Previsão em modelos ARIMA

Airlane Pereira Alencar

13 de Abril de 2020

Índice

ARMA(p,q)

$$\begin{aligned}
\tilde{Z}_t &= \phi_1 \tilde{Z}_{t-1} + \ldots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} \ldots - \theta_q a_{t-q} \\
(1 - \phi_1 B - \cdots - \phi_p B)\tilde{Z}_t &= (1 - \theta_1 B \ldots - \theta_q B^q) a_t \\
\Phi(B)\tilde{Z}_t &= \Theta(B) a_t
\end{aligned}$$

com $a_t \sim RB$

- Estacionário: se as raízes de $\Phi(B) = 0$ estiverem fora do círculo unitário
- Invertível: se as raízes de $\Theta(B) = 0$ estiverem fora
- FAC e FACP parece cair para zero e deve propor valores (p,q) até ter resíduo como RB
- Deve propor valores (p,q) até ter resíduo como RB

AR(1) como MA(∞)

$\Phi(B)Z_t = \Theta(B)a_t$
Por causa da estacionariedade, podemos escrever o modelo linear geral (MA(∞))

$$Z_t = \psi(B)a_t = a_t + \psi_1 a_{t-1} + \dots$$

Exemplo: AR(1)

$$Z_t = \phi Z_{t-1} + a_t \Rightarrow Z_t - \phi Z_{t-1} = a_t \Rightarrow (1 - \phi B)Z_t = a_t$$

Queremos

$$Z_t = a_t + \psi_1 a_{t-1} + \dots \Rightarrow Z_t = \psi(B)a_t = (1 + \psi_1 B + \psi_2 B^2 + \dots)a_t$$

Substituo a_t por $(1 - \phi B)Z_t$

$$Z_t = (1 + \psi_1 B + \psi_2 B^2 + \dots)(1 - \phi B)Z_t$$
$$(1 + \psi_1 B + \psi_2 B^2 + \dots)(1 - \phi B) = 1 \Rightarrow$$
$$(1 + \psi_1 B - \phi B + \psi_2 B^2 - \phi\psi_1 B^2 + \dots) = 1 \Rightarrow$$
$$\psi_1 = \phi, \psi_2 = \phi^2, \dots \Rightarrow \psi_j = \phi^j$$

Previsão

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \ldots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$$

Por causa da estacionariedade, escrevemos

$$Z_t = \psi(B) a_t = a_t + \psi_1 a_{t-1} + \ldots$$

No instante n, observamos $Z_n, Z_{n-1}, Z_{n-2}, \ldots$ e queremos prever Z_{n+l} como combinação de $Z_n, Z_{n-1}, Z_{n-2}, \ldots$
Como Z_t é estacionário podemos escrever como MA(∞). A previsão para
$Z_{n+l} = \sum_{j=0}^{\infty} \psi_j a_{n+l-j} = \psi_0 a_{n+l} + \psi_1 a_{n+l-1} + \ldots + \psi_l a_n + \psi_{l+1} a_{n-1} + \ldots$
com a informação até n é $\hat{Z}_n(l) = \psi_l^* a_n + \psi_{l+1}^* a_{n-1} + \psi_{l+2}^* a_{n-2} + \ldots$
Queremos ψ_l^* que minimiza o Erro Quadrático Médio de Previsão:

$$E(Z_{n+l} - \hat{Z}_n(l))^2 = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 + \sigma_a^2 \sum_{j=0}^{\infty} [\psi_{l+j} - \psi_{l+j}^*]^2$$

Minimizamos $\psi_{l+j}^* = \psi_{l+j}$.

Variância da previsão

O erro da previsão é

$$e_n(l) = Z_{n+l} - \hat{Z}_n(l) = \sum_{j=0}^{l-1} \psi_j a_{n+l-j} = \psi_0 a_{n+l} + \ldots + \psi_{l-1} a_{n+1}$$

$\hat{Z}_n(l)$ é não viesado pois $E(e_n(l)) = 0$

$$Var(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$$

Para processo gaussiano

$$IC = [\hat{Z}_n(l) \mp z\sqrt{Var(e_n(l))}]$$

Variância da Previsão

- **Os erros de previsão são correlacionados**

$$Cov(e_n(2), e_n(1)) = E[(a_{n+2} + \psi_1 a_{n+1})a_{n+1}] = \psi_1 \sigma_a^2$$

- Para ARIMA(p,d,q), em Wei(2006), baseado em Wegman(1986) Another Look at Box-Jenkins forecasting procedure, Comm.Stat.Simula, 15(2) 523-530: A variância da previsão fica igual ao do processo ARMA! Escrevemos como MA(∞), mesmo sendo o arima não estacionário. Em geral, as variâncias crescem com o horizonte.

$$Var(e_n(l)) = \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2$$

$$\Psi(B) = \frac{\Theta(B)}{\Psi(B)(1-B)^d}$$

Exemplos ARIMA(0,1,1)

Erros futuros são previstos iguais à média zero.

$$(1-B)Z_t = \theta_0 + a_t - \theta a_{t-1}$$
$$Z_t = \theta_0 + Z_{t-1} + a_t - \theta a_{t-1}$$

$$\begin{aligned}
Z_{t+h} &= \theta_0 + Z_{t+h-1} + a_{t+h} - \theta a_{t-1} \\
\hat{Z}_T(1) &= \theta_0 + Z_T - \theta a_T \\
\hat{Z}_T(2) &= \theta_0 + \hat{Z}_T(1) = 2\theta_0 + Z_T - \theta a_t \\
\hat{Z}_T(3) &= \theta_0 + \hat{Z}_T(2) = 3\theta_0 + Z_T - \theta a_T \\
\hat{Z}_T(h) &= h\theta_0 + Z_T - \theta a_T
\end{aligned}$$

ARIMA(0,1,1)

Apesar de Z_t ser não estacionário, vamos encontrar os primeiros termos $\psi_1, \psi_2, \ldots$

$$Z_t = \Psi(B)a_t = (\psi_0 + \psi_1 B + \psi_2 B^2 + \ldots)a_t$$

$$(1 - B)Z_T = (1 - \theta B)a_t$$

Substituindo Z_t, temos

$$(1 - B)(\psi_0 + \psi_1 B + \psi_2 B^2 + \ldots)a_t = (1 - \theta B)a_t$$

$$\psi_0 + (\psi_1 - \psi_0)B + (\psi_2 - \psi_1)B^2 + (\psi_3 - \psi_2)B^3 \ldots = (1 - \theta B)$$

Então igualando os polinômios temos:

$\psi_0 = 1; \psi_1 - 1 = -\theta \Rightarrow \psi_1 = 1 - \theta;\ \ \psi_2 - \psi_1 = 0 \Rightarrow \psi_2 = \psi_1 = 1 - \theta$

ARIMA(0,1,1)

$\psi_0 = 1; \psi_1 - 1 = -\theta \Rightarrow \psi_1 = 1 - \theta; \;\; \psi_2 - \psi_1 = 0 \Rightarrow \psi_2 = \psi_1 = 1 - \theta$

$$\begin{aligned} Var(e_T(l)) &= \sigma_a^2 \sum_{j=0}^{l-1} \psi_j^2 \\ Var(e_n(1)) &= \sigma_a^2 \\ Var(e_n(2)) &= \sigma_a^2(1 + \psi_1^2) = \sigma_a^2(1 + (1 - \theta)^2) \end{aligned}$$

Checando os erros de previsão são
$e_T(1) = Z_{T+1} - \hat{Z}_T(1) = \theta_0 + Z_T + a_{T+1} - \theta a_T - [\theta_0 + Z_T - \theta a_T] = a_{T+1}$
$e_T(2) = Z_{T+2} - \hat{Z}_T(2) = \theta_0 + \hat{Z}_T(1) + a_{T+2} - \theta a_{T+1} - [2\theta_0 + Z_T - \theta a_T] =$
$\theta_0 + [\theta_0 + Z_T - \theta a_T] + a_{T+2} - \theta a_{T+1} - [2\theta_0 + Z_T - \theta a_T] = a_{T+2} + (1 - \theta) a_{T+1}$
Agora calcule as variâncias desses erros para checar com as obtidas.

Transformações

Calculei $Y_t = g(Z_t)$
Propomos modelo para Z_t e calculamos sua previsão $\hat{Z}_T(h)$ para depois calcular as previsões para Y_{t+h}

$$\hat{Y}_T(h) = g(\hat{Z}_t(h)) \Rightarrow \hat{Z}_t(h) = g^{-1}(\hat{Y}_T(h))$$

ex: $Y_t = ln(Z_t)$ e calculamos $\hat{Z}_t(h) = e^{\hat{Y}_T(h)}$
Granger e Newbold (1976) (MT2006) mostram que se Y_t for gaussiano, a previsão ótima é $e^{\hat{Y}_T(h)+\frac{1}{2}Var(e_T(h))}$. Se usar só $e^{\hat{Y}_T(h)}$, a previsão é viesada e com maior erro quadrático médio.
Se $Y_t = lnZ_t \sim ARIMA$, tenho $Y_{t+h}|Y_T \sim N(\hat{Y}_T(h), Var(e_T(h)))$

$$IC = [\hat{Y}_T(h) \mp z\sqrt{\widehat{Var}(e_T(h))}]$$

IC para $Z_{t+h} = [e^{\hat{Y}_T(h)-z\sqrt{\widehat{Var}(e_T(h))}}, e^{\hat{Y}_T(h)+z\sqrt{\widehat{Var}(e_T(h))}}]$
$\widehat{Var}$ porque substitui σ^2 por sua estimativa

Referências

All Time series analysis

- Morettin e Toloi
- Shumway and Stoffer
- Wei item Hyndman. Forecasting