

Homework 2 (Modelos ARMA)

Marcelo Santos Carielo

jul/2023

- Questão 1
- Resposta 1
- Questão 2
- Resposta 2
- Questão 3
- Resposta 3
- Referências

Questão 1

Suponha que a série diária de log-retorno de um ativo siga o seguinte modelo: $y_t = 0.01 + 0.2 y_{t-2} + \epsilon_t$, onde ϵ_t é um ruído branco Gaussiano com média zero e variância 0.02. (a) Determine a média e a variância da série de retornos y_t . (b) Calcule as autocorrelações de ordem 1 e 2 de y_t . (c) Assuma que o valor de y em $t = 100$, y_{100} , seja -0.01 e o valor de y em $t = 99$, y_{99} , seja 0.02. Calcule a previsão um passo à frente da série de retornos a partir da origem $t = 100$. Calcule também a previsão dois passos à frente a partir de $t = 100$.

Resposta 1

Para simplificar as manipulações algébricas, iremos usar a seguinte notação para a série y_t definida neste exercício:

$$y_t = 0.01 + 0.2 y_{t-2} + \epsilon_t = \alpha + \phi y_{t-2} + \epsilon_t,$$

onde $\alpha = 0.01$, $\phi = 0.2$ e ϵ_t é um ruído branco com média zero e variância $\sigma^2 = 0.02$.

(a) Cálculo da média $E(y_t)$:

Usando as propriedades do valor esperado E , a média de y_t é dada por

$$\begin{aligned} E\{y_t\} &= E\{\alpha + \phi y_{t-2} + \epsilon_t\} \\ &= E\{\alpha\} + \phi E\{y_{t-2}\} + E\{\epsilon_t\}, \text{ pela linearidade de } E \\ &= \alpha + \phi E\{y_{t-2}\} + 0, \text{ pois } \epsilon_t \text{ possui média zero} \end{aligned}$$

Usando a estacionaridade de y_t , temos que $E\{y_t\} = E\{y_{t-2}\} = \mu$, onde μ é uma constante. Substituindo na expressão acima temos

$$\mu = \alpha + \phi \mu$$

Logo, a média μ de y_t é

$$\mu = \frac{\alpha}{1 - \phi} = \frac{0.01}{1 - 0.02} = \frac{0.01}{0.98} = 0.010204.$$

Portanto, $\mu = 0.010204$.

Cálculo da variância $var(y_t)$:

$$\begin{aligned} var(y_t) &= var(\alpha + \phi y_{t-2} + \epsilon_t), \text{ usando da definição de } y_t \\ &= \phi^2 var(y_{t-2}) + var(\epsilon_t) + 2 cov(y_{t-2}, \epsilon_t), \text{ pelas propriedades da variância} \\ &= \phi^2 var(y_{t-2}) + \delta^2 + 0, \text{ pois } var(\epsilon_t) = \delta^2 \text{ e } y_t \text{ e } \epsilon_t \text{ não são correlacionados} \end{aligned}$$

Logo,

$$var(y_t) = \frac{\delta^2}{1 - \phi^2}.$$

Substituindo os valores das constantes na expressão acima, temos que

$$\begin{aligned} var(y_t) &= \frac{0.02}{1 - (0.2)^2} = \frac{0.02}{1 - 4 * 10^{-2}} \\ &= \frac{0.02}{0.96} = 0.021 \end{aligned}$$

Portanto, $var(y_t) = 0.021$. Como $var(y_t) = cov(y_t, y_t) = \gamma_0$, então $\gamma_0 = 0.021$.

(b) Inicialmente, vamos calcular uma expressão para a autocovariância de lag k , denotada por γ_k :

$$\begin{aligned} \gamma_k &= cov(y_t, y_{t-k}) \\ &= E(y_t y_{t-k}) - \mu^2, \text{ pois } E(y_t) = E(y_{t-k}) = \mu \text{ 'média constante de } y_t' \\ &= E[(\alpha + \phi y_{t-2} + \epsilon_t) y_{t-k}] - \mu^2, \text{ já que } y_t = \alpha + \phi y_{t-2} + \epsilon_t \\ &= E(\alpha y_{t-k} + \phi y_{t-2} y_{t-k} + \epsilon_t y_{t-k}) - \mu^2, \text{ após multiplicarmos} \\ &= \alpha E(y_{t-k}) + \phi E(y_{t-2} y_{t-k}) + E(\epsilon_t y_{t-k}) - \mu^2, \text{ pois } E(aX + bY) = aE(X) + bE(Y), \forall a, b \text{ constantes e } \forall X, Y \text{ variáveis aleatórias} \\ &= \alpha \mu + \phi E(y_{t-2} y_{t-k}) + 0 - \mu^2, \text{ pois série } y_t \text{ e o erro } \epsilon_t \text{ não são correlacionados} \\ &= \alpha \mu + \phi E(y_{t-2} y_{t-k}) - \mu^2 + \phi \mu^2 - \phi \mu^2, \text{ após somar e subtrair um termo que desejamos} \\ &= \alpha \mu + \phi [E(y_{t-2} y_{t-k}) - \mu^2] - \mu^2 + \phi \mu^2, \text{ agrupando os termos} \\ &= \alpha \mu + \phi \gamma_{k-2} - \mu^2 + \phi \mu^2, \text{ pois } E(y_{t-2} y_{t-k}) - \mu^2 = \gamma_{k-2} \text{ (adiante verificaremos essa identidade)} \end{aligned}$$

Portanto,

$$\gamma_k = \alpha \mu + \phi \gamma_{k-2} - \mu^2 + \phi \mu^2. (*)$$

Usando $k = 1$ na expressão acima, temos

$$\gamma_1 = \alpha \mu + \phi \gamma_{-1} - \mu^2 + \phi \mu^2.$$

Como y_t é estacionária, vale que $\gamma_k = \gamma_{-k}$. Então,

$$\gamma_1 = \alpha \mu + \phi \gamma_1 - \mu^2 + \phi \mu^2 = \alpha \mu + \phi \gamma_1 - \mu^2(1 - \phi).$$

Isolando γ_1 , ficamos com

$$\gamma_1 = \frac{\alpha \mu}{1 - \phi} - \mu^2 = 0.$$

Então, substituindo os valores das constantes α, μ, ϕ , a correlação de lag 1 é:

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{0}{\gamma_0} = 0.$$

Portanto, a autocorrelação de ordem 1 é $\rho_1 = 0$.

Usando (*), obtemos a autocovariância de ordem 2, γ_2 :

$$\begin{aligned} \gamma_2 &= \alpha \mu + \phi \gamma_0 + \mu^2 (\phi - 1) \\ &= 0.01 * 0.0125 + 0.2 * 0.021 + (0.0125)^2 * (0.2 - 1) \\ &= 1.25 * 10^{-4} + 4.2 * 10^{-3} - 1.25 * 10^{-4} \\ &= 4.2 * 10^{-3} \end{aligned}$$

Logo, a autocorrelação de ordem 2 é

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \frac{4.2 * 10^{-3}}{0.021} = 0.2.$$

Portanto, $\rho_2 = 0.2$.

Para verificar que $E(y_{t-2} y_{t-k}) - \mu^2 = \gamma_{k-2}$, podemos fazer a substituição de variáveis $t' = t - 2$ (ou seja, $t = t' + 2$), então

$$\begin{aligned} E(y_{t-2} y_{t-k}) - \mu^2 &= \text{cov}(y_{t-2}, y_{t-k}), \text{ pois } y_t \text{ é estacionária e tem média constante} \\ &= \text{cov}(y_{t'}, y_{(t'+2)-k}), \text{ substituindo } t = t' + 2 \\ &= \text{cov}(y_{t'}, y_{t'-(k-2)}), \text{ agrupando os termos} \\ &= \text{cov}(y_t, y_{t-(k-2)}), \text{ pois } t \text{ é um rótulo arbitrário para o subíndice e pode ser alterado} \\ &= \gamma_{k-2} \end{aligned}$$

Questão 2

Descreva como as FAC e FACP são úteis para identificar um modelo ARMA. Como os critérios de informação podem ser usados na construção de um modelo?

Resposta 2

Um modelo ARMA(p, q) pode ser escrito como:

$$y_t = \alpha + \sum_{j=1}^p \phi_j y_{t-j} + \sum_{i=0}^q \theta_i \epsilon_{t-i},$$

onde os coeficientes e variáveis são conforme vistos em aula. Dessa maneira, o valor atual da série y_t para um modelo AR(p,q), pode depender do intercepto (drift) α , dos valores passados da série y_k para $k < t$ e do erro atual ϵ_t (assumido como ruído branco). Para medir a correlação entre as variáveis atuais e passadas utilizamos a função de autocorrelação (abreviada por ACF, em inglês) e a função de autocorrelação parcial (PACF, em inglês). A ACF mede a correlação entre duas variáveis (atual e passado) e a PACF mede a correlação entre duas variáveis eliminando o efeito de outras variáveis (uma forma de se fazer isso, é aplicando uma regressão linear).

Conforme vimos ao longo do curso, a ACF de uma série temporal y_t é formada por $\rho_1, \rho_2, \rho_k, \dots, \rho_k$ onde $\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)} = \frac{\gamma_k}{\gamma_0}$. Além disso, a PACF é dada por $\phi_1, \phi_2, \dots, \phi_k$ em que ϕ_k satisfaz a equação $y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_k y_{t-k} + v_t$.

Adicionalmente, as propriedades dos processos AR(p) e MA(q) podem ser usados na identificação de um modelo ARMA(p, q). De fato, um processo AR(p) é descrito por: (i) uma ACF infinita e declinante; e (ii) uma PACF aproximadamente igual a zero (ou truncada) para lags maiores que p. Já um processo MA(q) é descrito por: (i) uma ACF aproximadamente igual a zero (ou truncada) para lags maiores que q; e (ii) uma PACF infinita e declinante. Analisando o gráfico da ACP e FACP podemos "estimar" visualmente os valores de p e q num modelo ARMA.

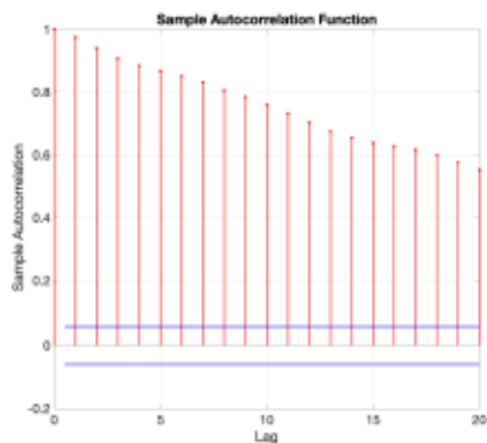
Uma propriedade interessante é que $\text{MA}(\infty) = \text{AR}(1)$ e $\text{MA}(1) = \text{AR}(\infty)$.

Conforme vimos em aula, modelos AR podem ser estimados por Mínimos Quadrados Ordinários (MQO). Já os modelos MA e ARMA são estimados através do método da Máxima Verossimilhança (MV), pois os erros que atuam como variáveis independentes nesses modelos não são diretamente observáveis.

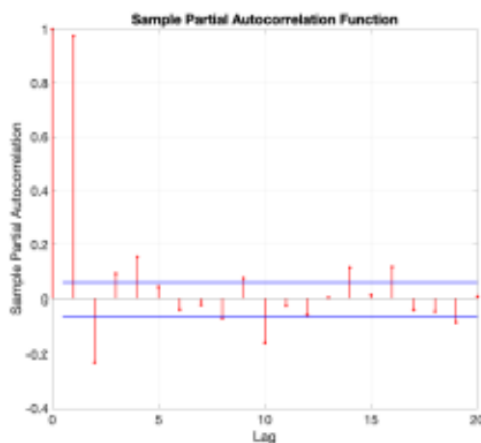
Para usarmos os critérios de informação Akaike (AIC) e Bayesiano (BIC), calculamos o AIC (ou BIC) para $k = 0, \dots, n$, onde n para um inteiro positivo. O valor da ordem k que possui o valor AIC (ou BIC) mínimo indica a ordem k que deve ser escolhida para termos o melhor modelo.

Questão 3

Seja a FAC e FACP (ACF e PACF em inglês, respectivamente) de uma série de log-retornos mostrada na página seguinte. Qual seria um modelo do tipo ARMA razoável para descrever a dinâmica dessa série? Como você poderia verificar se o modelo escolhido é adequado? NOTA: As FAC e FACP mostradas abaixo, começam no lag = zero. Então, o primeiro elemento é a autocorrelação de ordem zero, que deve ser desconsiderada para a análise proposta.



(m) ACF



(n) PACF

fonte: Homework 2 - Questão 3.

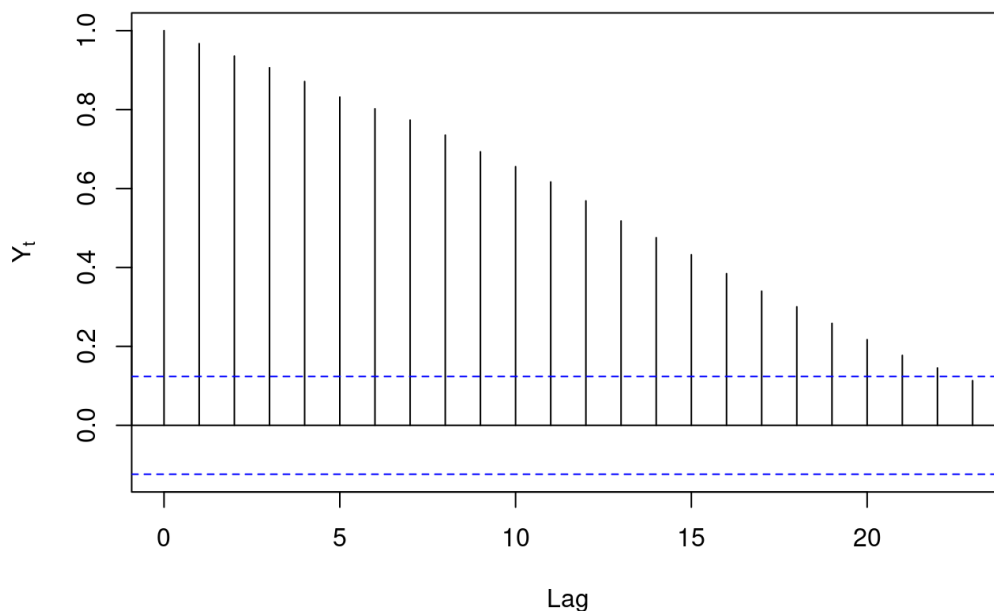
Resposta 3

Olhando para o gráfico da PACF notamos um alto valor para o lag = 1, sugerindo que o modelo ARMA possui uma parte AR(1). Enquanto isso o decaimento indicado no gráfico da ACF, sugere uma parte MA(∞).

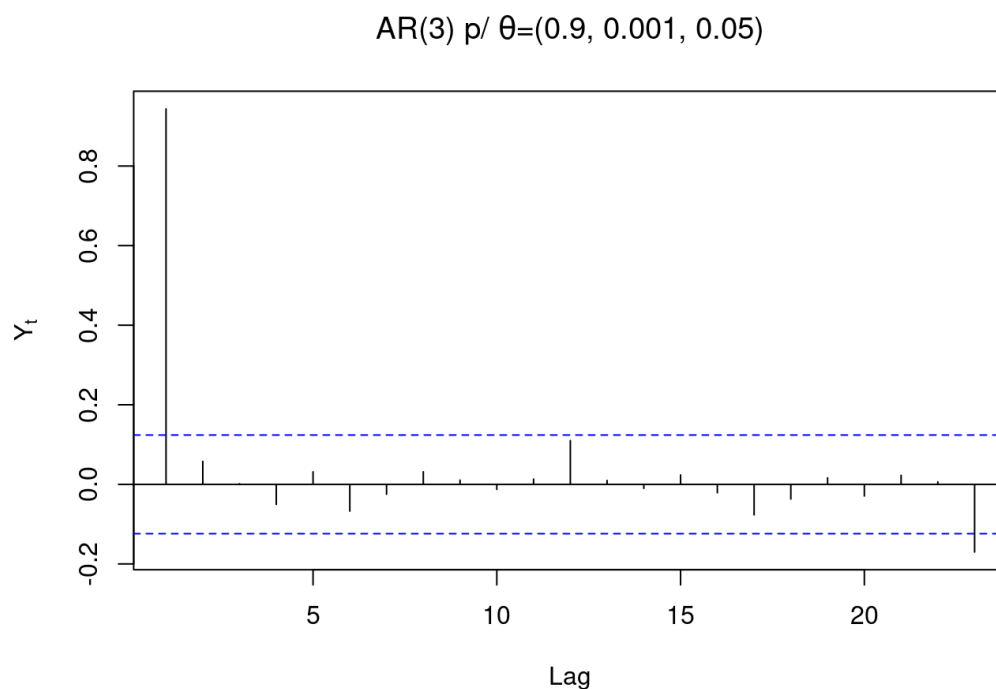
Vejamos alguns testes computacionais antes de decidir qual ordem escolheremos para nosso modelo ARMA.

```
TT <- 250
#?arima.sim
ma <- acf(arima.sim(model=list(ma=seq(from=1, to=10000, by = 1)), n=TT),
  ylab=expression(Y[t]),
  main=expression(paste("MA(10000) p/ ", theta, "=1")))
```

MA(10000) p/ $\theta=1$



```
ar <- pacf(arima.sim(model=list(ar=c(0.9, 0.001, 0.05)), n=TT),
  ylab=expression(Y[t]),
  main=expression(paste("AR(3) p/ ", theta, "=(0.9, 0.001, 0.05)")))
```



Os testes computacionais sugerem que nossa o modelo é formado por $MA(\infty)$ e $AR(1)$. Como $MA(\infty) = AR(1)$, então escolhemos o modelo $ARMA(1, \infty)$ que equivale a um $ARMA(1, 0)$.

Para saber se o modelo $ARMA(1, 0)$ proposto está adequado, podemos analisar os seus resíduos. Se os resíduos tiverem comportamento de ruído branco, ou seja, média zero, variância constante e ausência de autocorrelação, então o modelo proposto foi adequado. Podemos fazer isso através da análise das FAC e FACP dos resíduos, além dos testes de Ljung-Box e Box-Pierce. No caso de haver autocorrelação na série dos resíduos, devemos revisar o modelo proposto.

Referências

- Materiais das aulas (profa. Andreza Palma)
- CAP. 2 do livro "TSAY, Ruey S. *An introduction to analysis of financial data with R*. John Wiley & Sons, 2014."