SERIE STORICHE

Padova, 30 gennaio 2024

Nome Cognome N. matricola

Corso di Laurea

Domanda 1

Per a serie storica $\{y_t\}_{t=1}^{?}200$ è risultato: $\bar{y} = 3.82,\ se(\bar{y}) = 0.023$, $\hat{\gamma}_0 = 1.15$, $\hat{\gamma}_1 = 0.427$, $\hat{\gamma}_2 = 0.475$ e $\hat{\gamma}_3 = 0.169$. Sulla base di queste stime, stabilire se si può affermare che il processo generatore dei dati di $Y_t - \mu$ è white noise?

Soluzione: La media risulta significativa, ma questo non ha un impatto sull'autocorrelazione dei dati, quindi l'autocorrelazione di Y_t e quella di $Y_t - \mu$ è la stessa.
Se $Y_t - \mu$ fosse WN, allora dovrebbe essere la sua funzione di autocorrelazione dovrebbe essere $\hat{\rho}(k) = 0$ per $k \neq 0$. Sappiamo che, sotto ipotesi di white noise, $\hat{\rho}(k) \sim \mathcal{N}(0, 1/n)$ dove, nel nostro caso, $n = 200$. Quindi, se il processo è white noise, un intervallo di confidenza al 95% per $\rho(k)$ è dato da:

$$\pm 1.96 \cdot \frac{1}{\sqrt{200}} = \pm 0.139$$

.

Essendo $\hat{\rho}_k = \hat{\gamma}_k / \hat{\gamma}_0$ si ha:

$\hat{\rho}_1 = 0.427/1.15 = 0.371$, $\hat{\rho}_2 = 0.475/1.15 = 0.413$ e $\hat{\rho}_3 = 0.169/1.15 = 0.147$

che sono tutti esterni alla bande di confidenza. Dunque, i coefficienti di autocorrelazione sono significativamente diversi da zero. Questo ci porta a dire che il processo $Y_t - \mu$ non è WN.

Domanda 2

Si dispone di una serie storica giornaliera che presenta una componente stagionale costante di periodo 7.

1. Stabilire se una media mobile aritmetica semplice a 7 termini elimina tale stagionalità.
2. Derivare l'associato rapporto di riduzione della varianza.
3. Mostrare che la media mobile al punto (1) conserva trend polinomiali di primo grado ma non quelli di secondo grado.

Soluzione: Esercizio 3.2 dell'eserciziario.

1. 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.
2. 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.
3. 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.

Domanda 3 Dopo aver definito il previsore ottimo per un processo AR(1): $Y_t = \phi_0 + \phi_1 Y_{t-1} + \epsilon_t$ con $\epsilon_t \sim WN(0, \sigma_\epsilon^2)$, si consideri una serie storica di $N = 100$ osservazioni a cui si è adattato un modello AR(1) senza costante con $\hat{\phi}_1 = -0.5$. Supponendo che $y_{100} = 12.2$, si trovino $\hat{y}_{101}, \hat{y}_{102}$ e $\hat{y}_{110}$.

Soluzione:

Per un processo AR(1) il previsore ottimo risulta essere:

$$\hat{Y}_{n+h} = E_n[\phi_0 + \phi_1 Y_{n+h-1} + \epsilon_{n+h}] = \phi_0 + \phi_1 E_n[Y_{n+h-1}] + E_n[\epsilon_{n+h}]$$

quindi, ricordando che

$$E_n[Y_{n+h}] = \begin{cases} y_{n+h} & h \leq 0 \\ \hat{y}_{n+h} & h > 0 \end{cases}$$

e, posto $e_t = y_t - \hat{y}_t$

$$E_n[\epsilon_{n+h}] = \begin{cases} e_{n+h} & h \leq 0 \\ 0 & h > 0 \end{cases}$$

si ottiene:

$$\hat{Y}_{n+h} = \begin{cases} \phi_0 + \phi_1 y_n & h = 1 \\ \phi_0(1 + \phi_1 + \phi_1^2 + \ldots + \phi_1^{h-1}) + \phi_1^h y_n & h > 1 \end{cases}$$

Per un processo AR(1) senza costante si ha dunque:

$$\hat{y}_{n+1} = \phi_1 y_n$$
$$\hat{y}_{n+2} = \phi_1 \hat{y}_{n+1} = \phi_1^2 y_n$$
$$\hat{y}_{n+3} = \phi_1 \hat{y}_{n+2} = \phi_1^3 y_n$$
$$\ldots.$$
$$\hat{y}_{n+h} = \phi_1 \hat{y}_{n+h-1} = \phi_1^h y_n$$

Pertanto:

$$\hat{y}_{101} = -0.5 \cdot 12.2 = -6.1, \quad \hat{y}_{102} = (-0.5)^2 \cdot 12.2 = 3.05 \quad e \quad \hat{y}_{110} = (-0.5)^{10} \cdot 12.2 = 0.012$$

Domanda 4

Per ognuno dei seguenti processi:

(a) $Y_t = Y_{t-1} - 0.25Y_{t-2} + \epsilon_t - 0.1\epsilon_{t-1}$

(b) $Y_t = 2Y_{t-1} - Y_{t-2} + \epsilon_t$

(c) $Y_t = 0.8Y_{t-1} - 0.15Y_{t-2} + \epsilon_t - 0.3\epsilon_{t-1} - 0.1\epsilon_{t-2}$

si chiede di:

1. riscriverlo utilizzando l'operatore ritardo B;
2. dire che tipo di modello ARIMA(p, d, q) lo descrive.

Soluzione:

(1.a) $(1 - B + 0.25B^2)Y_t = (1 - 0.1B)\epsilon_t$

(1.b) $(1 - 2B + B^2)Y_t = \epsilon_t$ quindi $(1 - B)^2 Y_t = \epsilon_t$

(1.c) $(1 - 0.8B + 0.15B^2)Y_t = (1 - 0.3B - 0.1B^2)\epsilon_t$

(2.a) Il processo sembra un ARIMA$(2, 0, 1)$, tuttavia per esserne certo (per verificare che non ci siano radici unitarie e/o che non si possa procedere ad una semplificazione) trovo le radici dell'equazione caratteristica

$$1 - B + 0.25B^2 = 0.$$

Essendo $1-B+0.25B^2 = (1-0.5B)^2$ ci sono due radici reali coincidenti e pari a $B_{1,2} = 2$ quindi in modulo maggiori di uno. Resta dunque confermato che il processo è un ARIMA$(2, 0, 1)$.

(2.b) Il processo possiede due radici unitarie, quindi è un ARIMA$(0, 2, 0)$.

(2.c) Il processo sembra un ARIMA$(2, 0, 2)$, tuttavia per esserne certo trovo le radici delle equazioni caratteristiche delle parti AR e MA. Per la parte AR abbiamo:

$$1 - 0.8B + 0.15B^2 = 0$$

che avendo $\Delta = 0.04$ porta alle due soluzioni $B_1 = 2$ e $B_2 = 10/3$. Per la parte MA abbiamo, invece:

$$1 - 0.3B - 0.1B^2 = 0$$

che porta a $\Delta = 0.49$ e quindi a $B_1 = 2$ e $B_2 = -5$. Si nota che abbiamo una radice in comune, e che quindi si può scrivere:

$$1 - 0.8B + 0.15B^2 = (1 - 0.5B)(1 - 0.3B)$$

e

$$1 - 0.3B - 0.1B^2 = (1 - 0.5B)(1 + 0.2B),$$

il che porta a semplificare il processo che risulta dunque essere un ARMA$(1, 1)$ (stazionario e invertibile).

Domanda 5

Per la serie y_t è stato stimato un modello SARIMA$(1,1,1)(1,0,1)_{12}$. Descrivere quali analisi diagnostiche vanno effettuate per stabilire se il modello stimato è accettabile e i relativi strumenti statistici.

Soluzione (traccia):

- va controllata la stazionarietà e l'invertibilità: spiegare come.
- si deve verificare la significatività delle stime parametri: definire test
- bisogna controllare l'incorrelazione dei residui. Strumenti:
ACF e PACF: test sottostante;
Test di Ljung-Box: specificare l'ipotesi nulla del test;
Trattandosi di un modello stagionale bisogna porre particolare attenzione ai ritardi che sono multipli del periodo
Test di Box-Pierce: specificare ipotesi nulla sottostante.
- si può verificare la gaussianità dei residui con il test di Jarque-Bera

Domanda 6

Si descriva in cosa consiste la cosiddetta trappola delle dummy e quali azioni devono essere messe in atto per non cadere in tale trappola.

Soluzione:

Consideriamo un modello di per serie storiche in cui sia presente solo una componente stagionale (ad es. di periodo 4), oltre alla componente erratica

$$y_t = \gamma_0 + \gamma_1 d_{1t} + \gamma_2 d_{2t} + \gamma_3 d_{3t} + \gamma_4 d_{4t} + \varepsilon_t, \quad t = 1, 2, \ldots, n.$$

Così scritto, cioè includendo sia la costante sia le 4 variabili dummy, il modello assume la seguente forma

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ \vdots \\ y_{12} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \vdots \\ \varepsilon_{12} \end{bmatrix} \tag{1}$$

E' chiaro che, in questo caso, la matrice dei regressori non ha rango pieno dato che al prima colonna è una combinazione lineare delle altre 4 colonne. Questo implica che, nel calcolo delle stime LS (minimi quadrati) la matrice $(X'X)^{-1}$ non è più invertibile. Questa circostanza è quella che viene chiamata trappola delle dummy

.

Ci sono due modi per evitare la trappola delle dummy:
1) evitare di includere la costante;
2) inserire costante ed eliminare una dummy (ad es. la prima). In questo caso il modello diventa:

$$y_t = \gamma_0 + \gamma_2^* d_{2t} + \gamma_3^* d_{3t} + \gamma_4^* d_{4t} + \varepsilon_t, \quad t = 1, 2, \ldots, n.$$

dove

$$\begin{aligned} \gamma_0 &= \gamma_1 \\ \gamma_2^* &= \gamma_2 - \gamma_0 \\ \gamma_3^* &= \gamma_3 - \gamma_0 \\ \gamma_4^* &= \gamma_4 - \gamma_0 \end{aligned}$$

rappresentano degli effetti differenziali rispetto al livello tipico del primo periodo (in questo caso).