Econometria de Séries Temporais

Identificação e seleção de modelos

João Ricardo Costa Filho

“The most important questions of life are, for the most part, really only problems in probability."

Laplace (1812)

“In God we trust. All others must bring data."

William Edwards Deming

Como escolher a ordem de um ARMA(p,q)?

A metodologia Box-Jenkins

A metodologia Box-Jenkins

1) Estacionariedade

A metodologia Box-Jenkins

1) Estacionariedade
2) Identificar a ordem do modelo

A metodologia Box-Jenkins

1) Estacionariedade
2) Identificar a ordem do modelo

- FAC e FACP

A metodologia Box-Jenkins

1) Estacionariedade
2) Identificar a ordem do modelo

- FAC e FACP
- Criérios de informação

A metodologia Box-Jenkins

1) Estacionariedade
2) Identificar a ordem do modelo

- FAC e FACP
- Criérios de informação

3) Estimar os parâmetros do modelo

A metodologia Box-Jenkins

1) Estacionariedade
2) Identificar a ordem do modelo

- FAC e FACP
- Criérios de informação

3) Estimar os parâmetros do modelo
4) Diagnóstico dos resíduos

A metodologia Box-Jenkins

1) Estacionariedade
2) Identificar a ordem do modelo

- FAC e FACP
- Criérios de informação

3) Estimar os parâmetros do modelo
4) Diagnóstico dos resíduos
5) Projeções

(1) Estacionariedade

(1) Estacionariedade

Se for necessário (e fizer sentido), ela pode ser induzida:

(1) Estacionariedade

Se for necessário (e fizer sentido), ela pode ser induzida:

- Log, raiz quadrada, diferenciação.
  - Tranformação Box-Cox (Box and Cox 1964):

$$Y_t^{(\lambda)} = \begin{cases} \frac{Y_t^{\lambda}-1}{\lambda} & \text{for } \lambda \neq 0 \\ \log(Y_t) & \text{for } \lambda = 0 \end{cases}$$

(2) Identificar a ordem do ARMA(p,q)

Seja y_t uma variável aleatória tal que o seu processo estocástico é dado por:

(2) Identificar a ordem do ARMA(p,q)

Seja y_t uma variável aleatória tal que o seu processo estocástico é dado por:

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}.$$

onde ε_t é um ruído branco.

(2) Identificar a ordem do ARMA(p,q)

Seja y_t uma variável aleatória tal que o seu processo estocástico é dado por:

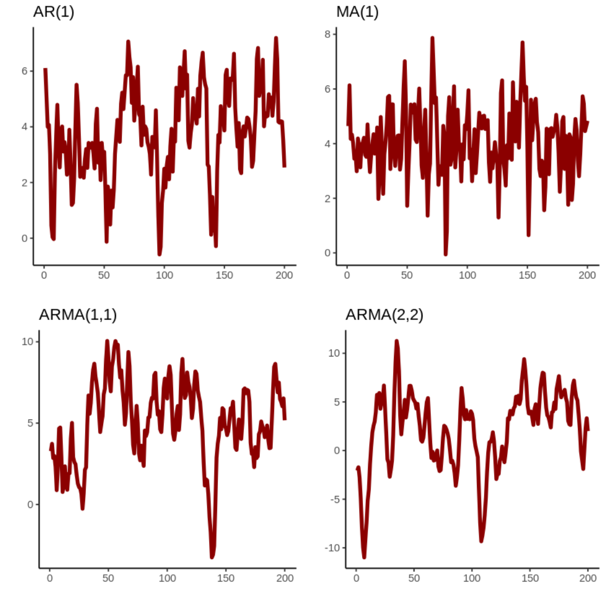
$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \theta_j \varepsilon_{t-j}.$$

onde ε_t é um ruído branco. A ordem do processo impacta muito o "formato" do gráfico?

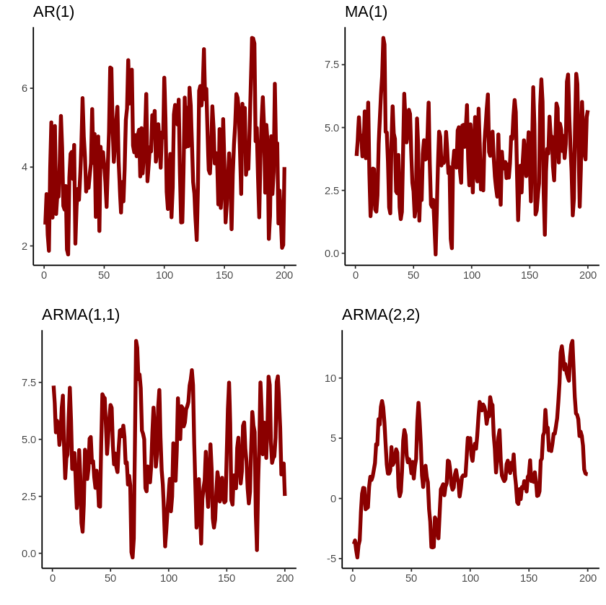
(2) Identificar a ordem do ARMA(p,q)

Vamos simular quatro processos: AR(1), MA(1), ARMA(1,1) e ARMA(2,2).

ARMA(p,q) – Simulação 1



ARMA(p,q) – Simulação 2



(2) Identificar a ordem do ARMA(p,q)

Sem os títulos dos gráficos, vocês conseguiriam identificar qual é a ordem de cada processo estocásticos?

(2) Identificar a ordem do ARMA(p,q): funções de autocorrelação

Valores defasados de y_t (como $y_{t-1}, y_{t-2}, \ldots$) podem influenciar y_t de maneira **direta** e/ou **indireta**.

(2) Identificar a ordem do ARMA(p,q): funções de autocorrelação

Valores defasados de y_t (como $y_{t-1}, y_{t-2}, \ldots$) podem influenciar y_t de maneira **direta** e/ou **indireta**.

- Exemplo: Como y_{t-2} influencia y_t?

(2) Identificar a ordem do ARMA(p,q): funções de autocorrelação

Valores defasados de y_t (como $y_{t-1}, y_{t-2}, \ldots$) podem influenciar y_t de maneira **direta** e/ou **indireta**.

- Exemplo: Como y_{t-2} influencia y_t?
  - **Indiretamente**: $y_{t-2} \to y_{t-1} \to y_t$

(2) Identificar a ordem do ARMA(p,q): funções de autocorrelação

Valores defasados de y_t (como $y_{t-1}, y_{t-2}, \ldots$) podem influenciar y_t de maneira **direta** e/ou **indireta**.

- Exemplo: Como y_{t-2} influencia y_t?
  - **Indiretamente**: $y_{t-2} \to y_{t-1} \to y_t$
  - **Diretamente**: $y_{t-2} \to y_t$

(2) Identificar a ordem do ARMA(p,q): funções de autocorrelação

Valores defasados de y_t (como $y_{t-1}, y_{t-2}, \ldots$) podem influenciar y_t de maneira **direta** e/ou **indireta**.

- Exemplo: Como y_{t-2} influencia y_t?
  - **Indiretamente**: $y_{t-2} \rightarrow y_{t-1} \rightarrow y_t$
  - **Diretamente**: $y_{t-2} \rightarrow y_t$

Temos duas formas de captar esses efeitos: as funções de **autocorrelação** e de **autocorrelação parcial**.

Autocorrelação (FAC)

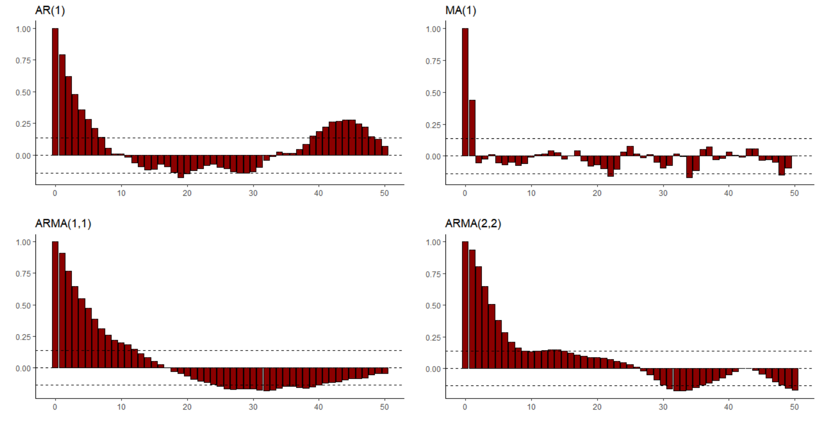
- **Função de autocorrelação**: ordene y_t e y_{t-2} lado-a-lado e calcule a correlação linear.

Autocorrelação (FAC)

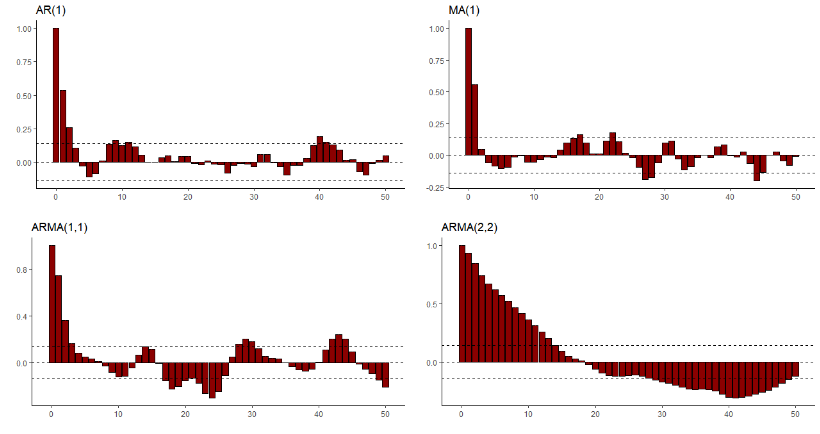
- **Função de autocorrelação**: ordene y_t e y_{t-2} lado-a-lado e calcule a correlação linear. Momentos teóricos (Bueno 2012):

Modelo	FAC
$MA(q)$	$\rho_j = \frac{\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q\theta_{q-j}}{\sum_{j=0}^{q}\theta_j^2}, \quad j = 1, 2, \ldots, q$
$AR(1)$	$\rho_j = \phi^j, \quad j = 1, 2, \ldots$
$AR(p)$	$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \cdots + \phi_p\rho_{j-p}, \quad j = 1, 2, \ldots$
ARMA $(1, 1)$	$\begin{cases} \rho_1 = \frac{(1+\phi_1\theta_1)(\phi_1+\theta_1)}{1+\theta_1^2+2\phi_1\theta_1} \\ \rho_j = \phi_1\rho_{j-1} = \phi_1^{j-1}\rho_1, \quad j > 1. \end{cases}$

Autocorrelação (FAC) – Simulação 1



Autocorrelação (FAC) – Simulação 2



Autocorrelação parcial (FACP)

- **Função de autocorrelação parcial**: como calcular efeitos parciais controlando por outros fatores?

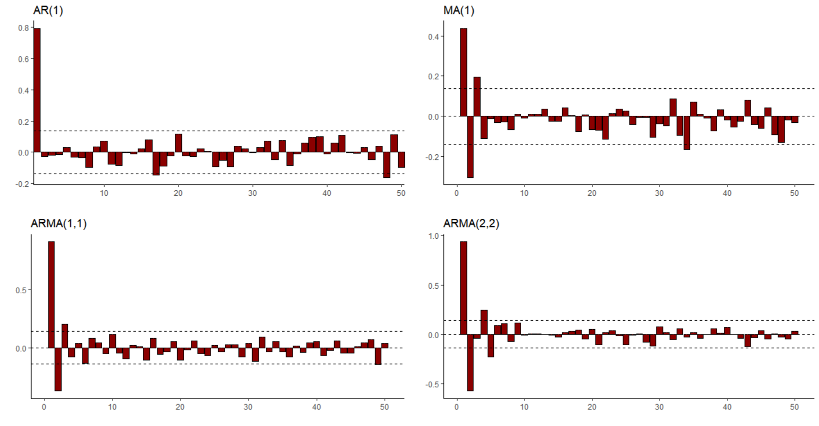
Autocorrelação parcial (FACP)

- **Função de autocorrelação parcial**: como calcular efeitos parciais controlando por outros fatores? Estime uma regressão!

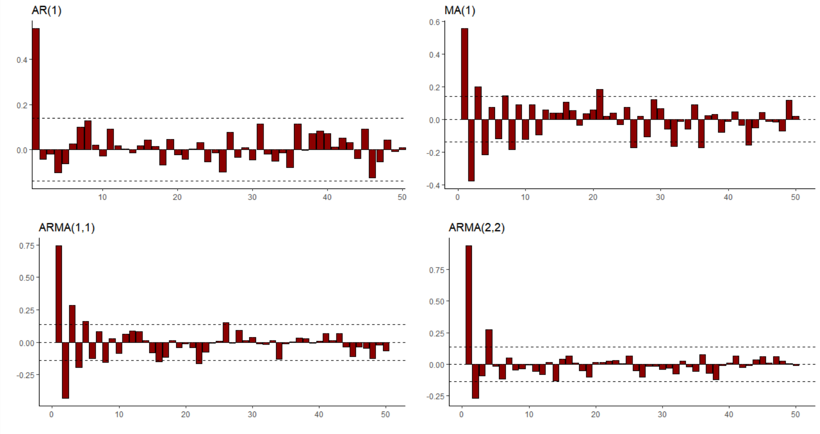
Autocorrelação parcial (FACP)

- **Função de autocorrelação parcial**: como calcular efeitos parciais controlando por outros fatores? Estime uma regressão!
  - $y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + \epsilon_t$

Autocorrelação (FACP) – Simulação 1



Autocorrelação (FACP) – Simulação 2



Seleção de defasagens (Bueno 2012)

Seleção de defasagens (Bueno 2012)

- A FAC define a defasagem do componente MA.

Seleção de defasagens (Bueno 2012)

- A FAC define a defasagem do componente MA.
- A FACP define a defasagem do componente AR.

Seleção de defasagens (Bueno 2012)

- A FAC define a defasagem do componente MA.
- A FACP define a defasagem do componente AR.

Modelo	FAC	FACP
$AR(p)$	Decai	Truncada na defasagem p
$MA(q)$	Truncada na defasagem q	Decai
$ARMA(p, q)$	Decai se $j > q$	Decai se $j > p$

Teste de Ljung-Box

Teste de Ljung-Box

Uma alternativa é verificar, estatísticamente, se a série possui alguma autocorrelação:

Teste de Ljung-Box

Uma alternativa é verificar, estatísticamente, se a série possui alguma autocorrelação:

$$\mathcal{H}_0 : \sum_{j=1}^{n} \rho_j = 0 \mathcal{H}_a : \sum_{j=1}^{n} \rho_j \neq 0$$

$$Q = T(T+2) \sum_{j=1}^{n} \frac{\widehat{\rho}_j^2}{T-j} \xrightarrow{d} \chi_n^2,$$

(2) Identificar a ordem do ARMA(p,q): critérios de informação

Como encontrar o número ideal de parâmetros de um modelo?

(2) Identificar a ordem do ARMA(p,q): critérios de informação

Como encontrar o número ideal de parâmetros de um modelo?
"Penalizando" a inclusão de novos parâmetros.

(2) Identificar a ordem do ARMA(p,q): critérios de informação

Como encontrar o número ideal de parâmetros de um modelo? "Penalizando" a inclusão de novos parâmetros.

Genericamente, as estatísticas utilizadas para esse fim possuem a seguinte forma funcional (Bueno 2012):

(2) Identificar a ordem do ARMA(p,q): critérios de informação

Como encontrar o número ideal de parâmetros de um modelo? "Penalizando" a inclusão de novos parâmetros.

Genericamente, as estatísticas utilizadas para esse fim possuem a seguinte forma funcional (Bueno 2012):

$$C = \ln \widehat{\sigma}^2(T) + c_T \varphi(T),$$

onde $\widehat{\sigma}^2(T) = T^{-1} \sum_{t=1}^{T} \widehat{\varepsilon}_t^2$ é a variância dos resíduos, c_T representa o número de parâmetros estimados e $\varphi(T)$ é a ordem do processo.

Critérios de informação

Critérios de informação

- **Akaike (AIC)**
  - $\text{AIC}(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T}$.

Critérios de informação

- **Akaike (AIC)**
  - $\text{AIC}(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T}$.
- **Hannan-Quinn (HQ)**
  - $HQ(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T} \ln \ln T$.

Critérios de informação

- **Akaike (AIC)**
  - $\mathrm{AIC}(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T}$.
- **Hannan-Quinn (HQ)**
  - $HQ(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T} \ln \ln T$.
- **Schwarz (BIC)**
  - $BIC(p,q) = \ln \widehat{\sigma}^2 + n\frac{\ln T}{T}$.

Critérios de informação

- **Akaike (AIC)**
  - $\text{AIC}(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T}$.
- **Hannan-Quinn (HQ)**
  - $HQ(p,q) = \ln \widehat{\sigma}^2 + n\frac{2}{T} \ln \ln T$.
- **Schwarz (BIC)**
  - $BIC(p,q) = \ln \widehat{\sigma}^2 + n\frac{\ln T}{T}$.

onde T é o número de observações é $n = p + q$ (ou $n = p + q + 1$ se há constante).

Leia os livros e os artigos, não fique só com os slides!!!!

Referências

Box, George EP, and David R Cox. 1964. "An Analysis of Transformations." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 26 (2): 211–43.

Bueno, Rodrigo De Losso da Silveira. 2012. *Econometria de Séries Temporais*. Cengage Learning.