Programa de Pós-graduação em Engenharia de Produção

Métodos Qualitativos

Análise de séries temporais:

Modelos de Box-Jenkins

Profa. Dra. Liane Werner

Metodologia de Box-Jenkins

- Para os modelos de decomposição e os modelos de suavização exponencial é usual assumir que os erros t (t =1,...,T) são não correlacionados,
- o que implica que as observações também não estão correlacionadas.
- Esta suposição raramente ocorre na prática.
- Correlação serial é esperada uma vez que os dados são coletados seqüencialmente no tempo.

Metodologia de Box-Jenkins

- Os modelos que capturam esta estrutura de correlação são uma classe especial de modelos
- chamados de modelos *Autoregressivos Integrados de Médias Móveis (ARIMA)*.
- Estes modelos apresentam uma variedade de diferentes estruturas de correlação,
- se esta estrutura de correlação for bem modelada, ele fornecerá boas previsões.

Metodologia de Box-Jenkins

- os modelos ARIMA foram popularizados por George Box e Gwilym Jenkins no início dos anos 70,
- e seus nomes tem sido usado como sinônimo destes modelos.
- Box e Jenkins colocaram de forma compreensiva a informação necessária para entender e usar os modelos ARIMA para séries temporais univariadas.

Conceitos Importantes

- Modelos Estocásticos:
  - a idéia de usar modelos matemáticos para descrever o comportamento de fenômenos físicos esta bem estabelecida.
  - Por exemplo, quando se lança um míssil é possível conhecer sua trajetória se sua direção e a velocidade são conhecidas. Ao precisar a trajetória, tem-se um modelo determinístico.
  - Fatores desconhecidos podem ocorrer - tal como a velocidade do vento que pode desviar o míssil de seu curso.

Conceitos Importantes

- Quando o comportamento de uma série temporal está sujeita a fatores desconhecidos, ou seja, a efeitos aleatórios,
- Pode-se encontrar um modelo que é obtido com base no cálculo de probabilidade. Tal modelo é chamado de estocástico.
- Um processo estocástico é caracterizado por uma família de variáveis aleatórias que descrevem a evolução de um fenômeno de interesse, neste caso a evolução temporal da série em estudo.

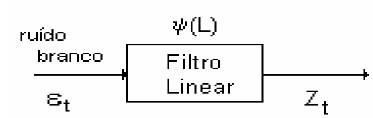
Conceitos Importantes

- *Modelo de Filtro Linear* :
  - Os modelos estocásticos são baseados na idéia de que séries temporais, em que valores sucessivos são altamente dependentes,
  - podem ser estimadas de séries provenientes de "choques" independentes ε_t.
  - Estes choques são aleatórios e provêem de uma distribuição fixa, usualmente assumida como Normal tendo média μ zero e variância σ^2_ε .
  - A seqüência de variáveis aleatórias ε_t, ε_{t-1}, ε_{t-2}, ... é chamada de processo de ruído branco, ou simplesmente ruído branco.

Conceitos Importantes

- *Modelo de Filtro Linear* :
  - o ruído branco ε_t é transformado em Z_t pelo processo chamado filtro linear.
  - O filtro linear faz a soma ponderada das observações anteriores

$$Z_t = \mu + \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + ...$$



$$Z_t = \mu + \psi(L)\varepsilon_t$$

onde: μ é o parâmetro que determina o nível da série;

L é o operador de defasagem , expresso por $L^k\varepsilon_t = \varepsilon_{t-k}$;

$\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + ...$ é o operador linear que transforma ε_t em Z_t.

Conceitos Importantes

- Modelos Estocásticos Estacionários:
  - é uma classe especial de processos estocásticos que assumem que o processo está em “equilíbrio”.
  - Em caso de estacionariedade, a distribuição de probabilidade para $t_1, t_2, ..., t_m$ deve ser a mesma distribuição de probabilidade para os tempos em qualquer tempo $t_{1+k}, t_{2+k}, ..., t_{m+k}$, onde k é um deslocamento arbitrário ao longo do eixo do tempo.

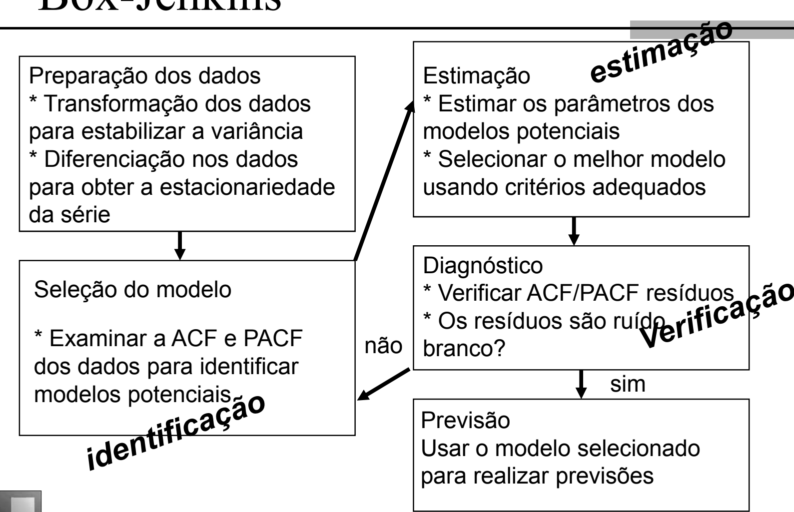
Conceitos Importantes

- Modelos Estocásticos Estacionários:
  - a estacionariedade implica que não existe inclinação nos dados,
  - eles devem permanecer ao redor de uma linha horizontal ao longo do tempo.
  - Os dados flutuam ao redor de uma média constante, independente do tempo e
  - a variação das flutuações permanece essencialmente constante sobre o tempo.
- Modelos Não-Estacionários=> tendência ou variabilidade não constante.

Etapas para a construção dos modelos Box-Jenkins

- A construção dos modelos Box-Jenkins, é baseada em um ciclo iterativo,
- no qual a escolha do modelo é feita com base nos próprios dados.
- São três etapas para construção do modelo:
  - (i) identificação;
  - (ii) estimação e
  - (iii) verificação.
- Makridakis et al. (1998) esquematizaram as etapas para a construção de modelos

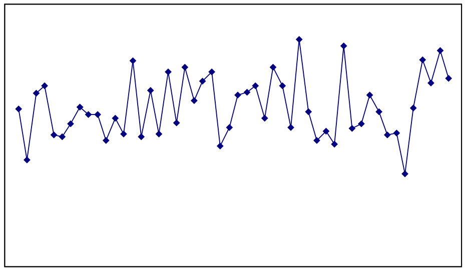
Etapas para a construção dos modelos Box-Jenkins



Fonte: Makidrakis et al. (1998)

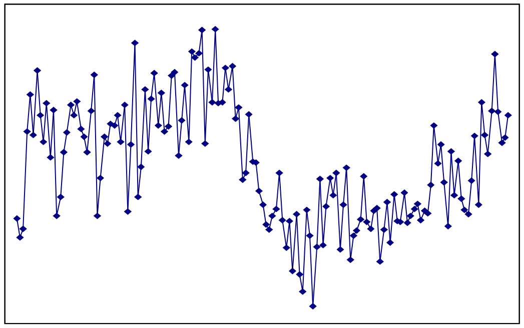
Preparação dos dados
Examinando a estacionariedade da série

- Para dectectar se uma série é estacionária ou não, sugere-se o uso do **time plot** (gráfico de linha onde é plotada a variável em estudo contra o tempo).
- Se a série temporal é plotada e não existe evidência de mudanças na média sobre o tempo, então diz-se que a série é estacionária na média.
- Se o gráfico não mostra mudança na variação ao longo do tempo, então diz-se que a série é estacionária na variância



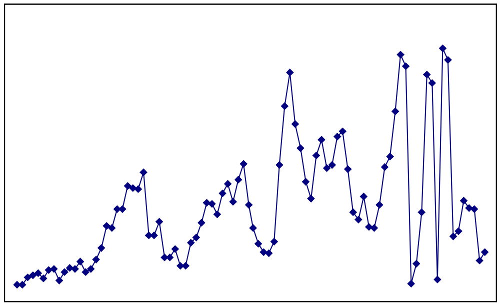
Preparação dos dados
Examinando a estacionariedade da série

- Se uma série apresenta uma certa inclinação ao longo do tempo ou então uma mudança de nível é por que a série é não-estacionária na média.



Preparação dos dados
Examinando a estacionariedade da série

- Uma série é não-estacionária na variância, quando esta não se mantém constante ao longo do tempo.



Preparação dos dados
Examinando a estacionariedade da série

- A forma analítica não é a única maneira de se detectar a estacionariedade de uma série,
- existem testes estatísticos, chamados de teste de raiz unitária, que verificam se uma série é ou não estacionária.
- O teste de raiz unitária mais usado é o teste de Dickey-Fuller.
- Maiores detalhes sobre o teste podem ser pesquisados em:
  - MAKRIDAKIS, S., WHEELWRIGHT, S. C. & HYNDMAN, R. J. **Forecasting. Methods and Applications**. Third Edition. John Wiley & Sons. New York, 1998.
  - ENDERS, W. **Applied Econometric Time Series.** John Wiley & Sons. New York, 1995.

Preparação dos dados
Removendo a não estacionariedade

- Caso ainda não sejam estacionárias, será necessário transformá-las.
- A transformação mais comum consiste em tomar diferenças sucessivas da série original, até obter uma série estacionária.
- A primeira diferença de Z_t é: $\Delta Z_t = Z_t - Z_{t-1}$,
- A segunda diferença é por:

$$\Delta^2 Z_t = \Delta\,[\Delta\, Z_t\,] = \Delta[Z_t - Z_{t-1}] = Z_t - 2Z_{t-1} - Z_{t-2}$$

- Em geral tomar uma ou duas diferenças é suficiente para que a série se torne estacionária.

Preparação dos dados
Removendo a não estacionariedade

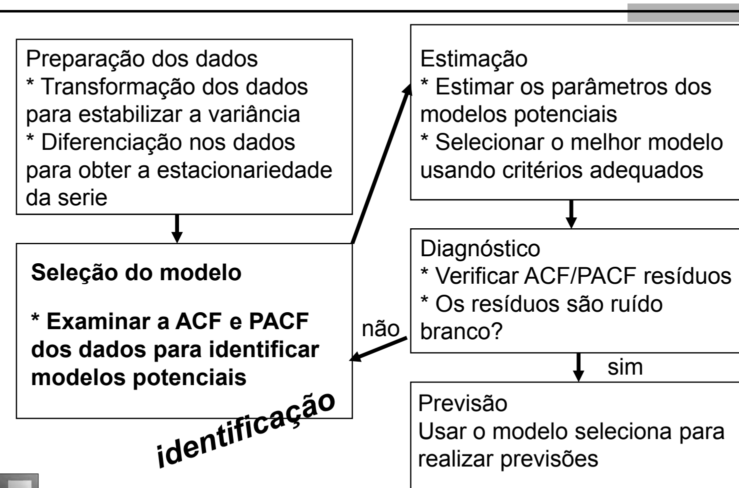
- Quando a série é sazonal e não estacionária é apropriado tomar as diferenças para o período de sazonalidade.
- Assim, dados mensais devem apresentar um padrão anual de sazonalidade, neste caso a estação de sazonalidade s é 12.
- As diferenças sazonais para dados mensais é dada por:

$$\Delta Z_t = Z_t - Z_{t-12}$$

Estacionariedade - Exercício

- O arquivo TEQ8_series.xls contém 5 planilhas.
- Em cada planilha uma série temporal,
- Construa os gráficos de tempo para cada série e
- verifique se cada uma das séries apresentam estacionariedade ou não.
- No caso da serie ser não-estacionária, verificar se é em relação ao nível ou a variância. Remova a estacionariedade, caso seja necessário.
- Após resolução

Etapas para a construção dos modelos Box-Jenkins



Fonte: Makridakis et al. (1998)

Seleção do modelo
Função de autocorrelação

- Outra implicação, da estacionariedade, que tem importante interpretação em termos do comportamento da série temporal
- é o fato que a autocovariância entre quaisquer duas observações depende somente do número de períodos de tempo que separam eles.
- A covariância entre Z_t e Z_{t+k} é chamada de autocovariância de defasagem (ou lag) k

$$\gamma_k = \text{cov}(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)]$$

Seleção do modelo
Função de autocorrelação

- como a autocorrelação é função da autocovariância, em séries estacionárias,
- esta também dependerá apenas do número de defasagem entre duas observações

$$\rho_k = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{\sqrt{E[(Z_t - \mu)^2]E[(Z_{t+k} - \mu)^2]}} = \frac{E[(Z_t - \mu)(Z_{t+k} - \mu)]}{\sigma_Z^2}$$

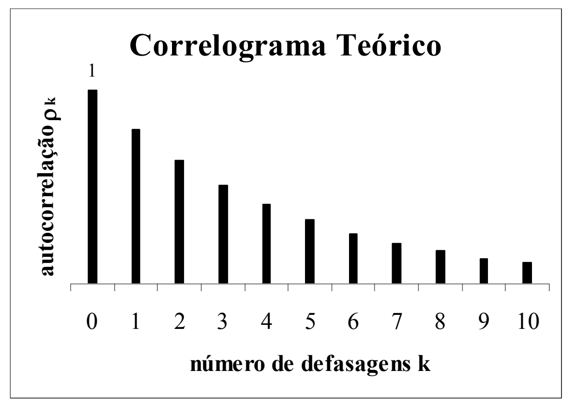
Cabe ressaltar que para k = 0 o valor de ρ_k será sempre igual a um.

Seleção do modelo
Função de autocorrelação

- O coeficiente de autocorrelação ρ_k, que mede a correlação entre observações distantes k períodos de tempo.
- Assim ρ_1 indica como valores sucessivos de Z_t se relacionam, ρ_2 indica como os valores a cada dois períodos se relacionam e assim por diante.
- Juntas, as autocorrelações de defasegem 1,2,..., constituem a função de autocorrelação ou ACF.
- O gráfico da função de autocorrelação é chamado de correlograma.

Seleção do modelo
Função de autocorrelação

Correlograma para o modelo: $Z_t = 0{,}8\, Z_{t-1} + \varepsilon_t$



Seleção do modelo
Função de autocorrelação

- Função de autocorrelação amostral
- Ao analisar uma série temporal é preciso obter estimativas para os valores da correlações ρ_k.
- Utiliza-se então o coeficiente de autocorrelação amostral:

$$r_k = \frac{\sum_{t=k+1}^{T}(Z_t - \overline{Z})(Z_{t+k} - \overline{Z})}{\sum_{t=1}^{T}(Z_t - \overline{Z})^2} \qquad \text{com } k = 0, 1, 2, ..., T$$

$$\text{onde: } \overline{Z} = \frac{1}{T}\sum_{t=1}^{T} Z_t$$

Seleção do modelo
Função de autocorrelação

- Nos correlogramas amostrais são traçados limites paralelos ao eixo do número de defasagens k,
- tais limites correspondem a $\pm 2\sigma_{r_k}$, que são úteis para verificar se as autocorrelações são significativamente diferente de zero.
- Além disto, a forma da ACF amostral irá auxiliar na especificação de um modelo em particular.

desvio padrão do coeficiente de autocorrelação amostral

Seleção do modelo
Função de autocorrelação

- Função de autocorrelação parcial
- Autocorrelações parciais são usadas para medir o graus de associação entre Z_t e Z_{t-k}, quando os efeitos das outras defasagens (de 1, 2, ..., k-1) são removidos.
- O coeficiente de autocorrelação parcial é presentada por ϕ_{kk}.
- Juntas, as autocorrelações parciais de defasegem 1,2,..., constituem a função de autocorrelação parcial ou PACF.
- O gráfico da função de autocorrelação parcial é chamado de correlograma parcial.

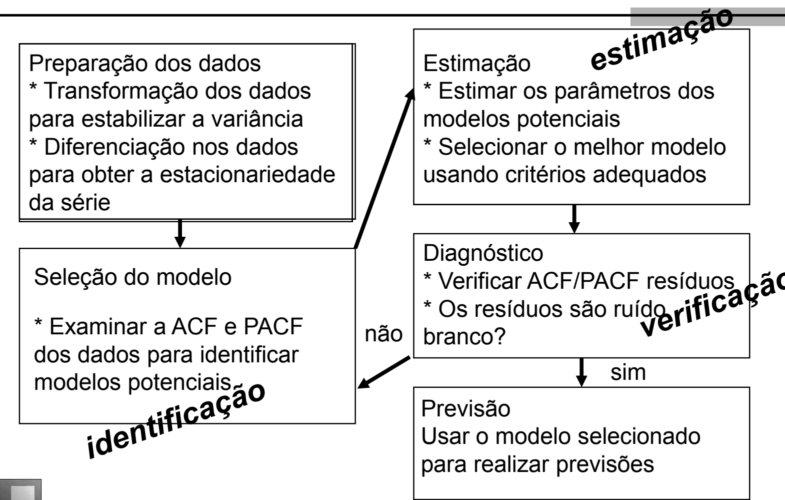
Seleção do modelo
Função de autocorrelação

- Autocorrelação e Autoc. Parcial e a Sazonalidade
- Caso os dados sejam observados em períodos mensais e se o padrão de sazonalidade é consistente,
- o coeficiente de autocorrelação de defasagens 12 deverá ser grande, indicando a existência de sazonalidade.
- A sazonalidade pode ser encontrada pela identificação de coeficientes de autocorrelação altos ou coeficientes de autocorrelação parciais altos para defasagens de sazonalidade.
- No caso de dados mensais para defasagens múltiplas de 12.

Função de autocorrelação - Exercício

- Copie para o SPSS, cada série do arquivo TEQ8_series.xls (a série sem influência da estacionariedade)
- Construa as ACF e PACF para cada série.

Etapas para a construção dos modelos Box-Jenkins



Fonte: Makidrakis et al. (1998)

Métodos Qualitativos

Análise de séries temporais:

Modelos de Box-Jenkins

Profa. Dra. Liane Werner