

Metodologia Box-Jenkins

Bibliografia Básica:

- Enders, W. *Applied Econometric Time Series*. Cap. 2.
- Bueno, R. L. S. *Econometria de Séries Temporais*. Cap. 2 e 3.
- Box, G.E. & Jenkins, G. M.. *Time series analysis: forecasting and control*. Cap. 3.
- Morettin, P. A. *Análise de Séries Temporais*. Cap. 2 e 5.

Um processo MA(q) pode ser escrito como:

$$y_t = \mu + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \dots - \theta_q\epsilon_{t-q} \quad (1)$$

Condição de Invertibilidade

A condição que garante a invertibilidade de um processo MA(q) é que o inverso das raízes do polinômio estejam fora (dentro) do círculo unitário.

Revisão: Processos AutoRegressivos

Um processo $AR(p)$, pode ser escrito como:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t \quad (2)$$

ou:

$$(1 - \phi_1 L - \phi_2 L^2 - \phi_3 L^3 - \dots - \phi_p L^p) y_t = \mu + \epsilon_t \quad (3)$$

Condição de Estacionariedade

O processo $AR(p)$ é estacionário se todas as raízes do polinômio resultante têm um módulo maior que 1, ou seja, se as raízes estiverem fora do círculo unitário.

Função de Autocorrelação (FAC ou ACF)

Ajuda a caracterizar o desenvolvimento de y_t ao longo do tempo.

Assumindo estacionaridade fraca, podemos definir a k -ésima ordem de autocovariância, γ_k , como:

$$\gamma_k = \text{cov}\{y_t, y_{t-k}\} = \text{cov}\{y_t, y_{t+k}\}$$

A autocovariância de um processo estocástico pode ser normalizada e apresentada como uma função de autocorrelação, ρ_k :

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Função de Autocorrelação

A ACF de um processo AR(1) é dado por:

$$\rho_k = \frac{\text{cov}\{y_t, y_{t-1}\}}{\text{Var}(y_t)} = \phi^k$$

Observação: em um modelo AR a função ACF é infinita e com decrescimento exponencial

Para o processo MA(1) a ACF é:

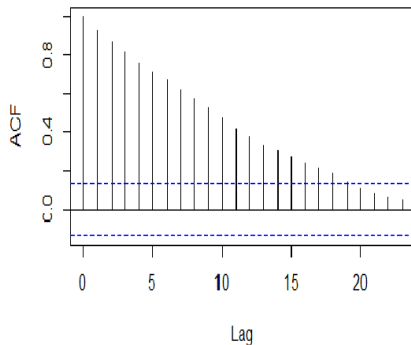
$$\rho_1 = \frac{\theta}{1 + \theta^2}$$

e $\rho_k = 0, k > 1$.

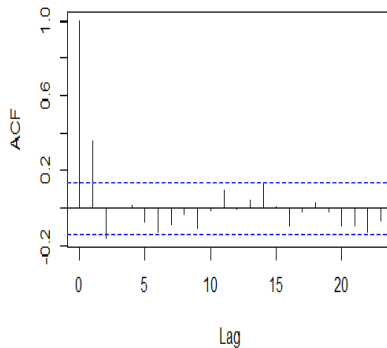
Observação: em um modelo MA a função ACF é finita, com valores iguais a zero para defasagens maiores que q .

Exemplos - ACF Autoregressivos

AR(1) com $\phi=0.9$



MA(1) com $\theta=0.9$



Função de Autocorrelação Parcial (FACP ou PACF)

Fonece a correlação entre a variável no instante t e uma de suas defasagens, retirado os efeitos das outras defasagens.

Processo AR(p): coeficientes de correlação parcial diferentes de zero até a ordem p e estatisticamente iguais a zero a partir de então.

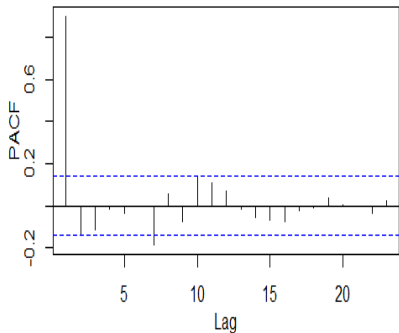
Processo MA(q): coeficientes de correlação parcial decaem.

Para um número de observações N , ϕ_{kk} tem distribuição aproximadamente normal, o que permite a construção de um intervalo de confiança para ϕ_{kk} , considerando que a estimativa é diferente de zero quando

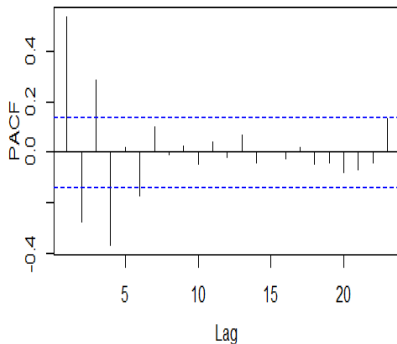
$$|\hat{\phi}_{kk}| > 2/\sqrt{N}$$

Exemplos - PACF - Autoregressivo

AR(1) com $\phi=0.9$



MA(1) com $\theta=0.9$



Processos AutoRegressivos e de Média Móvel

ARMA(p,q) pode ser formulado como:

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} + \epsilon_t \quad (4)$$

Ou, escrevendo com os operadores de defasagem, tem-se:

$$(1 + \phi_1 L + \dots + \phi_p L^p) y_t = (1 + \phi_1 L + \dots + \phi_p L^p) \epsilon_t \quad (5)$$

Ou,

$$\Phi(L) y_t = \Theta(L) \epsilon_t$$

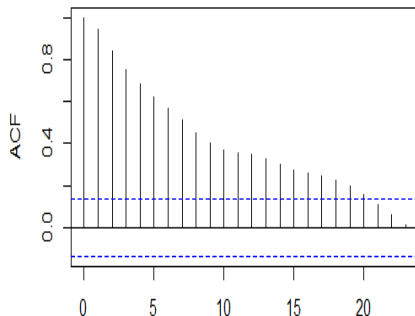
Os valores de ϕ_1, \dots, ϕ_p que tornam o processo estacionário são tais que as raízes de $\Phi(L)$ estão fora do círculo unitário.

Os valores de $\theta_1, \dots, \theta_q$ que tornam o processo inversível são tais que as raízes de $\Theta(L)$ estão fora do círculo unitário.

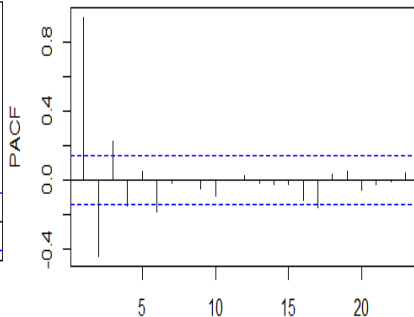
ACF e PACF de Processos ARMA

Para um processo ARMA(p,q) a função de autocorrelação tem um decaimento exponencial ou oscilatório após a defasagem q enquanto que a função de autocorrelação parcial tem o mesmo comportamento após a defasagem p (Box & Jenkins 1970, p. 79).

ARMA(1,1) com $\phi=0.9$ e $\theta=0.9$



ARMA(1,1) com $\phi=0.9$ e $\theta=0.9$



Construção do modelo de séries temporais:

- 1 **Identificação:** com base na análise de autocorrelação, autocorrelação parcial e/ou critérios de informação;
- 2 **Estimação:** os parâmetros do modelo identificado são estimados;
- 3 **Verificação do modelo ajustado:** por meio de uma análise de resíduos, averigua-se se este é adequado para os fins em vista, no caso, para a previsão.
- 4 **Previsão.**

Identificação

Em geral, a identificação é baseada no comportamento da FAC e FACP empíricas.

O primeiro passo na etapa de identificação é a verificação da propriedade de estacionariedade da série a $\{y_t\}$.

A metodologia de Box-Jenkins é restrita à séries estacionárias, porém, é possível aplicá-la à séries que se tornam estacionárias após a aplicação de diferenças.

Nestes casos, a série $\{y_t\}$ é dita diferença estacionária, e a ordem d é o número de raízes unitárias existentes no processo estocástico.

Identificação

Processo	FAC	FACP
AR(p)	declinante	truncada em $j = p$
MA(q)	truncada em $j = q$	declinante
ARMA(p,q)	declinante	declinante

Na prática, FAC e FACP exigirão a análise da significância estatística das autocorrelações.

Em geral, haverá mais de um modelo candidato a gerador da série.

Na prática, ajustamos um conjunto de modelos candidatos e utilizamos critérios de informação (ou de seleção) para selecionar o modelo mais adequado.

Modelos parcimoniosos são preferidos a modelos de ordens elevadas.

Identificação

Os critérios de informação mais utilizados são: Akaike (1981) e Schwartz (1978):

$$AIC(p, q) = \ln \sigma^2 + n \frac{2}{N}$$

$$BIC(p, q) = \ln \sigma^2 + n \frac{\ln N}{N}$$

em que $n = p + q$ e N o tamanho da amostra.

Sobre os critérios de informação:

1. não devem ser utilizados para comparar modelos com tamanhos de amostras diferentes;
2. o critério BIC tende a escolher um modelo mais parcimonioso do que o AIC;
3. o critério AIC é melhor quando a amostra é pequena, porém, tende a indicar modelos sobreparametrizados;

Uma vez determinados os valores p e q , o próximo passo é a estimação dos p parâmetros ϕ , dos q parâmetros θ e da variância σ_ϵ^2 .

1. Modelos puramente autoregressivos podem ser estimados de maneira consistente por OLS e por máxima verossimilhança.
2. Modelos puramente caracterizados por médias móveis podem ser estimados de maneira consistente por máxima verossimilhança.
3. Modelos ARMA podem ser estimados de maneira consistente por máxima verossimilhança.

Objetivo:

Avaliar adequação do modelo identificado/estimado.

Procedimento:

- a) Analisar a significância dos parâmetros estimados;
- b) Análise da correlação dos resíduos: modelo bem especificado, resíduos seguem um RB
- c) Critério de desempate:
 - Escolher modelo com menores: σ^2 , AIC , BIC

Utiliza-se a estatística Q de Box-Pierce (1970) ou a modificada por Ljung-Box (1978) para testar se os primeiros k coeficientes de autocorrelação amostrais são estatisticamente iguais a zero:

$$H_0 : \sum_{j=1}^K \rho_j = 0 \quad H_1 : \sum_{j=1}^K \rho_j \neq 0$$

A estatística do teste de Box-Pierce é:

$$Q = N \sum_{j=1}^K \hat{\rho}_j^2 \xrightarrow{d} \chi_{K-p-q}^2$$

A estatística do teste de Ljung-Box é:

$$Q = N(N+2) \sum_{j=1}^K \frac{\hat{\rho}_j^2}{N-j} \xrightarrow{d} \chi_{K-p-q}^2$$

Verificação - Teste de Normalidade

Teste Jarque-Bera: usado para grandes amostras

Teste Jarque-Bera (JB) analisa se os momentos da série estimada (no caso os resíduos) são iguais aos da normal. Sob essa hipótese, a assimetria é igual a zero e a curtose é igual a 3. As hipóteses nula e alternativa são, respectivamente,

H_0 : os resíduos têm distribuição normal

H_1 : os resíduos não têm distribuição normal

A estatística do teste é:

$$JB = N \left[\frac{A^2}{6} - \frac{(C - 3)^2}{24} \right] \xrightarrow{d} \chi^2_2$$

Verificação - Teste de Normalidade

Teste de Shapiro-Wilk (1965): pode ser utilizado para amostras de qualquer tamanho.

O teste é baseado no cálculo da estatística W que verifica se uma amostra aleatória de tamanho n provém de uma distribuição normal:

$$W = \frac{b^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

em que x_i são valores da amostra ordenados e b é uma constante determinada a partir das médias, variâncias e covariâncias de uma amostra de tamanho N . As hipóteses nula e alternativa são:

H_0 : a amostra provém de uma População Normal

H_1 : a amostra não provém de uma População Normal

Teste ARCH-LM

Objetivo:

Análise de heterocedasticidade condicional.

Regredir a equação:

$$\hat{\epsilon}_t^2 = \beta_1 \hat{\epsilon}_{t-1}^2 + \beta_2 \hat{\epsilon}_{t-2}^2 + \dots + \beta_h \hat{\epsilon}_{t-h}^2 + u_t$$

Sob as hipóteses

$$H_0 = \beta_1 = \beta_2 = \dots = \beta_h = 0$$

$$H_1 = \beta_1 \neq 0, \text{ ou } \beta_2 \neq 0, \text{ ou } \dots \beta_h \neq 0$$

Estatística do teste:

$$ARCH - LM = T \cdot R^2 \xrightarrow{d} \chi_h^2$$

Não rejeitar H_0 significa ausência de heterocedasticidade.

Transformação de Box-Cox

Razões para transformar os dados: estabilizar a variância e tornar o efeito sazonal aditivo;

No caso de séries econômicas e financeiras poderá ser necessário aplicar à série original alguma transformação não-linear, como a logarítmica ou, em geral, uma transformação da forma:

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

chamada **transformação de Box-Cox** (1964). Os parâmetros λ e c devem ser estimados.

Para ilustrar o processo de previsão, considere um modelo AR(1):

$$y_{t+1} = c + \phi y_t + \epsilon_{t+1}$$

As previsões são feitas tomando-se as esperanças:

$$E(y_{t+1}) = c + \phi y_t$$

$$E(y_{t+2}) = c + \phi E(y_{t+1}) = c + \phi(c + \phi y_t)$$

$$\vdots$$

$$E(y_{t+h}) = c \sum_{i=1}^{h-1} \phi^{i-1} + \phi^h y_t$$

O erro de previsão $e_t(H)$, $H = 1, 2, \dots, h$ passos à frente, é dado por:

$$e_t(1) = y_{t+1} - E(y_{t+1}) = \epsilon_{t+1}$$

$$\begin{aligned} e_t(2) &= y_{t+2} - E(y_{t+2}) = c + \phi y_{t+1} + \epsilon_{t+2} - c - \phi E(y_{t+1}) \\ &= \phi \epsilon_{t+1} + \epsilon_{t+2} \end{aligned}$$

$$\begin{aligned} e_t(3) &= y_{t+3} - E(y_{t+3}) = c + \phi y_{t+2} + \epsilon_{t+3} - c - \phi E(y_{t+1}) \\ &= \epsilon_{t+3} + \phi \epsilon_{t+2} + \phi^2 \epsilon_{t+1} \end{aligned}$$

\vdots

$$e_t(h) = y_{t+h} - E(y_{t+h}) = \epsilon_{t+h} + \phi \epsilon_{t+h-1} + \dots + \phi^{h-1} \epsilon_{t+1}$$

A variância do erro de previsão é dada por:

$$\begin{aligned} \text{Var}(e_t(H)) &= \text{Var}(\epsilon_{t+h} + \phi\epsilon_{t+h-1} + \dots + \phi^{h-1}\epsilon_{t+1}) \\ &= \sigma_\epsilon^2 (1 + \phi^2 + \phi^4 + \dots + \phi^{2(h-1)}) \end{aligned}$$

A variância aumenta com o horizonte de previsão, porém a taxas decrescentes.

Quando $H \rightarrow \infty$, $\text{Var}(e_t(H)) \rightarrow \frac{\sigma_\epsilon^2}{1 - \phi^2}$.

O intervalo de confiança aumenta à medida que h cresce.

O intervalo de confiança para um AR(1) com resíduos normais é dado por:

$$c \sum_{i=1}^{h-1} \phi^{i-1} + \phi^h y_t \pm 2\sigma_\epsilon \left(1 + \phi^2 + \phi^4 + \dots + \phi^{2(h-1)}\right)^{1/2}$$

Deixando-se H observações fora da estimação, tem-se as opções:

- **Estática:** previsão um passo à frente;
- **Dinâmica:** previsão vários passos à frente.

Para o tamanho de amostra H , calcula-se os erros de previsão:

Raiz do Erro quadrático médio:

$$RMSE = \sqrt{\frac{\sum_{h=1}^H e_t^2(h)}{H}}$$

Erro absoluto médio:

$$MAE = \frac{\sum_{h=1}^H |e_t(h)|}{H}$$

Avaliação da previsão - Exemplo

Obtenha uma série de erros de previsão (de um passo à frente), seguindo os seguintes passos:

1. Estimar modelos utilizando dados do IPCA referente ao período de janeiro de 2004 até dezembro de 2015;
2. Fazer previsão um passo à frente e salvar previsão;
3. Calcular e salvar erro de previsão;
4. Adicionar uma observação e reestimar o modelo;
5. Repetir passos 2-4 até a última observação da amostra.