3^* d_{3t} + \gamma_4^* d_{4t} \end{array}, \qquad \mathbf{X}^* = \begin{bmatrix} 2 & 0 & 1 & 0 & 0 \\ 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & 0 & 0 & 1 \\ 5 & 1 & 0 & 0 & 0 \\ 6 & 0 & 1 & 0 & 0 \\ 7 & 0 & 0 & 1 & 0 \\ 8 & 0 & 0 & 0 & 1 \\ 9 & 1 & 0 & 0 & 0 \\ 10 & 0 & 1 & 0 & 0 \end{bmatrix};$$

$$\bar{\gamma}^* = \sum_{j=1}^{4} \gamma_j^* \rightarrow \hat{\gamma}_j = \gamma_j^* - \bar{\gamma}^*, j = 1{,}2{,}3{,}4 \Rightarrow f(t) = \bar{\gamma}^* + \beta t + \hat{\gamma}_1 d_{1t} + \hat{\gamma}_2 d_{2t} + \hat{\gamma}_3 d_{3t} + \hat{\gamma}_4 d_{4t}.$$

E pertanto $y_t^d = y_t - (\hat{\gamma}_1 d_{1t} + \hat{\gamma}_2 d_{2t} + \hat{\gamma}_3 d_{3t} + \hat{\gamma}_4 d_{4t})$.

Volendo usare un modello con intercetta e privo della variabile *dummy* relativa al quarto trimestre, si avrebbe:

$$\tilde{\alpha} + \beta t + \tilde{\gamma}_1 d_{1t} + \tilde{\gamma}_2 d_{2t} + \tilde{\gamma}_3 d_{3t}, \quad \tilde{\mathbf{X}} = \begin{bmatrix} 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 1 \\ 4 & 1 & 0 & 0 & 0 \\ 5 & 1 & 1 & 0 & 0 \\ 6 & 1 & 0 & 1 & 0 \\ 7 & 1 & 0 & 0 & 1 \\ 8 & 1 & 0 & 0 & 0 \\ 9 & 1 & 1 & 0 & 0 \\ 10 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Le relazioni tra i coefficienti utili in fase di destagionalizzazione sono le seguenti:

$$\begin{array}{ll} \gamma_1^* = \tilde{\alpha} + \tilde{\gamma}_1 & \tilde{\alpha} = \gamma_4^* \\ \gamma_2^* = \tilde{\alpha} + \tilde{\gamma}_2 & \tilde{\gamma}_1 = \gamma_1^* - \gamma_4^* \\ \gamma_3^* = \tilde{\alpha} + \tilde{\gamma}_3, & \tilde{\gamma}_2 = \gamma_2^* - \gamma_4^* \\ \gamma_4^* = \tilde{\alpha} & \tilde{\gamma}_3 = \gamma_3^* - \gamma_4^* \end{array} \quad \Rightarrow \quad \bar{\gamma}^* = \frac{1}{4}\sum_{j=1}^{4} \gamma_j^* = \tilde{\alpha} + \frac{\tilde{\gamma}_1 + \tilde{\gamma}_2 + \tilde{\gamma}_3}{4}$$

$$\begin{aligned} \hat{\gamma}_1 &= \gamma_1^* - \bar{\gamma}^* = \frac{3\tilde{\gamma}_1 - \tilde{\gamma}_2 - \tilde{\gamma}_3}{4} \\ \hat{\gamma}_2 &= \gamma_2^* - \bar{\gamma}^* = \frac{3\tilde{\gamma}_2 - \tilde{\gamma}_1 - \tilde{\gamma}_3}{4} \\ \hat{\gamma}_2 &= \gamma_2^* - \bar{\gamma}^* = \frac{3\tilde{\gamma}_3 - \tilde{\gamma}_1 - \tilde{\gamma}_2}{4} \\ \hat{\gamma}_4 &= \gamma_4^* - \bar{\gamma}^* = -\frac{\tilde{\gamma}_1 + \tilde{\gamma}_2 + \tilde{\gamma}_3}{4} \end{aligned}.$$

**C2.5** Supponendo di aver caricato i dati della serie storica dei ricavi nella variabile `y` di Gretl, lo script

```
genr alfa = 0.1
genr beta = 0.1
y = alfa*exp(beta*time)
deriv alfa = exp(beta*time)
deriv beta = alfa*exp(beta*time)*time
```

produce il seguente risultato:

```
Sono state usate derivate analitiche
Tolleranza = 1,81899e-012
Convergenza raggiunta dopo 31 iterazioni
```

Modello 1: NLS, usando le osservazioni 2003–2009 ($T$ = 7)
`y = alfa*exp(beta*time)`

| | Stima | Errore Std. | rapporto $t$ | p-value |
|---|---|---|---|---|
| alfa | 202,597 | 39,8034 | 5,0899 | 0,0038 |
| $\beta$ | 0,225105 | 0,0339881 | 6,6231 | 0,0012 |

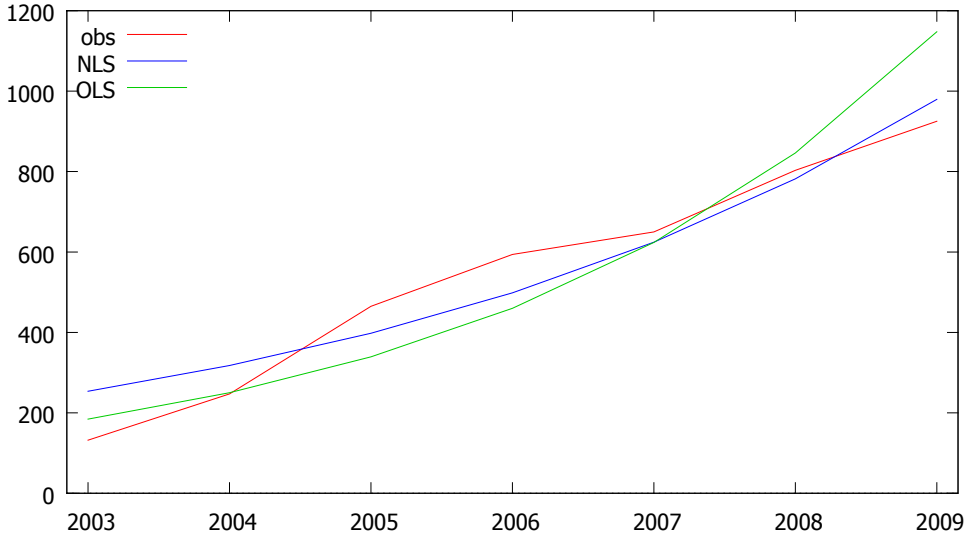
| | | | |
|---|---|---|---|
| Media var. dipendente | 545,1429 | SQM var. dipendente | 285,8213 |
| Somma quadr. residui | 37494,83 | E.S. della regressione | 86,59657 |
| $R^2$ | 0,923505 | $R^2$ corretto | 0,908206 |
| Log-verosimiglianza | −39,98374 | Criterio di Akaike | 83,96748 |
| Criterio di Schwarz | 83,85930 | Hannan–Quinn | 82,63040 |
| $\hat{\rho}$ | 0,350817 | Durbin–Watson | 0,879515 |

Avremmo ottenuto lo stesso risultato, ma in 10 iterazioni anziché 31, se avessimo usato come valori iniziali le stime ottenute coi minimi quadrati sul modello trasformato mediante logaritmo ($\hat{\alpha} = 135{,}8796$ e $\hat{\beta} = 0{,}3048$).

I valori delle 3 serie, osservata (obs) e stimate con i minimi quadrati non lineari (NLS) e coi minimi quadrati sul modello trasformato (OLS), nonché il loro grafico, sono riportati di seguito.

| Anno | obs | NLS | OLS |
|---|---|---|---|
| 2003 | 132 | 253,74 | 184,30 |
| 2004 | 247 | 317,80 | 249,98 |
| 2005 | 465 | 398,03 | 339,06 |
| 2006 | 594 | 498,52 | 459,88 |
| 2007 | 650 | 624,37 | 623,76 |
| 2008 | 803 | 781,99 | 846,04 |
| 2009 | 925 | 979,41 | 1147,53 |

Figura 8.1
Grafico della serie osservata e dei valori stimati




La radice quadrata dell'errore quadratico medio, $\sqrt{\frac{1}{7}\sum_{t=1}^{7}(y_t - \hat{y}_t)^2}$, è pari a 73,19 per la serie stimata coi minimi quadrati non lineari, a fronte di un valore dello stesso indice pari a 112,54 per la serie stimata a partire dal modello trasformato. Trova quindi conferma quanto già appariva visivamente, ovvero la migliore qualità dell'adattamento ai dati osservati delle stime ottenute coi minimi quadrati non lineari.

## Capitolo 3 - Le medie mobili

**C3.1** Vedi Di Fonzo e Lisi (2005, pp. 86-87).

**C3.2** Vedi Di Fonzo e Lisi (2005, p. 90).

**C3.3** Vedi Di Fonzo e Lisi (2005, pp. 116-119).

**C3.4** Vedi Di Fonzo e Lisi (2005, pp. 119-123).

**C3.5** $\sum_{i=-3}^{3} \theta_i = 1$ e $\sum_{i=-3}^{3} i\theta_i = \sum_{i=-3}^{3} i^2\theta_i = 0$.

**C3.6** $M = \left\{ [7]; \frac{1}{24}[1,3,5,\mathbf{6}] \right\}$.

**C3.7** $M = \left\{ [7]; \frac{1}{12}[-3,0,4,\mathbf{10}] \right\}$.

**C3.8** $M = \left\{ [7]; \frac{1}{39}[1,4,9,\mathbf{11}] \right\}$.

**C3.9** $M_0 = M_1 M_2$.

**C3.10** $M = \left\{ [7]; \frac{1}{21}[-2,3,6,\mathbf{7}] \right\}$. Il rapporto di riduzione della varianza residua è pari a 0,3333, superiore a quello di una media mobile aritmetica semplice di ordine 7 pari a 0,1429.

**C3.11** La media mobile $M_1$ ha un rapporto di riduzione pari a 0,3333, mentre per $M_2$ la riduzione è 0,3750. L'applicazione di $M_1$ fornisce quindi una serie storica più liscia.

**C3.12** La serie destagionalizzata è:

| *Anno* | *I trim* | *II trim* | *III trim* | *IV trim* |
|---|---|---|---|---|
| 1 | 103,62 | 99,90 | 97,08 | 96,88 |
| 2 | 98,12 | 97,70 | 97,78 | 97,48 |
| 3 | 97,32 | 95,00 | 97,38 | 97,48 |
| 4 | 98,02 | 97,80 | 97,38 | 96,18 |
| 5 | 94,12 | 96,70 | 96,68 | 94,58 |

## Capitolo 4 - Il lisciamento esponenziale

**C4.1** Interpolante lineare: $\hat{y}_{2001} = 103.307$ , $\hat{y}_{2002} = 101.736{,}4$;
Lisciamento esponenziale semplice: $\hat{y}_{2001} = 131.015{,}6$ ,
$\hat{y}_{2002} = \hat{y}_{2001} = 131.015{,}6$

**C4.2** $y^*_{10} = 0{,}00142y_6 + 0{,}004725y_7 + 0{,}01575y_8 + 0{,}3025y_9 + 0{,}675y_{10}$

## Capitolo 5 - Processi stocastici

**C5.1** a) Processo MA(1): $z_t = \xi_t - \theta_1\xi_{t-1}$, con $\xi_t = \varepsilon_t + \nu_t$;
b) $E(z_t) = 0$, $Var(z_t) = \sigma^2_\xi(1+\theta_1^2)$, $E(z_t z_{t-1}) = -\sigma^2_\xi\theta_1$.

**C5.2** La funzione di autocovarianza data è diversa da zero solo per il primo ritardo e poi è sempre nulla.

(a) Questo è coerente con un modello a media mobile di ordine 1, MA(1), o ARIMA(0,0,1).

(b) Un modello MA(1) può scriversi come:

$$X_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

con $\varepsilon \sim WN(0, \sigma^2)$. Pertanto dobbiamo trovare i valori dei due parametri $\theta$ e $\sigma^2$.
Sappiamo che, per un processo MA(1), si ha:

$$\gamma_0 = Var(X_t) = (1 + \theta^2)\sigma^2 = 2$$

e

$$\gamma_1 = Cov(X_t,X_{t-1}) = -\theta\sigma^2 = 0{,}5$$

In pratica ci ritroviamo con un sistema di due equazioni in due incognite. Per trovare $\theta$ possiamo osservare che:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = -\frac{\theta\sigma^2}{(1+\theta^2)\sigma^2} = -\frac{\theta}{1+\theta^2} = \frac{0{,}5}{2} = 0{,}25$$

da cui otteniamo l'equazione di secondo grado:

$$0{,}25\theta^2 + \theta + 0{,}25 = 0$$

che risolta porta alle due soluzioni:

$$\theta_1 \approx -3{,}73 \qquad e \qquad \theta_2 \approx -0{,}27.$$

Di queste due soluzioni si considera solo $\theta_2$, perché $\theta_1$ è in modulo maggiore di 1 e porta ad un processo MA(1) non invertibile.
Sostituendo $\theta_2$ in una delle due equazioni precedenti, si trova il valore di $\sigma^2 \approx 1{,}85$.
In sintesi il modello MA(1) stazionario e invertibile cui corrisponde la funzione di autocovarianza data è:

$$X_t = \varepsilon_t + 0{,}27\varepsilon_{t-1}$$

con $\varepsilon \sim WN(0, 1{,}85)$.

**C5.3** Vale quanto segue:

(a) Il primo processo è un WN, quindi gli corrispondono le funzioni di autocorrelazione della serie 3, infatti tutti i coefficienti di autocorrelazione globale e parziale risultano interni alle bande di confidenza e quindi non significativamente diversi da zero.

(b) Il processo 2 è un processo AR(1), pertanto ci si aspetta che la funzione di autocorrelazione globale decada a zero velocemente e che vi sia un unico ritardo significativo, il primo, nella funzione di autocorrelazione parziale. Questo comportamento lo si ritrova nelle autocorrelazioni empiriche della serie 4.

(c) Il processo 3 è un processo MA(1), a cui sono associati i correlogrammi empirici della serie 1. Infatti, per tale serie solo il primo coefficiente della funzione di autocorrelazione globale è significativo, mentre la funzione di autocorrelazione parziale va a zero dopo un po' di ritardi.

(d) Il processo 4 è un processo ARMA(1,1). A questo processo rimangono associate le funzioni di autocorrelazione della serie 2 che, infatti, decrescono entrambe a zero.

**C5.4** Vale quanto segue:

(a) Sostituendo $X_t$ in $Y_t$ si ottiene:

$$Y_t = \frac{5}{12}Y_{t-1} + \frac{1}{6}Y_{t-2} + \epsilon_t - \frac{7}{6}\epsilon_{t-1} + \frac{1}{3}\epsilon_{t-2}$$

che è un processo $ARMA(2,2)$.

(b) Il processo può essere riscritto in termini dell'operatore ritardo come:

$$\left(1 - \frac{5}{12}B - \frac{1}{6}B^2\right) Y_t = \left(1 - \frac{7}{6}B + \frac{1}{3}B^2\right) \epsilon_t$$

Per verificare se il processo è stazionario si deve verificare se le radici del polinomio associato alla parte AR sono in modulo maggiori di 1. Risolvendo l'equazione:

$$1 - \frac{5}{12}B - \frac{1}{6}B^2 = 0$$

si trovano le due radici $B_1 = 1{,}5$ e $B_2 = -4$. Essendo entrambe in modulo maggiori di uno, il processo è stazionario.
Per verificare se il processo è invertibile bisogna verificare se le radici del polinomio associato alla parte MA sono in modulo maggiori di 1. Risolvendo l'equazione:

$$1 - \frac{7}{6}B + \frac{1}{3}B^2 = 0$$

si trovano le due radici $B_1 = 1{,}5$ e $B_2 = 2$. Essendo entrambe in modulo maggiori di uno, il processo è invertibile.

(c) Si osservi che i due polinomi hanno una radice in comune che può essere eliminata, pertanto il processo è in realtà un processo $ARMA(1,1)$ che può scriversi come:

$$\left(1 - \frac{2}{3}B\right)\left(1 + \frac{1}{4}B\right) Y_t = \left(1 - \frac{2}{3}B\right)\left(1 - \frac{1}{2}B\right) \epsilon_t$$

ossia

$$\left(1 + \frac{1}{4}B\right) Y_t = \left(1 - \frac{1}{2}B\right) \epsilon_t$$

o anche

$$Y_t = -\frac{1}{4}Y_{t-1} + \epsilon_t - \frac{1}{2}\epsilon_{t-1}$$

pertanto la varianza è:

$$Var(Y_t) = \frac{1}{16}Var(Y_{t-1}) + Var(\epsilon_t) + \frac{1}{4}Var(\epsilon_{t-1}) + \frac{1}{4}Cov(Y_{t-1},\epsilon_{t-1})$$

essendo il processo stazionario $Var(Y_t) = Var(Y_{t-1})$, inoltre $Cov(Y_{t-1},\epsilon_{t-1}) = Var(\epsilon_t) = 1$, quindi:

$$(1 - \frac{1}{16})Var(Y_t) = 1 + \frac{1}{4} + \frac{1}{4}$$

o

$$\frac{15}{16}Var(Y_t) = \frac{3}{2}$$

da cui si ricava, infine, $Var(Y_t) = \frac{8}{5} = 1{,}6$.

**C5.5** Si ha:

(a)
$$E[W_t] = \int_0^2 \frac{1}{2} x dx = \left[\frac{1}{2}\frac{x^2}{2}\right]_0^2 = 1$$

$$Var(W_t) = E[W_t^2] - E[W_t]^2 = \int_0^2 \frac{1}{2} x^2 dx - 1 = \left[\frac{1}{2}\frac{x^3}{3}\right]_0^2 - 1 = \frac{1}{3}$$

di conseguenza:

$$E[X_t] = E[W_t] + \frac{1}{2}E[W_{t-1}] = 1 + \frac{1}{2} = \frac{3}{2}$$

e

$$Var(X_t) = Var\left(W_t + \frac{1}{2}W_{t-1}\right) = Var(W_t) + \frac{1}{4}Var(W_{t-1}) = \frac{1}{3}\left(1 + \frac{1}{4}\right) = \frac{5}{12}$$

Si osservi che sia il valore atteso, sia la varianza di $X_t$ non dipendono dal tempo e restano costanti per ogni $t$.

(b) Per calcolare la funzione di autocorrelazione devo prima calcolare la funzione di autocovarianza del processo. Si osservi che il processo è un processo $MA(1)$ con errori non gaussiani, quindi la funzione di autocovarianza del processo è nulla per $k > 1$. Per $k = 1$ abbiamo:

$$\begin{aligned} Cov(X_t, X_{t-1}) &= Cov\left[\left(W_t + \frac{1}{2}W_{t-1}\right)\left(W_{t-1} + \frac{1}{2}W_{t-2}\right)\right] \\ &= Cov\left(\frac{1}{2}W_{t-1}, W_{t-1}\right) = E\left[\frac{1}{2}W_{t-1}^2\right] - \frac{1}{2}\cdot 1 \\ &= \frac{1}{2}\left[E[W_{t-1}^2] - 1\right] = \frac{1}{2}Var(W_{t-1}) = \frac{1}{6} \end{aligned}$$

Quindi

$$\rho(1) = \frac{1/6}{5/12} = 0{,}4$$

mentre $\rho(k) = 0$ per $k > 1$.

(c) Il processo è stazionario, perché valore atteso, varianza e autocovarianze del processo non dipendono dal tempo.

## Capitolo 6 - Processi ARIMA

**C6.1** Vedi Di Fonzo e Lisi (2005, pp. 205-208).

**C6.2** $B_1 = -2{,}241, B_2 = 1{,}116$; il processo è stazionario.

**C6.3** Radici AR: $B_1 = -2, B_2 = 3{,}333$. Radice MA: $B_1 = 3{,}333$. Il processo è stazionario ed invertibile.

**C6.4** a) Radici MA: $B_1 = -3{,}812, B_2 = 1{,}312$. Il processo è stazionario ed invertibile. b) Funzione di autocorrelazione globale: $\rho_1 = -0{,}31008$, $\rho_2 = -0{,}15504$, $\rho_k = 0$ per $k > 2$.

**C6.5** Sì, esiste una radice comune tra i polinomi AR e MA ($B = 2$). Il processo si semplifica nell'ARMA(1,1)

$$(1 - 0{,}25B)y_t = (1 - 0{,}3B)\varepsilon_t.$$

**C6.6** La funzione di autocorrelazione è pari a $\rho_1 = \rho_2 = 0$, $\rho_3 = 0{,}4$, $\rho_k = 0$ per $k > 3$.

**C6.7** La funzione di autocorrelazione è pari a $\rho_1 = 0$, $\rho_2 = 0{,}5$, $\rho_3 = 0$, $\rho_k = 0{,}5\rho_{k-2}$ per $k > 3$.

**C6.8** Possiamo scrivere $X_t$ come:

$$(1 - \phi_1 B)(1 - \Phi_1 B^4)X_t = \varepsilon_t$$

con $\varepsilon_t \sim WN(0, \sigma^2$. Sviluppando il prodotto si ottiene:

$$(1 - \phi_1 B - \Phi_1 B^4 + \phi_1\Phi_1 B^5)Y_t = \varepsilon_t$$

da cui:

$$Y_t = \phi_1 Y_{t-1} + \Phi_1 Y_{t-4} - \phi_1\Phi_1 Y_{t-5}$$

che è un processo $AR(5)$ con i coefficienti relativi al secondo e terzo ritardo pari a zero.

## Capitolo 7 - Previsioni ARIMA

**C7.1** Le previsioni a $k$ passi, per $k = 1{,}2{,}3{,}4$, sono date da:

$$\begin{aligned}
\hat{y}_{n+1|n} &= y_n + y_{n-2} - y_{n-3} - \theta_1 e_n - \Theta_1 e_{n-2} + \theta_1\Theta_1 e_{n-3} \\
\hat{y}_{n+2|n} &= \hat{y}_{n+1|n} + y_{n-1} - y_{n-2} - \Theta_1 e_{n-1} + \theta_1\Theta_1 e_{n-2} \\
&= y_n + y_{n-1} - y_{n-3} - \theta_1 e_n - \Theta_1 e_{n-1} + \Theta_1(\theta_1 - 1)e_{n-2} + \theta_1\Theta_1 e_{n-3} \\
\hat{y}_{n+3|n} &= \hat{y}_{n+2|n} + y_n - y_{n-1} - \Theta_1 e_n + \theta_1\Theta_1 e_{n-1} \\
&= 2y_n - y_{n-3} - (\theta_1 + \Theta_1)e_n + \Theta_1(\theta_1 - 1)e_{n-1} + \Theta_1(\theta_1 - 1)e_{n-2} + \\
&\quad + \theta_1\Theta_1 e_{n-3} \\
\hat{y}_{n+4|n} &= \hat{y}_{n+3|n} + \hat{y}_{n+1|n} - y_n + \theta_1\Theta_1 e_n \\
&= 3y_n + y_{n-2} - 2y_{n-3} - (2\theta_1 + \Theta_1)e_n + \Theta_1(\theta_1 - 1)e_{n-1} + \\
&\quad + \Theta_1(\theta_1 - 2)e_{n-2} + 2\theta_1\Theta_1 e_{n-3}.
\end{aligned}$$

**C7.2** Si ha:

(a)

$$\hat{Y}_{101} = E[Y_{101} \mid I_{100}] = E[\varepsilon_{101} + 0{,}2\varepsilon_{100}]$$

ed essendo, per la regola dei valori attesi condizionati, $E[\varepsilon_{101} \mid I_{100}] = 0$ risulta:

$$\hat{Y}_{101} = 0{,}2 * 0{,}01 = 0{,}002.$$

L'errore di previsione è

$$e_{101} = Y_{101} - \hat{Y}_{101} = \varepsilon_{101} + 0{,}2\varepsilon_{100} - 0{,}2\varepsilon_{100} = \varepsilon_{101}$$

per cui

$$Var(\varepsilon_{101}) = \sigma_\varepsilon^2 = 0{,}000625$$

da cui l'errore standard di previsione un passo in avanti risulta essere:

$$\sigma_\varepsilon = \sqrt{0{,}000625} = 0{,}025.$$

(b) Essendo il processo un $MA(1)$, la previsione $h$ passi in avanti per $h > 1$ è pari a zero. L'errore di previsione risulta quindi pari a:

$$e_{100+h} = Y_{100+h} = \varepsilon_{100+h} + 0{,}2\varepsilon_{100+h-1}$$

da cui la varianza dell'errore di previsione:

$$Var(e_{100+h}) = Var(\varepsilon_{100+h} + 0{,}2\varepsilon_{100+h-1}) = Var(\varepsilon_{100+h}) + 0{,}04Var(\varepsilon_{100+h}) = 0.00065$$

e quindi l'errore standard risulta essere 0.0255.

**C7.3** Usando l'operatore ritardo $B$ il processo può essere riscritto come:

$$(1 - 0{,}5B)Y_t = (1 - 1{,}2B + 0{,}35B^2)\varepsilon_t.$$

(a) Essendo il coefficiente autoregressivo pari a 0,5 - e quindi inferiore ad uno - il processo $Y_t$ è certamente stazionario: la radice della parte autoregressiva vale infatti $B = 1/0{,}5 = 2$. Per verificare l'invertibilità dobbiamo calcolare le due radici dell'equazione caratteristica:

$$1 - 1{,}2B + 0{,}35B^2 = 0$$

e verificare se sono entrambe in modulo maggiori di uno. Essendo $\Delta = 1{,}2^2 - 4 \cdot 0{,}35 = 0{,}04$ le due radici sono:

$$B_1 = \frac{1{,}2 - \sqrt{0{,}04}}{0{,}7} = \frac{10}{7} \qquad \text{e} \qquad B_2 = \frac{1{,}2 + \sqrt{0{,}04}}{0{,}7} = \frac{14}{7} = 2.$$

Essendo entrambe le radici in modulo maggiori di uno, possiamo affermare che il processo è anche invertibile.

(b) Osserviamo che la seconda radice $B_2 = 2$ è uguale alla radice della componente autoregressiva e quindi possiamo riscrivere il processo come:

$$(1 - 0{,}5B)Y_t = (1 - 0{,}7B)(1 - 0{,}5B)\varepsilon_t$$

ossia

$$Y_t = (1 - 0{,}7B)\varepsilon_t.$$

Il processo $Y_t$ appartiene quindi alla classe dei processi ARIMA(0,0,1) o MA(1).

(c) Essendo il processo un MA(1), è noto che $\rho_k = 0$ per $k > 1$, mentre per $k = 1$ abbiamo $\rho_1 = \gamma_1/\gamma_0$. Abbiamo:

$$\gamma_0 = Var(Y_t) = Var(\varepsilon_t - 0{,}7\varepsilon_{t-1}) = (1 + 0{,}7^2)\sigma_\varepsilon^2 = 1{,}49\sigma_\varepsilon^2$$

$$\gamma_1 = Cov(Y_t, Y_{t-1}) = E[(\varepsilon_t - 0{,}7\varepsilon_{t-1})(\varepsilon_{t-1} - 0{,}7\varepsilon_{t-2})] = -0{,}7\sigma_\varepsilon^2$$

da cui

$$\rho_1 = -\frac{0{,}7\sigma_\varepsilon^2}{1{,}49\sigma_\varepsilon^2} = -0{,}47$$

(d) Per quanto riguarda le previsioni, queste si calcolano a partire da:

$$\hat{Y}_{n+h} = E[Y_{n+h} \mid I_n] = E[\varepsilon_{n+h} - 0{,}7\varepsilon_{n+h-1} \mid I_n] = E_n[\varepsilon_{n+h} - 0{,}7\varepsilon_{n+h-1}]$$

Essendo

$$E_n[\varepsilon_{n+j}] = \begin{cases} \varepsilon_{n+j} & j \leq 0 \\ 0 & j > 0 \end{cases}$$

dove, in pratica $\varepsilon_{n+j}$, $j \leq 0$, viene stimato da $e_{n+j} = y_{n+j} - \hat{y}_{n+j}$, abbiamo:

$$\hat{Y}_{n+1} = E_n[\varepsilon_{n+1} - 0{,}7\varepsilon_n] = E_n[\varepsilon_{n+1}] - 0{,}7E_n[\varepsilon_n] = -0{,}7e_n$$

mentre per $h > 1$ tutte le previsioni sono uguali a 0.

**C7.4** Anzitutto, per verificare che $X_t \sim WN(0,1)$, è necessario calcolare valore atteso, varianza e funzione di autocovarianza:

(a) Si ha:

$$E[X_t] = \begin{cases} E[\varepsilon_t] = 0 \quad \text{per } t \text{ pari} \\ E\left[\frac{\epsilon_{t-1}^2 - 1}{\sqrt{2}}\right] = \frac{1}{2}\left(E[\varepsilon_{t-1}^2] - 1\right) = 0 \quad \text{per } t \text{ dispari} \end{cases}$$

(Infatti $E[\varepsilon_{t-1}^2] = Var(\varepsilon_{t-1}) = 1$).
Inoltre:

$$Var[X_t] = \begin{cases} Var[\varepsilon_t] = 1 \quad \text{per } t \text{ pari} \\ Var\left[\frac{\epsilon_{t-1}^2 - 1}{\sqrt{2}}\right] = \frac{1}{2}Var(\varepsilon_{t-1}^2) = \frac{1}{2}(E[\varepsilon_{t-1}^4] - 1) = 1 \quad \text{per } t \text{ dispari} \end{cases}$$

Infine:

$$Cov[X_t, X_{t-1}] = \begin{cases} Cov\left[\varepsilon_t, \frac{\varepsilon_{t-1}^2 - 1}{\sqrt{2}}\right] = E\left[\varepsilon_t \cdot \frac{\varepsilon_{t-1}^2 - 1}{\sqrt{2}}\right] = \frac{1}{\sqrt{2}}E[\varepsilon_t \cdot \varepsilon_{t-1}^2 - \varepsilon_t] = 0 \quad \text{per } t \text{ pari} \\ Cov\left[\frac{\varepsilon_{t-1}^2 - 1}{\sqrt{2}}, \varepsilon_{t-1}\right] = E\left[\frac{\varepsilon_{t-1}^2 - 1}{\sqrt{2}} \cdot \varepsilon_{t-1}\right] = \frac{1}{\sqrt{2}}E[\varepsilon_{t-1}^3 - \varepsilon_{t-1}] = 0 \quad \text{per } t \text{ dispari} \end{cases}$$

Quindi abbiamo dimostrato quanto si voleva.

(b) Se $n$ è pari allora:

$$\hat{X}_{n+1} = E\left[\frac{\varepsilon_n^2 - 1}{\sqrt{2}} \mid \varepsilon_n, \cdots, \varepsilon_0\right] = \frac{\varepsilon_n^2 - 1}{\sqrt{2}} = \frac{X_n^2 - 1}{\sqrt{2}}$$

Se, invece, $n$ è dispari, allora:

$$\hat{X}_{n+1} = E[\varepsilon_{n+1} \mid \varepsilon_{n-1} \cdots \varepsilon_0] = E[\varepsilon_{n+1}] = 0$$

Si noti che questo succede perché $X_t$ è WN, e quindi è un processo con componenti incorrelate che, tuttavia, non sono indipendenti.

# Appendice 1. Uno sguardo ad alcuni risultati presentati in Di Fonzo e Lisi (2005)

**Effetto di Slutsky-Yule: come si ricava l'espressione a pag. 68**

$$
\begin{aligned}
Cov\left[\varepsilon_t^*, \varepsilon_{t+h}^*\right] &= \sum_{i=-m}^{m} \sum_{j=-m}^{m} \vartheta_i \vartheta_j E\left[\varepsilon_{t+i}^* \varepsilon_{t+h+j}^*\right] \\
&= \begin{cases} \sigma^2 \displaystyle\sum_{i=h-m}^{m} \vartheta_i \vartheta_{i-h} & \text{se } h \leq 2m \\ 0 & \text{altrimenti} \end{cases}
\end{aligned}
$$

Per comprendere questa espressione conviene verificarne la correttezza per qualche valore di $m$. Il caso più semplice è ovviamente $m = 1$:

$$
\varepsilon_t^* = \vartheta_{-1}\varepsilon_{t-1} + \vartheta_0\varepsilon_t + \vartheta_1\varepsilon_{t+1}, \quad \varepsilon_{t+h}^* = \vartheta_{-1}\varepsilon_{t+h-1} + \vartheta_0\varepsilon_{t+h} + \vartheta_1\varepsilon_{t+h+1}
$$

Osserviamo anzitutto che $E\left[\varepsilon_t^*\right] = E\left[\varepsilon_{t+h}^*\right] = 0$, per cui $Cov\left[\varepsilon_t^*, \varepsilon_{t+h}^*\right] = E\left[\varepsilon_t^* \varepsilon_{t+h}^*\right]$. Sviluppando il prodotto $\varepsilon_t^* \varepsilon_{t+h}^*$ si ha:

$$
\begin{aligned}
\varepsilon_t^* \varepsilon_{t+h}^* &= \vartheta_{-1}\varepsilon_{t-1}\varepsilon_{t+h}^* + \vartheta_0\varepsilon_t\varepsilon_{t+h}^* + \vartheta_1\varepsilon_{t+1}\varepsilon_{t+h}^* \\
&= \quad A_1 \quad + \quad A_2 \quad + \quad A_3
\end{aligned}
$$

con

$$
\begin{aligned}
A_1 &= \vartheta_{-1}^2\varepsilon_{t-1}\varepsilon_{t+h-1} + \vartheta_{-1}\vartheta_0\varepsilon_{t-1}\varepsilon_{t+h} + \vartheta_{-1}\vartheta_1\varepsilon_{t-1}\varepsilon_{t+h+1} \\
A_2 &= \vartheta_0\vartheta_{-1}\varepsilon_t\varepsilon_{t+h-1} + \vartheta_0^2\varepsilon_t\varepsilon_{t+h} + \vartheta_0\vartheta_1\varepsilon_t\varepsilon_{t+h+1} \\
A_3 &= \vartheta_1\vartheta_{-1}\varepsilon_{t+1}\varepsilon_{t+h-1} + \vartheta_1\vartheta_0\varepsilon_{t+1}\varepsilon_{t+h} + \vartheta_1^2\varepsilon_{t+1}\varepsilon_{t+h+1}
\end{aligned}
$$

Il calcolo del valore atteso $E\left[\varepsilon_t^*\varepsilon_{t+h}^*\right] = E[A_1] + E[A_2] + E[A_3]$ va fatto per diversi valori di $h$, come riportato nella tavola seguente, in cui vengono messi in evidenza gli addendi di $A_i$, $i = 1,2,3$, che danno contributo diverso da zero al valore atteso.

| Addendi di $\varepsilon_t^*\varepsilon_{t+h}^*$ | Contributi al valore atteso diversi da 0 | | |
|---|---|---|---|
| | $h=0$ | $h=1$ | $h=2$ |
| $A_1$ | $\vartheta_{-1}\varepsilon_{t-1}^2$ | — | — |
| $A_2$ | $\vartheta_0\varepsilon_t^2$ | $\vartheta_{-1}\vartheta_0\varepsilon_t^2$ | — |
| $A_3$ | $\vartheta_1\varepsilon_{t+1}^2$ | $\vartheta_1\vartheta_0\varepsilon_{t+1}^2$ | $\vartheta_1\vartheta_{-1}\varepsilon_{t+1}^2$ |
| $E\left[\varepsilon_t^*\varepsilon_{t+h}^*\right]$ | $\sigma_\varepsilon^2\sum_{-1}^{1}\vartheta_i^2$ | $\sigma_\varepsilon^2(\vartheta_0\vartheta_{-1}+\vartheta_1\vartheta_0)$ | $\sigma_\varepsilon^2\vartheta_1\vartheta_{-1}$ |

Infine, si noti che per $m=1$ si ha:

$$\sum_{i=h-m}^{m}\vartheta_i\vartheta_{i-h}=\begin{cases}\displaystyle\sum_{i=-1}^{1}\vartheta_i^2=\vartheta_{-1}^2+\vartheta_0^2+\vartheta_1^2 & \text{per } h=0\\ \displaystyle\sum_{i=0}^{1}\vartheta_i\vartheta_{i-1}=\vartheta_0\vartheta_{-1}+\vartheta_1\vartheta_0 & \text{per } h=1\\ \displaystyle\sum_{i=1}^{1}\vartheta_i\vartheta_{i-2}=\vartheta_1\vartheta_{-1} & \text{per } h=2\end{cases}$$

**Medie mobili di Spencer: come si ricavano i coefficienti di $M_{15}$ a pag. 72**

La m.m. di Spencer di ordine 15 si ottiene componendo 4 m.m., ossia $M_{15}=M_4M_3M_2M_1$, con

$$M_1=M_2=\left\{[4];\frac{1}{4}[1,1,1,1]\right\}$$
$$M_3=\left\{[5];\frac{1}{5}[1,1,\mathbf{1}]\right\}$$
$$M_4=\left\{[5];\frac{1}{4}[-3,3,\mathbf{4}]\right\}$$

**1. Derivazione dei coefficienti della m.m.** $M_2M_1$

$M_2M_1$ è una m.m. di ordine 7 (4+4-1) con coefficienti 'triangolari':

$$M_2M_1=\left\{[7];\frac{1}{16}[1,2,3,\mathbf{4}]\right\}.$$

Questo risultato può essere ricavato nel modo seguente:

$$y_4^*=\frac{1}{4}\left[\frac{y_1+y_2+y_3+y_4}{4}+\frac{y_2+y_3+y_4+y_5}{4}+\frac{y_3+y_4+y_5+y_6}{4}+\frac{y_4+y_5+y_6+y_7}{4}\right]$$
$$y_5^*=\frac{1}{4}\left[\frac{y_2+y_3+y_4+y_5}{4}+\frac{y_3+y_4+y_5+y_6}{4}+\frac{y_4+y_5+y_6+y_7}{4}+\frac{y_5+y_6+y_7+y_8}{4}\right]$$
$$\vdots$$
$$y_t^*=\frac{1}{4}\left[\frac{y_{t-3}+y_{t-2}+y_{t-1}+y_t}{4}+\frac{y_{t-2}+y_{t-1}+y_t+y_{t+1}}{4}+\frac{y_{t-1}+y_t+y_{t+1}+y_{t+2}}{4}+\frac{y_t+y_{t+1}+y_{t+2}+y_{t+3}}{4}\right]$$

e pertanto

$$y_t^*=\frac{y_{t-3}+2y_{t-2}+3y_{t-1}+4y_t+3y_{t+1}+2y_{t+2}+3y_{t+3}}{16},t=4,\ldots,n-3.$$

**2. Derivazione dei coefficienti della m.m.** $M_3M_2M_1$

$M_3M_2M_1$ è una m.m. di ordine 11 (5+7-1) i cui coefficienti si possono ricavare sviluppando l'espressione

$$\begin{aligned} y_6^{**} &= \frac{y_4^* + y_5^* + y_6^* + y_7^* + y_8^*}{5} \\ y_7^{**} &= \frac{y_5^* + y_6^* + y_7^* + y_8^* + y_9^*}{5} \\ &\vdots \\ y_t^{**} &= \frac{y_{t-2}^* + y_{t-1}^* + y_t^* + y_{t+1}^* + y_{t+2}^*}{5} \end{aligned}$$

ossia

$$\begin{aligned} y_t^{**} = \frac{1}{5 \cdot 16} [&y_{t-5} + 2y_{t-4} + 3y_{t-3} + 4y_{t-2} + 3y_{t-1} + 2y_t + y_{t+1} \\ &+ y_{t-4} + 2y_{t-3} + 3y_{t-2} + 4y_{t-1} + 3y_t + 2y_{t+1} + y_{t+2} \\ &+ y_{t-3} + 2y_{t-2} + 3y_{t-1} + 4y_t + 3y_{t+1} + 2y_{t+2} + y_{t+3} \\ &+ y_{t-2} + 2y_{t-1} + 3y_t + 4y_{t+1} + 3y_{t+2} + 2y_{t+3} + y_{t+4} \\ &+ y_{t-1} + 2y_t + 3y_{t+1} + 4y_{t+2} + 3y_{t+3} + 2y_{t+4} + y_{t+5}] \end{aligned}$$

da cui si ricava

$$y_t^{**} = \frac{1}{80} \left(y_{t-5} + 3y_{t-4} + 6y_{t-3} + 10y_{t-2} + 13y_{t-1} + 14y_t + 13y_{t+1} + 10y_{t+2} + 6y_{t+3} + 3y_{t+4} + y_{t+5}\right).$$

Pertanto,

$$M_3M_2M_1 = \left\{ [11]; \frac{1}{80} [1,3,6,10,13,\mathbf{14}] \right\}.$$

**3. Derivazione dei coefficienti della m.m.** $M_{15} = M_4M_3M_2M_1$

$M_4M_3M_2M_1$ è una m.m. di ordine 15 (5+11-1) i cui coefficienti si possono ricavare sviluppando l'espressione

$$\begin{aligned} y_8^{***} &= \frac{-3y_6^{**} + 3y_7^{**} + 4y_8^{**} + 3y_9^{**} - 3y_{10}^{**}}{4} \\ y_9^{***} &= \frac{-3y_7^{**} + 3y_8^{**} + 4y_9^{**} + 3y_{10}^{**} - 3y_{11}^{**}}{4} \\ &\vdots \\ y_t^{***} &= \frac{-3y_{t-2}** + 3y_{t-1}^{**} + 4y_t^{**} + 3y_{t+1}^{**} - 3y_{t+2}^{**}}{4} \end{aligned}$$

ossia

$$\begin{aligned} y_t^{***} = \frac{1}{80 \cdot 4} [&-3y_{t-7} - 9y_{t-6} - 18y_{t-5} - 30y_{t-4} - 39y_{t-3} - 42y_{t-2} - 39y_{t-1} - 30y_t - 18y_{t+1} - 9y_{t+2} - 3y_{t+3} \\ &+ 3y_{t-6} + 9y_{t-5} + 18y_{t-4} + 30y_{t-3} + 39y_{t-2} + 42y_{t-1} + 39y_t + 30y_{t+1} + 18y_{t+2} + 9y_{t+3} + 3y_{t+4} \\ &+ 4y_{t-5} + 12y_{t-4} + 24y_{t-3} + 40y_{t-2} + 52y_{t-1} + 56y_t + 52y_{t+1} + 40y_{t+2} + 24y_{t+3} + 12y_{t+4} + 4y_{t+5} \\ &+ 3y_{t-4} + 9y_{t-3} + 18y_{t-2} + 30y_{t-1} + 39y_t + 42y_{t+1} + 39y_{t+2} + 30y_{t+3} + 18y_{t+4} + 9y_{t+5} + 3y_{t+6} \\ &- 3y_{t-3} - 9y_{t-2} - 18y_{t-1} - 30y_t - 39y_{t+1} - 42y_{t+2} - 39y_{t+3} - 30y_{t+4} - 18y_{t+5} - 9y_{t+6} - 3y_{t+7}] \end{aligned}$$

da cui si ricava

$$y_t^{***} = \frac{1}{320} \left(-3y_{t-7} - 6y_{t-6} - 5y_{t-5} + 3y_{t-4} + 21y_{t-3} + 46y_{t-2} + 67y_{t-1} + 74y_t + 67y_{t+1} + 46y_{t+2} + 21y_{t+3} + 3y_{t+4} - 5y_{t+5} - 6y_{t+6} - 3y_{t+7}\right).$$

### La media di un processo $AR(p)$: come si ricava l'espressione a pag. 121

Vogliamo ricavare l'espressione

$$E[Y_t] = \frac{\phi_0}{1 - \phi_1 - \ldots - \phi_p}$$

Per il processo $AR(p)$ vale quanto segue:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \quad \Rightarrow \quad E[Y_t] = \phi_0 + \phi_1 E[Y_{t-1}] + \ldots + \phi_p E[Y_{t-p}] + E]\varepsilon_t].$$

Per un processo stazionario si ha

$$E[Y_t] = E[Y_{t-1}] = \ldots = E[Y_{t-p}] = \mu$$

e pertanto

$$\mu = \phi_0 + \phi_1\mu + \ldots + \phi_p\mu \quad \Rightarrow \mu - \phi_1\mu - \ldots - \phi_p\mu = \phi_0 \quad \Rightarrow \mu = \frac{\phi_0}{1 - \phi_1 - \ldots \phi_p}.$$

### La funzione di autocovarianza di un processo $ARMA(1,1)$: come si ricava l'espressione a pag. 126

Vogliamo ricavare l'espressione

$$Var[Y_t] = \gamma_0 = \frac{(1 + \theta_1^2 - 2\theta_1\phi_1)}{1 - \phi_1^2}\sigma_\varepsilon^2$$

$$\gamma_k = \begin{cases} \dfrac{(\phi_- \theta_1)(1 - \phi_1\theta_1)}{(1 - \phi_1^2}\sigma_\varepsilon^2 \; k = 1 \\ \phi_1\gamma_{k-1} = \phi_1^{k-1}\gamma_1 \qquad k \geq 2 \end{cases}$$

Assumendo per comodità $\phi_0 = 0$ (il processo ha media nulla), si ha:

$$\begin{aligned} Var[Y_t] = \gamma_0 = E[Y_t^2] &= E\left[(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}^2\right] \\ &= E\left[\phi_1^2 Y_{t-1}^2 + \varepsilon_t^2 + \theta_1^2\varepsilon_{t-1}^2 + 2\phi_1 Y_{t-1}\varepsilon_t - 2\theta_1\phi_1 Y_{t-1}\varepsilon_{t-1} - 2\theta_1\varepsilon_t\varepsilon_{t-1}\right] \end{aligned}$$

Nello sviluppare questa espressione conviene tenere presente che:

- $E[Y_{t-1}^2] = E[Y_t^2] = Var[Y_t]$, perché il processo è stazionario;
- $E[\varepsilon_{t-1}^2] = E[\varepsilon_t^2] = \sigma_\varepsilon^2$;
- $E[Y_{t-1}\varepsilon_t] = 0$ perché $Y_{t-1}$ dipende da $\varepsilon_{t-1}$,$\varepsilon_{t-2}\varepsilon_{t-3}, \ldots$;
- $E[Y_{t-1}\varepsilon_{t-1}] = E\left[(\phi_1 Y_{t-2} + \varepsilon_{t-1} - \theta_1\varepsilon_{t-2})\varepsilon_{t-1}\right] = \sigma_\varepsilon^2$ .

Pertanto

$$Var[Y_t] = \phi_1^2 Var[Y_t] + \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 - 2\theta_1\phi_1\sigma_\varepsilon^2$$

da cui si ricava

$$(1 - \phi_1^2)Var[Y_t] = (1 + \theta_1^2 - 2\theta_1\phi_1)\sigma_\varepsilon^2 \quad \Rightarrow \quad Var[Y_t] = \frac{(1 + \theta_1^2 - 2\theta_1\phi_1)}{1 - \phi_1^2}\sigma_\varepsilon^2.$$

Calcoliamo ora $\gamma_k$ per $k > 0$.
$k = 1$

$$E\left[Y_{t-1}Y_t\right] = E\left[Y_{t-1}(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1})\right] = E\left[\phi_1 Y_{t-1}^2 + Y_{t-1}\varepsilon_t - \theta_1 Y_{t-1}\varepsilon_{t-1}\right]$$

$$= \phi_1\gamma_0 - \theta_1\sigma_\varepsilon^2 = \left(\phi_1\frac{1+\theta_1^2-2\theta_1\phi_1}{1-\phi_1^2} - \theta_1\right)\sigma_\varepsilon^2$$

e pertanto

$$\gamma_1 = \frac{(\phi_1-\theta_1)(1-\phi_1\theta_1)}{1-\phi_1^2}\sigma_\varepsilon^2.$$

$k > 1$

$$E\left[Y_{t-k}Y_t\right] = E\left[Y_{t-k}(\phi_1 Y_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1})\right]$$

$$= E\left[\phi_1 Y_{t-k}Y_{t-1} + Y_{t-k}\varepsilon_t - \theta_1 Y_{t-k}\varepsilon_{t-1}\right] = \phi_1\gamma_{k-1}$$

e pertanto

$$\gamma_k = \phi_1^{k-1}\gamma_1.$$

# Appendice 2

**Scomposizione del trinomio** $ax^2 + bx + c$

Assumiamo $\Delta = b^2 - 4ac \geq 0$. Indicate con

$$x_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \qquad x_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

le soluzioni (reali, eventualmente coincidenti se $\Delta = 0$) dell'equazione $ax^2 + bx + c = 0$, il trinomio $ax^2 + bx + c$ può essere scomposto nel modo seguente:

$$ax^2 + bx + c = c\left(1 - \frac{1}{x_1}x\right)\left(1 - \frac{1}{x_2}x\right).$$

Tale risultato, utile per semplificare un modello ARMA, deriva dalla più nota fattorizzazione

$$ax^2 + bx + c = a\left(x - x_1\right)\left(x - x_2\right).$$

Infatti, dividendo quest'ultima espressione per $a$, possiamo scrivere

$$\frac{1}{a}\left(ax^2 + bx + c\right) = \left(x - x_1\right)\left(x - x_2\right).$$

Dividendo quindi per $x_1x_2$, si ottiene

$$\frac{1}{ax_1x_2}\left(ax^2 + bx + c\right) = \left(1 - \frac{1}{x_1}x\right)\left(1 - \frac{1}{x_2}x\right).$$

Poiché, com'è noto, $x_1x_2 = \dfrac{c}{a}$, l'espressione precedente diventa

$$\frac{1}{c}\left(ax^2 + bx + c\right) = \left(1 - \frac{1}{x_1}x\right)\left(1 - \frac{1}{x_2}x\right),$$

da cui si ricava $ax^2 + bx + c = c\left(1 - \dfrac{1}{x_1}x\right)\left(1 - \dfrac{1}{x_2}x\right)$.

Ad esempio, si consideri il processo ARMA(2,1)

$$Y_t = 0{,}2Y_{t-1} + 0{,}15Y_{t-2} + \varepsilon_t + 0{,}3\varepsilon_{t-1},$$

con $\varepsilon_t \sim WN(0{,}\sigma^2)$. Le radici – che si ottengono risolvendo l'equazione caratteristica $\phi(B) = 0$ – del polinomio autoregressivo

$$\phi(B) = \left(1 - 0{,}2B - 0{,}15B^2\right)$$

sono pari a

$$B_{1/2} = \frac{0{,}2 \pm \sqrt{0{,}04 + 0{,}6}}{-0{,}3} = \frac{0{,}2 \pm 0{,}8}{-0{,}3} \rightarrow \begin{cases} B_1 = -\frac{10}{3} \\ B_2 = 2 \end{cases}.$$

Poiché in questo caso $c = 1$, possiamo scrivere

$$\phi(B) = \left(1 + \frac{3}{10}B\right)\left(1 - \frac{1}{2}B\right) = (1 + 0{,}3B)\,(1 - 0{,}5B)\,.$$

Il modello ARMA può dunque essere scritto come

$$(1 + 0{,}3B)\,(1 - 0{,}5B)\,Y_t = (1 + 0{,}3B)\,\varepsilon_t$$

che, semplificando $(1 + 0{,}3B)$, porta al modello AR(1)

$$Y_t = 0{,}5Y_{t-1} + \varepsilon_t.$$

# Appendice 3

**La regione di stazionarietà di un processo AR(2)**[1]

Si consideri il generico processo AR(2)

$$(1 - \phi_1 B - \phi_2 B^2) Y_t = \varepsilon_t$$

ossia

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t.$$

Il processo AR(2), essendo un modello autoregressivo di ordine finito, è sempre invertibile. Per essere stazionario, le radici dell'equazione caratteristica $(1 - \phi_1 B - \phi_2 B^2) = 0$ devono giacere all'esterno del cerchio di raggio unitario (ossia, i loro valori assoluti devono essere maggiori di uno). Per esempio, il processo $(1-1.5B+0.56B^2)Y_t = \varepsilon_t$ è stazionario perché le due radici dell'equazione $(1-1.5B+0.56B^2) = 0$, date da

$$B_{1/2} = \frac{1.5 \pm \sqrt{2.25 - 2.24}}{1.12} = \begin{cases} \frac{10}{7} \\ \frac{5}{4} \end{cases}$$

sono entrambe maggiori di 1. Il processo $(1 - 0.2B - 0.8B^2)Y_t = \varepsilon_t$, invece, è non stazionario, perché una delle due radici dell'equazione $(1 - 0.2B - 0.8B^2) = 0$ è $B_1 = 1$, che non è esterna al cerchio di raggio unitario (mentre lo è il valore assoluto della seconda radice, $B_2 = -\frac{5}{4}$)[2].

La condizione di stazionarietà di un modello AR(2) può essere espressa anche con riferimento ai suoi parametri $\phi_1$ e $\phi_2$. Indichiamo con $B_1$ e $B_2$ le radici di $(1 - \phi_1 B - \phi_2 B^2) = 0$ o, in maniera equivalente, di $(\phi_2 B^2 + \phi_1 B - 1) = 0$. Si ha:

$$B_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \quad B_2 = \frac{-\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2}.$$

Ora, si noti che

$$\frac{1}{B_1} = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}. \tag{8.1}$$

L'espressione (8.1) si ottiene nel modo seguente:

$$B_1 = \frac{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2\phi_2} \Rightarrow \frac{1}{B_1} = \frac{2\phi_2}{-\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}} \Rightarrow \frac{1}{B_1} = \frac{2\phi_2(\phi_1 + \sqrt{\phi_1^2 + 4\phi_2})}{-\phi_1^2 + \phi_1^2 + 4\phi_2},$$

1. Rielaborazione da Wei, W.W.S., *Time Series Analysis. Univariate and Multivariate Methods*, 2nd ed., Addison Wesley, 2005, pp. 39-41

2. Il lettore è pregato di verificare questo risultato.

da cui, semplificando, si arriva alla (8.1). Procedendo in modo analogo, si ricava:

$$\frac{1}{B_2} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}. \tag{8.2}$$

La condizione di stazionarietà $|B_i| > 1$, $i = 1{,}2$, implica $\left|\frac{1}{B_i}\right| < 1$, $i = 1{,}2$. Deve pertanto essere

$$\left|\frac{1}{B_1} \cdot \frac{1}{B_2}\right| < 1 \Rightarrow \left|\frac{\phi_1^2 - \phi_1^2 - 4\phi_2}{4}\right| = |\phi_2| < 1$$

e

$$\left|\frac{1}{B_1} + \frac{1}{B_2}\right| < 2 \Rightarrow \left|\frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2} + \phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}\right| = |\phi_1| < 2.$$

Abbiamo dunque derivato due condizioni *necessarie* per la stazionarietà del processo AR(2), indipendentemente dal fatto che le radici siano reali o complesse:

$$\begin{cases} -1 < \phi_2 < 1 \\ -2 < \phi_1 < 2 \end{cases}. \tag{8.3}$$

Si hanno radici reali quando $\phi_1^2 + 4\phi_2 \geq 0$, il che avviene sempre per $\phi_2 > 0$ (che, tenuto conto dell'espressione precedente, equivale a $0 < \phi_2 < 1$), e ciò comporta:

$$-1 < \frac{1}{B_2} = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2} \leq \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} = \frac{1}{B_1} < 1. \tag{8.4}$$

Mostriamo che l'espressione (8.4) è equivalente a

$$\begin{cases} \phi_2 + \phi_1 < 1 \\ \phi_2 - \phi_1 < 1 \end{cases}. \tag{8.5}$$

Consideriamo prima la parte destra della diseguaglianza (8.4). Procedendo con i calcoli si ha:

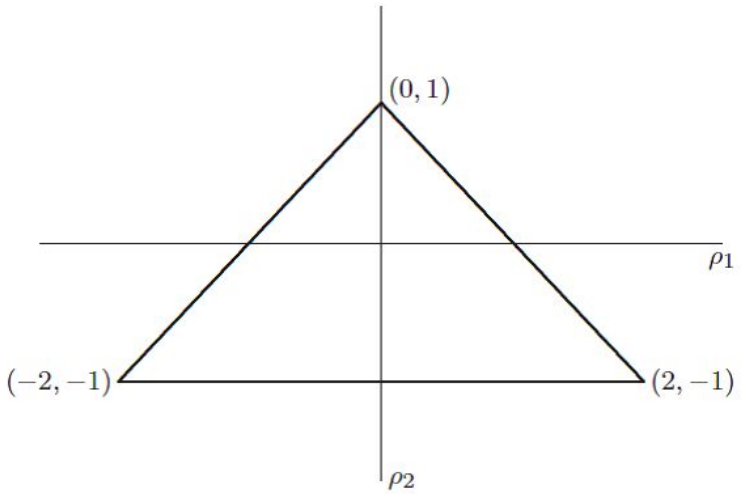
$$\frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} < 1 \Rightarrow \sqrt{\phi_1^2 + 4\phi_2} < 2 - \phi_1 \Rightarrow \phi_1^2 + 4\phi_2 < 4 - 4\phi_1 + \phi_1^2 \Rightarrow \phi_2 + \phi_1 < 1.$$

Analogamente, sviluppando la parte sinistra della diseguaglianza (8.4) si ricava:

$$-1 < \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2} \Rightarrow \sqrt{\phi_1^2 + 4\phi_2} < 2 + \phi_1 \Rightarrow \phi_1^2 + 4\phi_2 < 4 + 4\phi_1 + \phi_1^2 \Rightarrow \phi_2 - \phi_1 < 1.$$

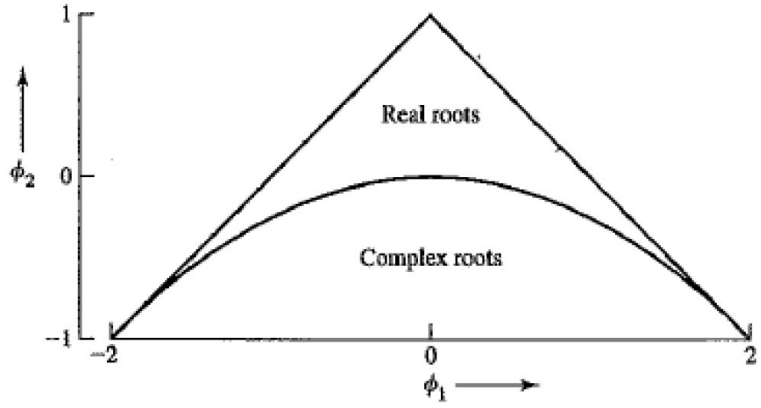
La sezione di piano limitata dalla (8.3) e dalla (8.5) costituisce la regione entro cui possono variare i parametri $\phi_1$ e $\phi_2$ di un processo AR(2) stazionario. Data la forma triangolare (figura 8.3), tale regione viene detta *Triangolo di stazionarietà* di un processo AR(2).

Figura 8.3
Triangolo di stazionarietà di un processo AR(2)




Ora, si noti che quando $\phi_1^2 + 4\phi_2 < 0$ si hanno radici complesse. Ciò avviene per $-1 < \phi_2 < 0$ e $\phi_2 < -\frac{1}{4}\phi_1^2$, ossia nella porzione del triangolo delimitata superiormente dalla parabola $\phi_2 = -\frac{1}{4}\phi_1^2$ (figura 8.5).

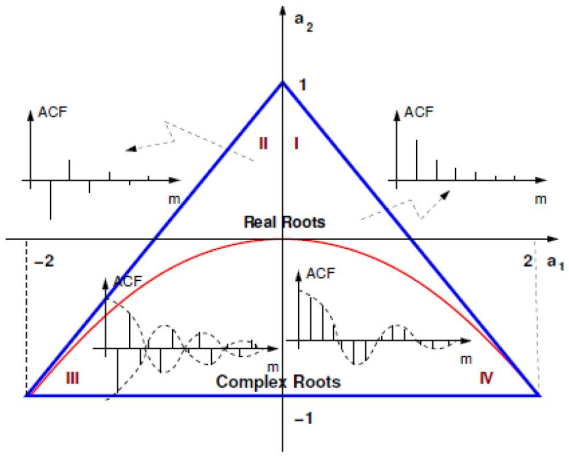
Figura 8.5
Radici reali e complesse nel triangolo di stazionarietà del processo AR(2)




Nella parte restante del triangolo di stazionarietà si collocano invece le coppie di parametri $(\phi_1, \phi_2)$ dei processi AR(2) stazionari a cui corrispondono radici reali dell'equazione caratteristica $(1 - \phi_1 B - \phi_2 B^2) = 0$.

Qualche dettaglio sulla forma dei correlogrammi globali al variare delle coppie di parametri $(\phi_1, \phi_2)$ viene mostrato nella figura 8.7.

Figura 8.7
ACF del processo AR(2) nel triangolo di stazionarietà




i) **Real roots** Region 1: Monotonically decaying ACF
ii) **Real roots** Region 2: Decaying oscillating ACF
iii) **Complex roots** Region 3: Oscilating pseudoperiodic ACF
iv) **Complex roots** Region 4: Pseudoperiodic ACF

**Un esercizio**

(a) Trovare l'insieme dei valori di $\alpha$ tali per cui il processo AR(2)

$$Y_t = Y_{t-1} + \alpha Y_{t-2} + \varepsilon_t$$

è stazionario.

(b) Calcolare $\rho_3$ per $\alpha = -\frac{1}{2}$.

**Soluzione**

(a) Prima di procedere con i calcoli, notiamo anzitutto che per $\alpha = 0$ si ha un processo *random walk*, che non è stazionario. Pertanto, assumiamo $\alpha \neq 0$. Le radici dell'equazione caratteristica $-\alpha B^2 - B + 1 = 0$ o, che è lo stesso, $\alpha B^2 + B - 1 = 0$, sono pari a:

$$B_{1/2} = \frac{-1 \pm \sqrt{1+4\alpha}}{2\alpha}.$$

Perché il processo sia stazionario si deve avere $|B_1| > 1$ e $|B_2| > 1$. Se consideriamo i reciproci delle due soluzioni,

$$\frac{1}{B_1} = \frac{1+\sqrt{1+4\alpha}}{2} \quad \text{e} \quad \frac{1}{B_2} = \frac{1-\sqrt{1+4\alpha}}{2},$$

per un processo stazionario si deve avere $\left|\frac{1}{B_1} \cdot \frac{1}{B_2}\right| < 1$, ossia $\left|\frac{1}{B_1} \cdot \frac{1}{B_2}\right| = \left|\frac{1-1-4\alpha}{4}\right| = |\alpha| < 1$.
Se $1 + 4\alpha \geq 0$ (caso di radici reali) vale la seguente diseguaglianza:

$$-1 < \frac{1}{B_2} = \frac{1-\sqrt{1+4\alpha}}{2} < \frac{1+\sqrt{1+4\alpha}}{2} = \frac{1}{B_1} < 1.$$

La parte sinistra della diseguaglianza, $-1 < \dfrac{1-\sqrt{1+4\alpha}}{2}$, è verificata per $\alpha < 2$, mentre quella destra, $\dfrac{1+\sqrt{1+4\alpha}}{2} < 1$ , è verificata per $\alpha < 0$.
Combinando dunque le due condizioni trovate, $|\alpha| < 1$ e $\alpha < 0$, si ricava la condizione di stazionarietà

$$-1 < \alpha < 0.$$

Inoltre, l'equazione caratteristica del processo ha radici reali per $1 + 4\alpha \geq 0$, ossia per $-\dfrac{1}{4} \leq \alpha < 0$, e radici complesse per $-1 < \alpha < -\dfrac{1}{4}$.

(b) Per calcolare la funzione di autocorrelazione globale, si moltiplichino ambo i lati della forma esplicita del processo, $Y_t = Y_{t-1} + \alpha Y_{t-2} + \varepsilon_t$, per $Y_{t-k}$, si prenda il valore atteso

$$E(Y_{t-k}Y_t) = E(Y_{t-k}Y_{t-1}) + \alpha E(Y_{t-k}Y_{t-2}) + E(Y_{t-k}\varepsilon_t) \Rightarrow \gamma_k = \gamma_{k-1} + \alpha\gamma_{k-2}, \quad k \geq 1$$

e si divida il risultato per $\gamma_0$:

$$\rho_k = \rho_{k-1} + \alpha\rho_{k-2}, \quad k \geq 1.$$

Sostituendo $-\dfrac{1}{2}$ ad $\alpha$, e tenuto conto che $\rho_1 = \rho_{-1}$, si ottiene

$$\begin{aligned} \rho_1 &= 1 - \frac{1}{2}\rho_1 \Rightarrow \rho_1 = \frac{2}{3} \\ \rho_2 &= \rho_1 - \frac{1}{2} \Rightarrow \rho_2 = \frac{1}{6} \\ \rho_3 &= \rho_2 - \frac{1}{2}\rho_1 \Rightarrow \rho_3 = \frac{1}{6} - \frac{1}{3} = -\frac{1}{6} \end{aligned} .$$