A Metodologia de *Box & Jenkins*

Aula 03

Bueno, 2011, Capítulo 3

Enders, 2009, Capítulo 2

Morettin e Toloi, 2006, Capítulos 6 a 8

A Metodologia Box & Jenkins

Uma abordagem bastante utilizada para a construção de modelos paramétricos para séries temporais univariadas é conhecida como metodologia Box & Jenkins (1970).

Tal metodologia consiste em propor e ajustar modelos lineares estacionários (ou não-estacionários) a uma série de tempo observada.

A estratégia para a construção do modelo será baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos próprios dados.

A Metodologia Box & Jenkins

O ciclo iterativo é formado pelas seguintes etapas:

Especificação

Identificação

Estimação

Diagnóstico

Especificação

Uma classe geral de modelos é considerada para a análise. Tal escolha pode ser feita, por exemplo, usando a FAC e a FACP estimadas, que esperamos que representem adequadamente as respectivas teóricas, que são desconhecidas

Identificação

O objetivo da identificação é determinar os valores *p* e *q* do processo ARMA(*p*,*q*), que foi especificado na etapa anterior. Aqui, podemos nos basear nas autocorrelações e autocorrelações parciais estimadas (que esperamos que representem adequadamente as respectivas teóricas, que são desconhecidas).

Identificação

Dado o comportamento da função de auto-correlação (FAC) para processos de médias móveis (MA), tal ferramenta se torna bastante útil para identificar a ordem de tais processos.

A auto-correlação na defasagem h, para processos estacionários, indicada por ρ_h, é definida como

$$\rho_h = Corr(x_t, x_{t+h}) = \frac{Cov(x_t, x_{t+h})}{Var(x_t)} = \frac{\gamma_h}{\gamma_0}$$

Observação: -1 ≤ ρ_h ≤ +1, $h \geq 1$ e $\rho_h = 1$, para $h = 0$.

Identificação

Como na prática, temos somente uma única realização de um processo estocástico (isto é, uma única série temporal observada), podemos calcular apenas a função de auto-correlação amostral, dada por:

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0} = \frac{\frac{1}{n} \cdot \sum_{t=1}^{n-h} \left[\left(x_t - \bar{x} \right) \left(x_{t+h} - \bar{x} \right) \right]}{\frac{1}{n} \cdot \sum_{t=1}^{n} \left(x_t - \bar{x} \right)^2}$$

Identificação

Ainda, Bartlett (1946), mostrou que, se uma série temporal apresentar um comportamento puramente aleatório, então os estimadores dos coeficientes de auto-correlação $\hat{\rho}_h$ terão distribuição aproximadamente normal com média zero e variância 1/n, em que n é o tamanho da amostra.

Identificação

Assim, não é difícil demonstrar que o intervalo de confiança bicaudal, com por exemplo 95% de confiança, para qualquer ρ_h será dado por:

$$\hat{\rho}_h \pm 1{,}96/\sqrt{n}.$$

Ainda, poderíamos pensar numa hipótese conjunta de que todos os coeficientes de auto-correlação ρ_h, h = 1, 2, ..., m, são simultaneamente iguais a zero. Ou seja,

$$H_0: \rho_1 = \rho_2 = \ldots = \rho_m = 0.$$

Identificação

A hipótese nula anteriormente formulada pode ser testada utilizando a estatística Q

$$Q(m) = n \cdot \sum_{h=1}^{m} \hat{\rho}_h^2$$

em que

n – tamanho da amostra;

m – número de defasagens.

desenvolvida por Box e Pierce.

Identificação

Sob H_0: $\rho_1 = \rho_2 = ... = \rho_m = 0$,

$$Q(m) = n \cdot \sum_{h=1}^{m} \hat{\rho}_h^2$$

aproxima-se de uma distribuição Qui-quadrado com *m* graus de liberdade.

Observação: Este teste apresenta boas propriedades somente quando estamos lidando com grandes amostras.

Identificação

Uma variante da estatística Q de Box-Pierce é a estatística desenvolvida por Ljung-Box (1978), que apresenta melhores propriedades, que a estatística Q, ao menos para amostras pequenas.

A estatística de Ljung-Box (LB) é definida por

$$LB(m) = n \cdot (n+2) \cdot \sum_{h=1}^{m} \left(\frac{\hat{\rho}_h^2}{n-h} \right) \approx \chi_m^2$$

Observações

- **Todavia, a FAC de processos AR e ARMA podem apresentar um comportamento complicado e, por consequência, a utilização de tal ferramenta não parece muito adequada para a identificação de tais processos.**

- **Box, Jenkins e Reinsel (1994) propõem o uso da *função de auto-correlação parcial* (FACP), que é útil para identificar modelos AR.**

Identificação

Inicialmente, vamos denotar por ϕ_{kj} o j-ésimo coeficiente de uma autorregressão de ordem k, de tal modo que ϕ_{kk} seja o último coeficiente.

Ainda, é sabido que

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \ldots + \phi_{kk}\rho_{j-k}, \quad j = 1, \ldots, k$$

Do resultado anterior, são obtidas as equações de Yule-Walker

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}$$

Identificação

O sistema de equações anteriores, para k = 1, 2, 3, ..., apresenta, respectivamente, as seguintes soluções:

$$\phi_{11} = \rho_1 \qquad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \qquad \phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}$$

Identificação

Para um k genérico

$$\phi_{kk} = \frac{\left|\underset{\sim}{\mathbf{P}_k^*}\right|}{\left|\underset{\sim}{\mathbf{P}_k}\right|}$$

em que

P_k é a matriz de auto-correlações;

P_k* é a matriz P_k com a última coluna substituída pelo vetor de auto-correlações.

Estimação

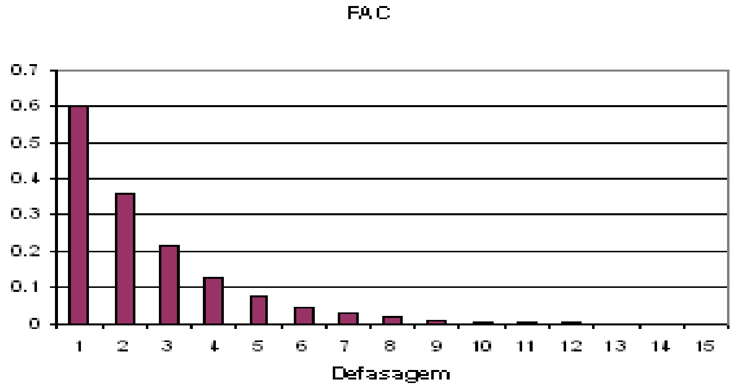
Após termos identificado um modelo provisório para a série temporal, o passo seguinte é estimar seus parâmetros. O método de estimação utilizado pode ser, por exemplo, o da máxima verossimilhança.

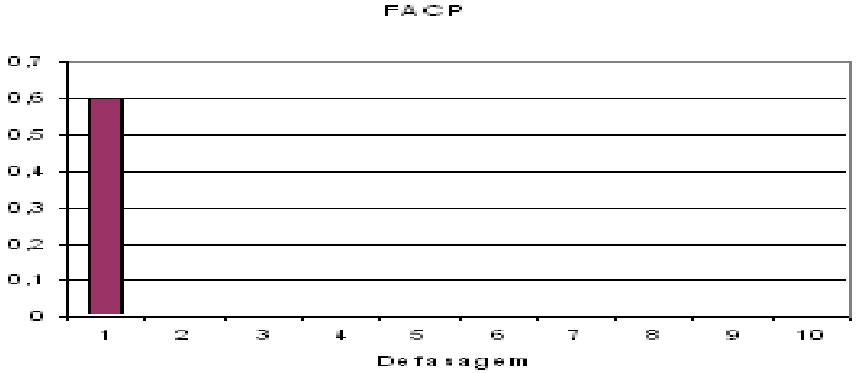
Observação

Além dos EMV, para o caso de modelos autorregressivos, também podemos considerar os estimadores de Yule-Walker.

Exercício 1

Uma série temporal Y_t, t = 1,...T, foi gerada por um processo da classe ARMA(p,q) e apresenta os seguintes formatos para a Função de Autocorrelação (FAC) e Função de Autocorrelação Parcial (FACP):





De acordo com os resultados observados, e levando-se em consideração que a média amostral da série seja 100, proponha um modelo adequado para a série Y_t e dê uma estimativa para os parâmetros de interesse.

Observação

Voltando à Identificação

Também, podemos utilizar os Critérios de Informação de Akaike e/ou Schwarz, por exemplo, para nos auxiliar na identificação dos valores p e q do processo ARMA(p,q).

Observação

Voltando à Identificação

Critério de Informação de Akaike (AIC)

$$AIC = -2\left(\frac{l}{T}\right) + 2\left(\frac{k}{T}\right)$$

em que

l – log-verossimilhança;

T – número de observações;

k – número de parâmetros estimados.

Observação

Voltando à Identificação

Critério de Informação de Schwarz (SIC)
ou
Critério de Informação Bayesiano (BIC)

$$SIC = -2\left(\frac{l}{T}\right) + k\frac{\log(T)}{T}$$

em que

l – log-verossimilhança;

T – número de observações;

k – número de parâmetros estimados.

Diagnóstico

Feita a estimação dos parâmetros do modelo proposto, temos que verificar se ele representa, ou não, adequadamente, os dados.

Tal verificação pode ser feita através de uma análise de resíduos, pois, a partir dos mesmos, teremos ideia se as suposições feitas sobre o modelo são ou não válidas.

Ainda, veremos, também, que qualquer insuficiência revelada pode sugerir um modelo alternativo como sendo adequado.

Diagnóstico

Se o modelo proposto for adequado, os resíduos deverão estar "próximos" dos erros e, portanto, deverão ter média zero, variância constante e ser aproximadamente não auto-correlacionados (ruído branco).

Uma ferramenta bastante útil para se verificar a suposição de que os erros são um ruído branco se baseia no cálculo da função de auto-correlação (FAC) dos resíduos.

Diagnóstico

Indicaremos por

$$\hat{r}_k$$

as auto-correlações entre os resíduos, que, aqui, serão denotados por

$$\hat{a}_t \,.$$

Assim, as auto-correlações, serão dadas pela seguinte expressão:

$$\hat{r}_k = \frac{\sum_{t=k+1}^{n} \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^{n} \hat{a}_t^2}$$

Diagnóstico

Ainda, é esperado que

$$\hat{r}_k \sim N\left(0, \frac{1}{n}\right)$$

sob a suposição que o modelo ajustado é apropriado.

Diagnóstico

Bartlett (1946), mostrou que, se a série residual apresentar um comportamento puramente aleatório, então os estimadores dos coeficientes de auto-correlação $\hat{r}_k$ terão distribuição aproximadamente normal com média zero e variância 1/n, em que n é o tamanho da amostra.

Diagnóstico

Assim, não é difícil demonstrar que o intervalo de confiança bicaudal, com por exemplo 95% de confiança, para qualquer ρ_h será dado por:

$$\hat{r}_k \pm 1{,}96/\sqrt{n}.$$

Todavia, para testar a hipótese conjunta de que todos os coeficientes de auto-correlação r_h, h = 1, 2, ..., m, são simultaneamente iguais a zero, podemos usar a estatística Q, desenvolvida por Box e Pierce.

Diagnóstico

A estatística Q, anteriormente citada, é dada por:

$$Q(m) = n \cdot \sum_{h=1}^{m} \hat{r}_h^2$$

em que

n – tamanho da amostra;

m – duração da defasagem.

Sob H_0,

$$Q \approx \chi^2_{m-p-q}$$

<u>Observação</u>: Este teste apresenta boas propriedades somente quando estamos lidando com grandes amostras.

Diagnóstico

Uma variante da estatística Q de Box-Pierce é a estatística desenvolvida por Ljung-Box (1978), que apresenta melhores propriedades, que a estatística Q, ao menos para amostras pequenas.

A estatística de Ljung-Box (LB) é definida por

$$LB(m) = n \cdot (n+2) \cdot \sum_{h=1}^{m} \left(\frac{\hat{r}_h^2}{n-h} \right) \approx \chi^2_{m-p-q}$$

Observações Finais

Muitas vezes os usuários da metodologia Box & Jenkins identificam não apenas um único modelo, mas alguns modelos que serão então estimados e verificados.

Assim, se o propósito for o de fazer previsões, escolher-se-á entre os modelos ajustados o melhor, por exemplo, no sentido de fornecer o <u>menor erro quadrático médio de previsão</u>.

Observações Finais

Em geral, os modelos postulados são parcimoniosos, pois contêm um número pequeno de parâmetros e as previsões obtidas são bastante precisas, comparando-se favoravelmente com os demais métodos de previsão.

Uma desvantagem do uso da metodologia proposta por Box & Jenkins é que a sua utilização requer experiência e algum conhecimento, além do uso automático de um *software* estatístico.