SERIE STORICHE

Padova, 10 febbraio 2025

Nome Cognome N. matricola

Domanda 1

Si consideri il processo Y_t, contenente un trend quadratico deterministico e una componente erratica:

$$Y_t = \alpha_0 + \alpha_1 \cdot t + \alpha_2 \cdot t^2 + \varepsilon_t$$

con $\varepsilon_t \sim WN(0, \sigma^2)$.
1. Si dica, giustificando la risposta, se tale processo è stazionario.
2. Si consideri il processo $X_t = (1-B)^2 Y_t$ e si dica a che classe dei processi ARIMA appartiene.
3. Si dica, giustificando la risposta, se il processo X_t è stazionario e/o invertibile.

Soluzione:

1. Dato che $E[Y_t] = \alpha_0 + \alpha_1 \cdot t + \alpha_2 \cdot t^2$ è evidente che il processo non è stazionario in quanto il suo valore atteso dipende dal tempo.

2. Per il processo X_t si ha:

$$X_t = (1-B)^2 Y_t = (1 - 2B + B^2) Y_t = Y_t - 2Y_{t-1} + Y_{t-2},$$

quindi

$$\begin{aligned} X_t &= \alpha_0 + \alpha_1 \cdot t + \alpha_2 \cdot t^2 + \varepsilon_t \\ &\quad -2\alpha_0 - 2\alpha_1 \cdot (t-1) - 2\alpha_2 \cdot (t-1)^2 - 2\varepsilon_{t-1} \\ &\quad +\alpha_0 + \alpha_1 \cdot (t-2) + \alpha_2 \cdot (t-2)^2 + \varepsilon_{t-2} \\ &= 2\alpha_2 + \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} \end{aligned}$$

che è un processo ARIMA(0,0,2) o MA(2)

3. Essendo un processo MA, X_t è certamente stazionario, mentre per vedere se è anche invertibile bisogna calcolare le radici dell'equazione caratteristica:

$$1 - 2B + B^2 = 0$$

Essendo $1 - 2B + B^2 = (1-B)^2$ si ottengono due radici unitarie e quindi il processo X_t non è invertibile.

Domanda 2

Dato un modello SARIMA(1,0,2)$\times$(1,0,0)$_4$ con parametri $\phi_1 = 0.4$, $\theta_1 = 0.5$, $\theta_2 = 0.2$, $\Phi_1 = -0.6$ e $\sigma^2 = 1.5$, calcolare la previsione con orizzonte temporale $k = 5$ sapendo che:

- l'ultimo valore osservato della serie storica è $y_n = 5$;
- $\hat{y}_{n+1|n} = 7$ e $\hat{y}_{n+4|n} = 4$.

Soluzione

Il processo indicato si esprime nella forma:

$$(1 - \phi_1 B)(1 - \Phi_1 B^4)Y_t = (1 - \theta_1 B - \theta_2 B^2)\varepsilon_t.$$

Poiché

$$(1 - \phi_1 B)(1 - \Phi_1 B^4) = (1 - \phi_1 B - \Phi_1 B^4 + \phi_1 \Phi_1 B^5)$$

il processo può essere riscritto nel modo seguente:

$$Y_t = \phi_1 Y_{t-1} + \Phi_1 Y_{t-4} - \phi_1 \Phi_1 Y_{t-5} - \theta_1 \varepsilon_{t-1} - \theta_1 \varepsilon_{t-2} + \varepsilon_t$$

Il previsore di Y_{n+5} è dunque pari a:

$$\begin{aligned} \hat{Y}_{n+5|n} &= E_n[Y_{n+5}] \\ &= \phi_1 E_n[Y_{n+4}] + \Phi_1 E_n[Y_{n+1}] - \phi_1 \Phi_1 E_n[Y_n] + \\ &\quad -\theta_1 E_n[\varepsilon_{n+4}] - \theta_2 E_n[\varepsilon_{n+3}] + E_n[\varepsilon_{n+5}]. \end{aligned}$$

Il calcolo dei valori attesi condizionati avviene secondo le seguenti espressioni:

$$\begin{aligned} E_n[Y_{n+j}] &= \begin{cases} y_{n+j} & j \leq 0 \\ \hat{y}_{n+j|n} & j > 0 \end{cases} \\ E_n[\varepsilon_{n+j}] &= \begin{cases} e_{n+j} & j \leq 0 \\ 0 & j > 0 \end{cases} . \end{aligned}$$

Pertanto, la previsione $\hat{y}_{n+5|n}$ è pari a

$$\begin{aligned} \hat{y}_{n+5|n} &= \phi_1 \hat{y}_{n+4|n} + \Phi_1 \hat{y}_{n+1|n} - \phi_1 \Phi_1 y_n \\ &= 0.4 \cdot (4) - 0.6 \cdot 7 - 0.4 \cdot (-0.6) \cdot 5 = -1.4 \end{aligned}$$

Domanda 3

1. Descrivere la media mobile M, ottenuta come *composizione* di due medie mobili aritmetiche semplici centrate.
2. Dire, giustificando la risposta, se la media mobile M può essere utilizzata per stimare il trend di una serie che presenta una componente di lungo periodo quadratica e una componente stagionale di periodo 5.
3. Dire, giustificando la risposta, se la media mobile M è opportuna per ridurre la variabilità della componente erratica, quantificandone gli effetti sulla variabilità stessa.

Soluzione

1.

$$M_1 y_t = y_t* = \frac{y_{t-1} + y_t + y_{t+1}}{3}$$

$$M_2 y_t = y_t* = \frac{y_{t-1} + y_t + y_{t+1}}{3}$$

$$\begin{aligned} M_2(M_1 y_t) = y_t^{**} &= \frac{y^*_{t-1} + y^*_t + y^*_{t+1}}{3} = \\ &= \frac{1}{3}\left(\frac{y_{t-2} + y_{t-1} + y_t}{3} + \frac{y_{t-1} + y_t + y_{t+1}}{3} + \frac{y_t + y_{t+1} + y_{t+2}}{3}\right) = \\ &= \frac{1}{9}y_{t-2} + \frac{2}{9}y_{t-1} + \frac{3}{9}y_t + \frac{2}{9}y_{t+1} + \frac{1}{9}y_{t+2} \qquad t = 3, \ldots, n-2 \end{aligned}$$

2. La media mobile M non è idonea a stimare il trend perché non preserva un trend quadratico, infatti:

$$\sum_{i=-2}^{2} i \cdot \theta_i = -\frac{2}{9} - \frac{2}{9} + \frac{2}{9} + \frac{2}{9} = 0$$

$$\sum_{i=-2}^{2} i^2 \cdot \theta_i = \frac{4}{9} + \frac{2}{9} + \frac{2}{9} + \frac{4}{9} = \frac{4}{3} \neq 0$$

3. La media mobile M riduce la variabilità della componente erratica, essendo il rapporto di riduzione della varianza residua pari a:

$$\sum_{i-2}^{2} \theta_i^2 = \frac{19}{81} = 0.234560$$

, dunque la varianza della componente erratica si riduce di circa il 76.5%.

Domanda 4

Per la serie storica y_t, $(t = 1, ..., 100)$ si vuole stimare un modello AR(2) a media nulla. Avendo osservato $Var(y_t) = 8.7$, $S_1 = \sum_{t=2}^{n}(y_t - \bar{y})(y_{t-1} - \bar{y}) = 650$ e $S_2 = \sum_{t=3}^{n}(y_t - \bar{y})(y_{t-2} - \bar{y}) = 580$, fornire delle stime dei parametri del modello.

Soluzione I parametri da stimare sono ϕ_1, ϕ_2 e σ^2_ε.
Si ha che:

$$\gamma_0 = Var(y_t) = 8.7$$

$$\gamma_1 = \frac{1}{100}\sum_{t=2}^{n}(y_t - \bar{y})(y_{t-1} - \bar{y}) = 6.5$$

$$\gamma_2 = \frac{1}{100}\sum_{t=3}^{n}(y_t - \bar{y})(y_{t-2} - \bar{y}) = 5.8$$

da cui

$$\rho_1 = \gamma_1/\gamma_0 = 0.747$$

$$\rho_2 = \gamma_2/\gamma_0 = 0.666.$$

Sfruttando le equazioni di Yule-Walker si ha:

$$\rho_1 = \phi_1\rho_0 + \phi_2\rho_1 = \phi_1 + \phi_2\rho_1$$

$$\rho_2 = \phi_1\rho_1 + \phi_2\rho_0 = \phi_1\rho_1 + \phi_2.$$

Dalla prima equazione si ha:

$$\phi_1 = (1 - \phi_2)\rho_1$$

che sostituito nella seconda equazione porta a :

$$\rho_2 = ((1 - \phi_2)\rho_1)\rho_1 + \phi_2.$$

e risolvendo in ϕ_2

$$\phi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{0.666 - 0.747^2}{1 - 0.747^2} = 0.244$$

e quindi

$$\phi_1 = (1 - \phi_2)\rho_1 = (1 - 0.244) \cdot 0.747 = 0.564$$

Infine, dato che

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2_\varepsilon$$

si ha:

$$\sigma^2_\varepsilon = 8.7 - 0.564 \cdot 6.5 - 0.244 \cdot 5.8 = 3.61.$$

Domanda 5

Dopo avere specificato cosa si intende per identificazione di un modello della classe SARIMA, si discutano brevemente i metodi con cui tale identificazione si concretizza.

Soluzione (traccia):
Con il termine identificazione di un modello SARIMA si intende l'individuazione degli ordini del modello e del periodo della stagionalità, ovvero p,d,q,P,D,Q ed S.
L'identificazione di un modello SARIMA avviene:
- tramite l'analisi delle sue funzioni di autocorrelazione globale e parziale (spiegare come.);
- tramite l'analisi della significatività dei parametri (spiegare come.)
- tramite criteri di identificazione automatica tipo AIB e BIC (spiegarne il funzionamento)

Domanda 6

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

Soluzione (traccia):
Ci sono 3 metodi per scegliere il grado p del polinomio:
- metodo dell'$\bar{R}^2$: si sceglie il grado p del polinomio che massimizza $\bar{R}_p^{\,2}$ (spiegare come);
- metodo della significatività dei parametri: si sceglie il polinomio di grado p massimo per il quale α_p è significativo.
- metodo delle differenze successive: si basa sul fatto che un trend polinomiale di grado p diventa costante dopo p differenziazioni. Quindi si sceglie p pari al numero di differenziazioni necessarie per ottenere una serie storica che oscilla intorno ad un trend costante.