

2

Metodologia de Previsão de Séries Temporais - Box & Jenkins

2.1

Introdução

A análise de séries temporais, segundo Box & Jenkins (1994), tem como objetivo principal a realização de previsão. Essa metodologia permite que valores futuros de uma série sejam previstos tomando por base apenas seus valores presentes e passados. Isso é feito através da correlação temporal existente entre os valores existentes.

Segundo Tápia, (2000), a realização do processo temporal pelo método de Box & Jenkins é representada por um conjunto de processos estocásticos denominados modelos ARIMA (*autoregressive integrated moving average*) onde, em cada instante de tempo t , existe um conjunto de valores que a série pode assumir, aos quais estão associadas possibilidades de ocorrência.

Para cada instante de tempo t , é possível que exista uma função de densidade de probabilidade; logo, cada variável aleatória Z_t , $t = t_1, t_2, \dots$ pode ter média e variância específicas.

O trabalho consiste em descobrir qual é o processo que gera a série em estudo, isto é, qual o modelo que representa melhor a série.

A metodologia Box & Jenkins é aplicada aos processos estocásticos que sejam estacionários. Um processo estocástico é dito estacionário de segunda ordem quando as seguintes condições forem satisfeitas para qualquer instante de tempo t :

$$E[z_t] = E[z_{t+k}] = \mu$$

$$\text{Var}[z_t] = E[(z_t - \mu)^2] = \sigma^2$$

$$\text{Cov}[z_t, z_{t+k}] = E[(z_t - \mu) \cdot (z_{t+k} - \mu)]$$

As duas primeiras condições indicam que a média e a variância de Z_t não variam com o tempo, e a terceira indica que as autocovariâncias não dependem do tempo e sim da distância k que separa as observações.

Se o processo estocástico não for estacionário, este pode se tornar estacionário por meio de sucessivas diferenciações da série original.

Quando a série recebe a influência de fatores sazonais, outro tipo de correlação passa a ter importância: a correlação entre os instantes de tempo distantes entre si por s ou múltiplos de s , onde s representa o período da sazonalidade.

2.2

Possíveis Modelos Não sazonais na Metodologia de Box & Jenkins

2.2.1

Modelo Auto-regressivo (AR)

Os modelos auto-regressivos foram criados com a idéia de que a presente observação da série Z_t pode ser explicada como uma função das p observações passadas, $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$, onde p determina o número de passos entre as observações passadas e a previsão da próxima observação.

A estrutura auto-regressiva geral é expressa por:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

Onde:

ϕ_i são parâmetros da estrutura, $i = 1, \dots, p$ (ordem da estrutura)

a_t é ruído branco com média zero e variância σ_a^2 .

Utilizando o operador de defasagem B , tem-se:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = a_t$$

onde $BZ_t = Z_{t-1}$

daí, $\phi(B)Z_t = a_t$

e o operador auto-regressivo é $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$.

2.2.2 Modelo Médias Móveis (MA)

Os modelos médias móveis são formados por combinação linear do ruído branco, a_t , ocorridos no período corrente e nos períodos passados.

A estrutura de médias móveis geral é expressa por:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} \dots - \theta_q a_{t-q}$$

Onde:

θ_i são parâmetros da estrutura, $i = 1, \dots, q$ (a ordem da estrutura)

a_t é ruído branco com média zero e variância σ_a^2 .

Utilizando o operador de defasagem B, tem-se:

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

daí, $Z_t = \theta(B) a_t$

e o operador médias móveis é $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$.

2.2.3 Modelo Auto-Regressivo Médias Móveis (ARMA)

Esse modelo é uma combinação dos dois anteriores onde Z_t é descrito por seus valores passados e pelos ruídos branco corrente e passados.

A estrutura geral ARMA(p,q) é expressa por:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Onde:

ϕ_i são os parâmetros da estrutura auto-regressiva, $i = 1, \dots, p$

θ_i são os parâmetros da estrutura médias móveis, $i = 1, \dots, q$

a_t ruído branco

Usando o operador de defasagem B, tem-se:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

ou $\phi(B) Z_t = \theta(B) a_t$.

2.2.4

Modelo Auto-Regressivo Integrado de Médias Móveis (ARIMA)

O modelo ARIMA (p, d, q) (*Auto-Regressive Integrated Moving Average*) é adequado para a previsão de séries temporais cujo processo estocástico não é estacionário. Logo, a série original passará por algumas diferenciações a fim de torná-la estacionária (Box & Jenkins, 1994).

O número necessário de diferença para tornar uma série estacionária é denominado ordem de integração (d).

A estrutura geral ARIMA(p, d, q) é expressa por:

$$\phi(B) \nabla^d Z_t = \theta(B) a_t$$

Onde:

$\phi(B)$ representa o operador auto-regressivo de ordem p

$\theta(B)$ representa o operador médias móveis de ordem q

a_t ruído branco

d representa o número de diferenças

$\nabla = 1 - B$ representa o operador diferença

Este operador diferença é definido como:

$$Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t = \nabla Z_t$$

$$\nabla^d = (1 - B)^d$$

2.2.5

Etapas da Metodologia de Box & Jenkins

O objetivo da metodologia Box & Jenkins é determinar os três componentes que configuram qualquer estrutura que são: p parâmetros auto-regressivos, d processos de diferenciação (integração) e q parâmetros de médias móveis.

De uma forma geral, a notação apresentada por Box & Jenkins é do tipo ARIMA (p, d, q). Por exemplo, a estrutura ARIMA (1, 2, 0) significa que esta possui um parâmetro auto-regressivo, duas diferenciações a partir da série original e nenhum parâmetro de médias móveis.

O procedimento padrão para utilização da metodologia de previsão ARIMA consiste nos seguintes passos, segundo Box & Jenkins:

1. Realiza-se a diferenciação da série original tantas vezes quanto necessário para torná-la estacionária;
2. Identifica-se os valores p e q através da análise das funções de autocorrelação (ACF) e de autocorrelação parcial (PACF) estimadas e a estimação dos parâmetros (ϕ_i , $i = 1, \dots, p$ e θ_i , $i = 1, \dots, q$).

A análise destas funções permite a tomada de decisão acerca do número p de parâmetros auto-regressivos e do número q de parâmetros de médias móveis que devem ser escolhidos de maneira a se obter uma estrutura parcimoniosa (ou seja, uma estrutura que tenha o menor número de parâmetros dentre todas as estruturas que se ajustem aos dados da série);

3. Realizam-se as previsões (obtenção dos novos valores da série) e os intervalos de confiança para as mesmas.

Para verificar se é necessária a diferenciação da série original, é comum utilizar o gráfico da mesma. Na maioria dos casos, a série pode ser não-estacionária quanto ao nível ou quanto à inclinação. Quando a série for não-estacionária quanto ao nível, para torná-la estacionária basta tomar a primeira diferença e, quando for não-estacionária quanto à inclinação, basta tomar a segunda diferença (Souza e Camargo, 2004).

Por vezes é necessário realizar uma transformação logarítmica na série original para torná-la estacionária.

Deve-se considerar; que, na metodologia desenvolvida por Box & Jenkins, os estádios da identificação e estimação se sobrepõem. Muitas vezes, ao se sobrefixar uma estrutura na etapa de identificação, aquela será corrigida na estimação, quando a significância dos parâmetros estimados será avaliada. Na maioria dos casos, o resultado da identificação não será preciso e mais de uma

estrutura será identificada. Esta imprecisão é consequência de que, na prática, o ajuste a situações reais induz a alguns erros.

A identificação visa encontrar uma classe de estruturas a serem ajustadas e validadas (Box & Jenkins, 1994; Souza e Camargo, 2004 e Machado, 2000). Procura-se então, encontrar um subconjunto adequado de estruturas ARMA (p, q),

$$\phi(B)z_t = \theta_0 + \theta(B)a_t ,$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p ,$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_p B^p , \text{ sendo } B \text{ o operador de atraso,}$$

dado por:

$$B^k z_t = z_{t-k} , k = 1, 2, \dots$$

através das funções de autocorrelação e autocorrelação parcial que representem a série temporal.

O coeficiente de autocovariância entre duas observações do modelo depende somente do número de períodos que as separam. Logo, a função de autocovariância,

$$\gamma_k = \text{COV}(z_t, z_{t+k}) = E \{ [z_t - E(z_t)] [z_{t+k} - E(z_{t+k})] \}, \text{ é reduzida à}$$

$\gamma_k = E \{ [z_t - \mu] [z_{t+k} - \mu] \}$, assumindo $\mu = E(z_t) = E(z_{t+k})$ (média constante para processos estacionários).

1) Função de autocorrelação

A função de autocorrelação é dada por:

$$\rho_k = \frac{\gamma_k}{\gamma_0}; \quad k = 0, 1, 2, \dots \quad (2.1)$$

Onde:

$$\gamma_0 = \text{var}(z_t)$$

Logo, ρ_k será uma medida padrão de dependência com $|\rho_k| \leq 1, k= 1,2,\dots$

O comportamento da ACF teórica para as estruturas Box & Jenkins é:

i) Estrutura auto-regressiva de ordem p

A estrutura AR(p) é dada por:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t$$

As autocovariâncias podem ser obtidas multiplicando-se esta equação por z_{t-k} e tomando-se os valores esperados (assumindo sem perda de generalidade $E\{z_t\} = \mu = 0$). Daí,

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}$$

Assim, em termos das autocorrelações, utilizando (2.1), temos:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}$$

isto é, ρ_k satisfaz à equação:

$$\phi(B) \cdot \rho_k = 0$$

Onde:

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

ii) Estrutura médias móveis de ordem q

A estrutura MA(q) é dada por:

$$z_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q},$$

supondo $E\{z_t\} = \mu = 0$, a autocovariância de z_t é :

$$\gamma_k = E\{z_t z_{t-k}\} = E\left\{[a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}] \cdot [a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}]\right\}$$

como os ruídos a_t são independentes, $\gamma_t = 0$ para $k > q$. Quando $k \leq q$,

$$\gamma_k = (-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q) \sigma_a^2$$

pois a independência dos ruídos equivale a :

$$E\{a_i a_j\} = \begin{cases} \sigma_a^2 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}.$$

A variância de z_t , γ_0 é obtida de :

$$\gamma_0 = E\{z_t z_t\} = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2$$

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

$$\rho_k = \begin{cases} 1, & k = 0 \\ \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} = \frac{-\theta_k + \sum_{j=1}^{p-k} \theta_j \theta_{j+k}}{1 + \sum_{j=1}^p \theta_j^2}, & k = 1, 2, \dots, q \\ 0, & k > q \end{cases}$$

Verifica-se, assim, que a ACF de uma estrutura MA(q) sofre um corte brusco no *lag* q.

iii) Estrutura ARMA(p, q)

A estrutura ARMA(p, q) é dada por:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Multiplicando-se por z_{t-k} e calculando-se os valores esperados de cada termo, do mesmo modo que na estrutura AR(p), a função de autocovariância é dada por:

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p}, \quad k \geq q+1$$

e a função de autocorrelação é dada por:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k \geq q+1$$

Para um ARMA(p, q), existem q autocorrelações, cujos valores dependem diretamente dos q parâmetros médias móveis e dos p parâmetros auto-regressivos. Pode-se concluir que a ACF da estrutura ARMA(p, q) é a combinação das ACF dos dois processos componentes AR e MA; ou seja, as autocorrelações nos lags 1 a q são afetadas pela parte MA da estrutura, a partir daí, a ACF se comporta como um processo AR(p).

2) Função de autocorrelação parcial

A identificação do grau do polinômio $\phi(B)$ da estrutura AR(p) é realizada através das funções de autocorrelação e de autocorrelação parcial. Para definir esta função, considere-se a função de autocorrelação da estrutura AR(p) dada por:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}$$

Onde:

$$\phi(B) \cdot \rho_k = 0.$$

Fazendo-se $k = 1, 2, \dots, p$ e levando-se em conta que $\rho_k = \rho_{-k}$, obtém-se o sistema conhecido como equações de Yule-Walker:

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} \\ &\dots \\ \rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \dots + \phi_p\rho_p\end{aligned}$$

Assim, é possível obter estimativas para os parâmetros ϕ_i 's pela substituição dos valores da ACF nas equações de Yule-Walker. Denotando os estimadores dos ϕ_i 's por ϕ_{ii} , define-se a função de autocorrelação parcial como a seqüência dos ϕ_{kk} 's obtidos de:

$$\begin{array}{ccccc} 1 & \rho_1 & \rho_{k-1} & \phi_{k1} & \rho_1 \\ \rho_1 & 1 & \rho_{k-2} & \phi_{k2} & \rho_2 \\ \dots & & & & \\ \rho_{k-1} & \rho_{k-2} & 1 & \phi_{kk} & \rho_k \end{array}$$

Onde:

ρ_k é a ACF de lag k , $k = 1, 2, \dots$

i) Estrutura AR(p)

Nas estruturas AR(p), a função de autocorrelação parcial, ϕ_{kk} , é finita para $k \leq p$ e anula-se bruscamente nos lags $k > p$.

ii) Estrutura MA(q)

A função de autocorrelação parcial é formada por exponenciais e/ou senóides amortecidas.

iii) Estrutura ARMA(p, q)

Como a estrutura ARMA(p, q) corresponde a uma estrutura AR(p) de ordem infinita ou a uma estrutura MA(q) de ordem também infinita, dos resultados

anteriores pode-se concluir que a função de autocorrelação parcial de uma estrutura ARMA(p,q) comporta-se de um modo misto mas sem particularidades notáveis (Souza e Camargo, 2004).

A identificação das estruturas Box & Jenkins baseia-se no comportamento teórico das funções de autocorrelação e da função de autocorrelação parcial. Na prática, as funções teóricas não são disponíveis sendo utilizados os estimadores amostrais das funções de autocorrelação e da função de autocorrelação parcial.

Nas figuras a seguir, encontram-se os gráficos representativos dos padrões comportamentais das funções de autocorrelação e autocorrelação parcial e a região de admissibilidade dos modelos comumente encontrados na prática (Souza e Camargo, 2004). A partir dos gráficos destas funções, pode-se determinar o comportamento dos parâmetros dos modelos.

1- Modelo AR(1)

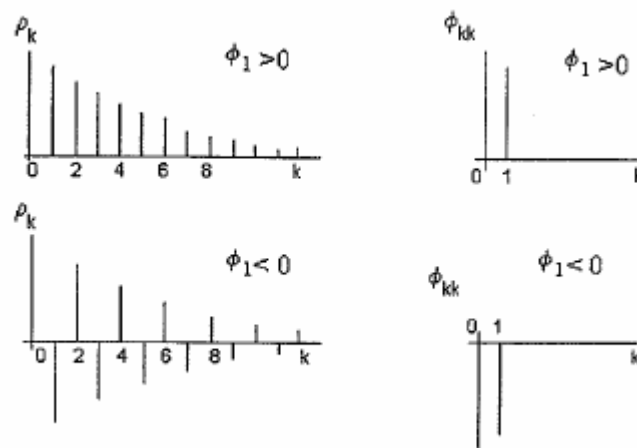


Figura 1 - ACF e PACF do modelo AR(1)

Fonte: Souza e Camargo (2004:69)

2 – Modelo AR(2)

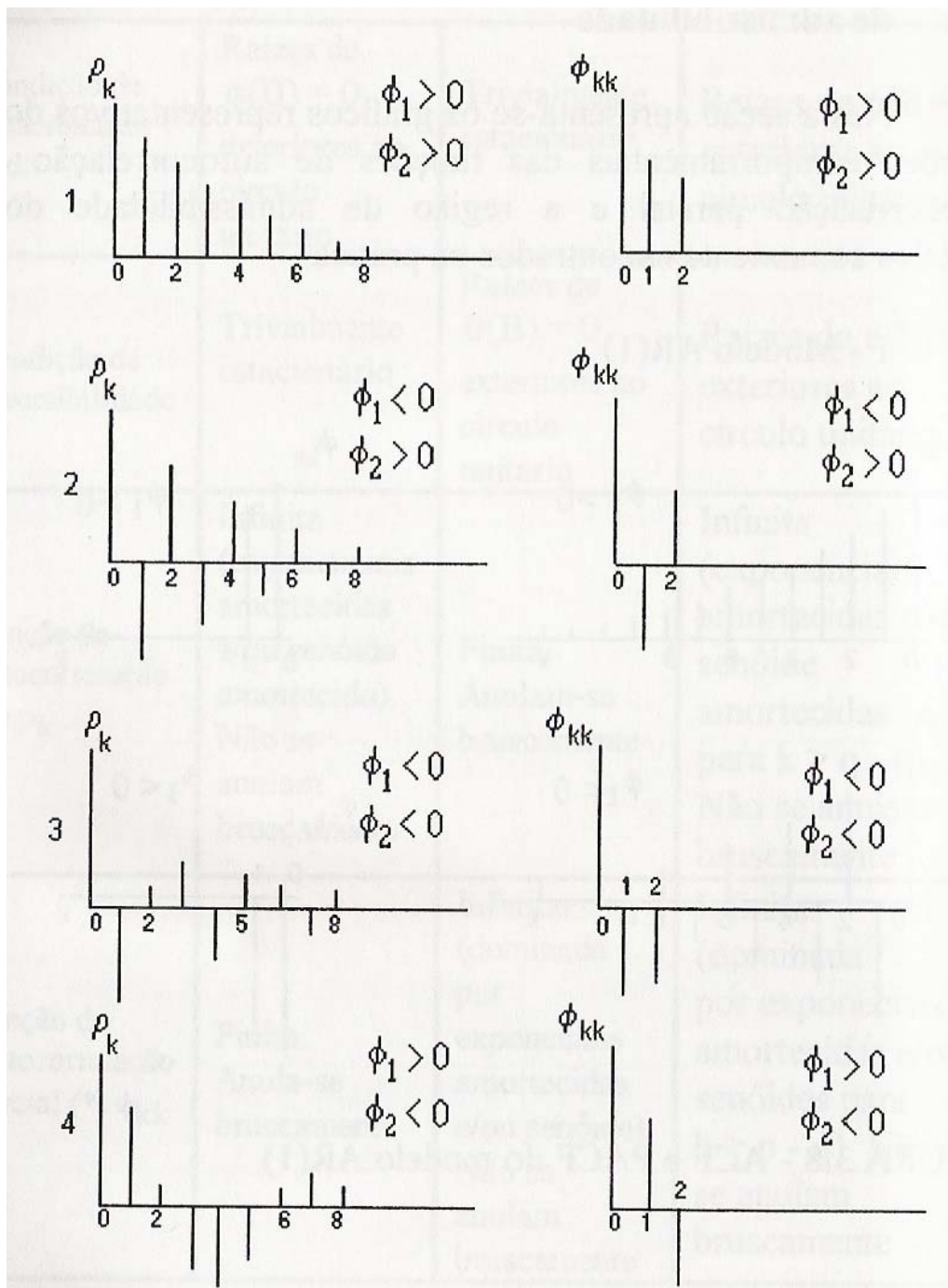


Figura 2 - ACF e PACF do modelo AR(2)

Fonte: Souza e Camargo (2004:70)

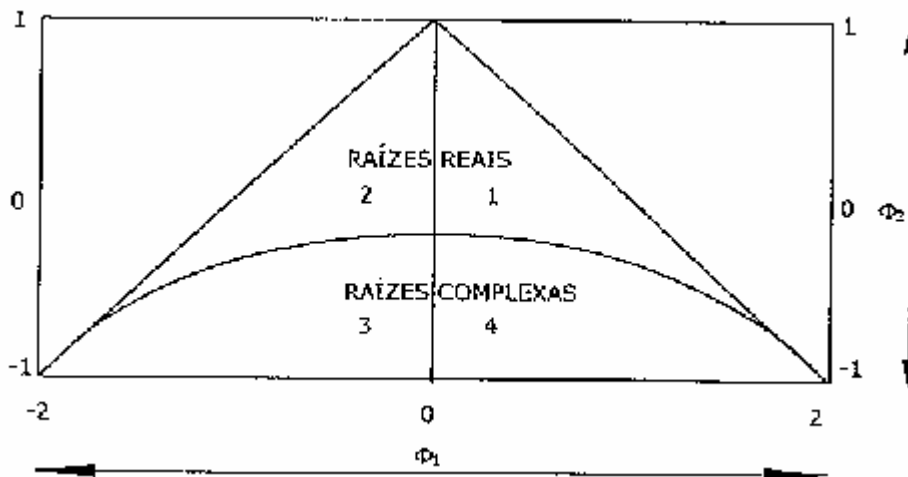


Figura 3 - Região de admissibilidade em função de ϕ_1 e ϕ_2 para o modelo AR(2)

Fonte: Souza e Camargo (2004:70)

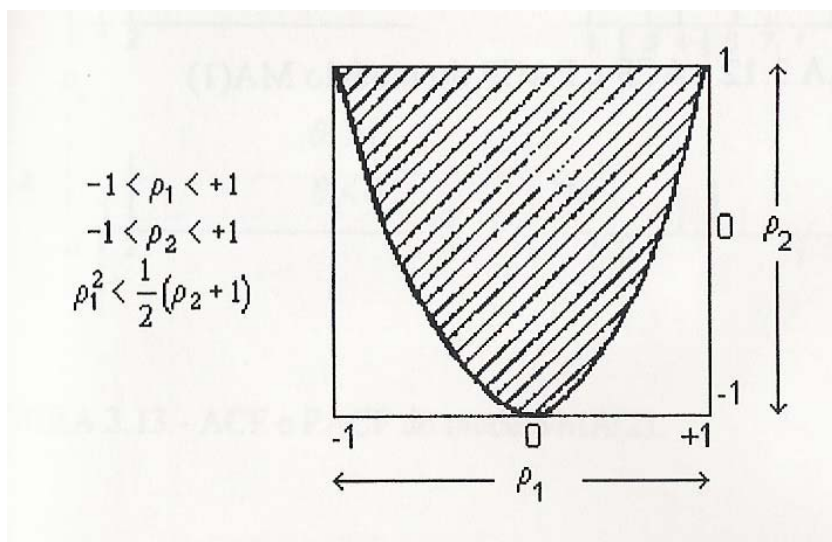


Figura 4 - Região de admissibilidade em função de ρ_1 e ρ_2 para o modelo AR(2)

Fonte: Souza e Camargo (2004:70)

3- Modelo MA(1)

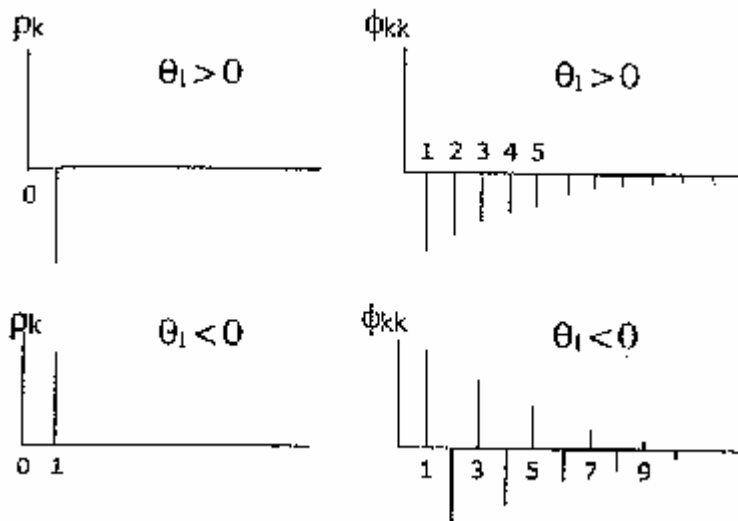


Figura 5 - ACF e PACF do modelo MA(1)

Fonte: Souza e Camargo (2004:71)

4 – Modelo MA(2)

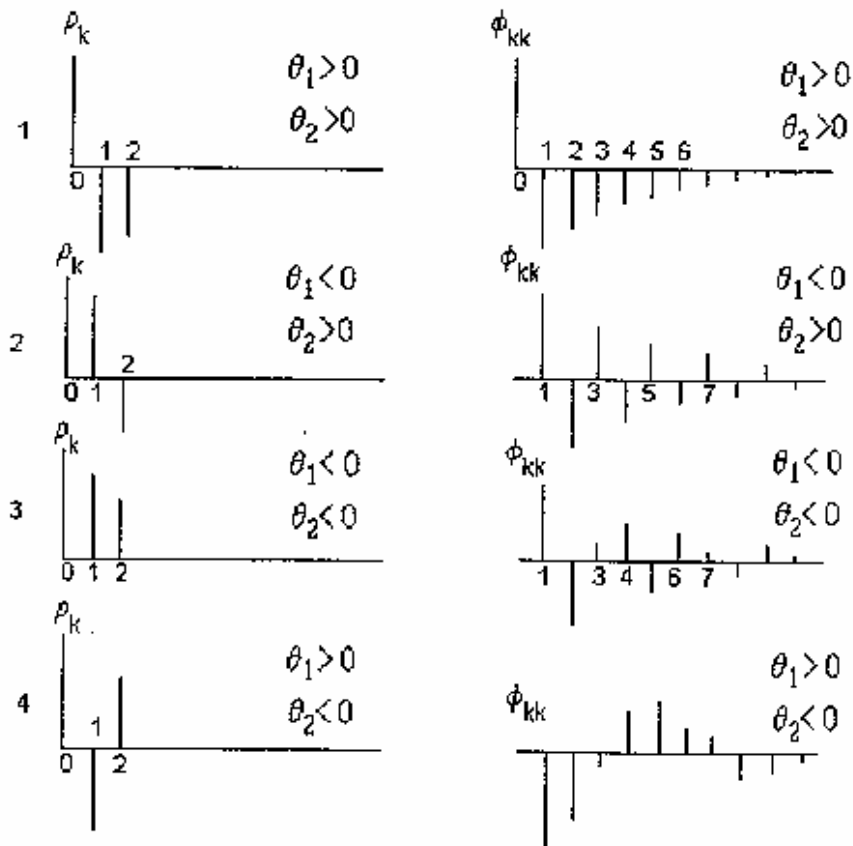


Figura 6 - ACF e PACF do modelo MA(2)

Fonte: Souza e Camargo (2004:71)

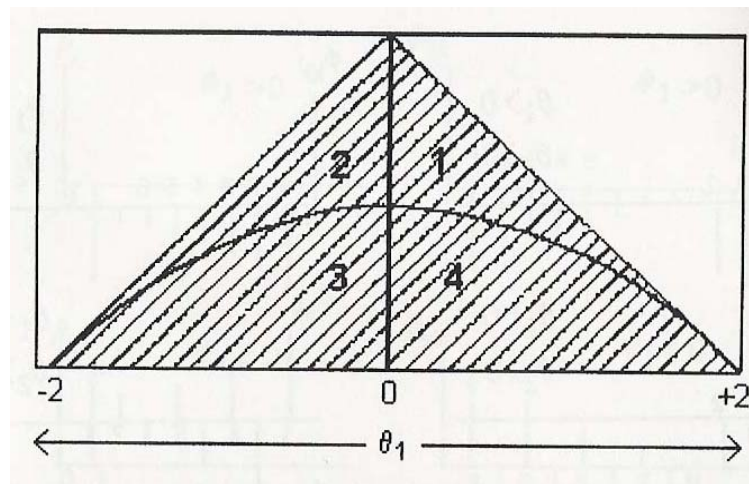


Figura 7 - Região de admissibilidade em função de θ_1 e θ_2 para o modelo MA(2)

Fonte: Souza e Camargo (2004:72)

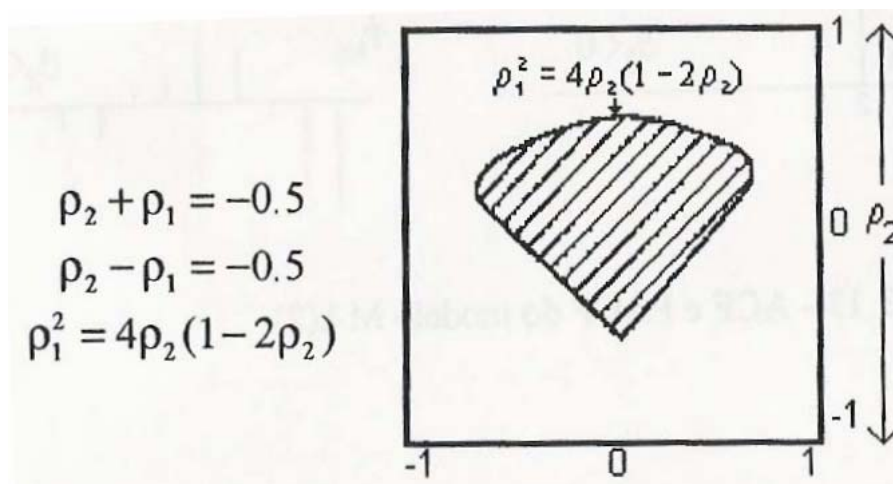


Figura 8 - Região de admissibilidade em função de ρ_1 e ρ_2 para o modelo MA(2)

Fonte: Souza e Camargo (2004:72)

5 – Modelo ARMA(1, 1)

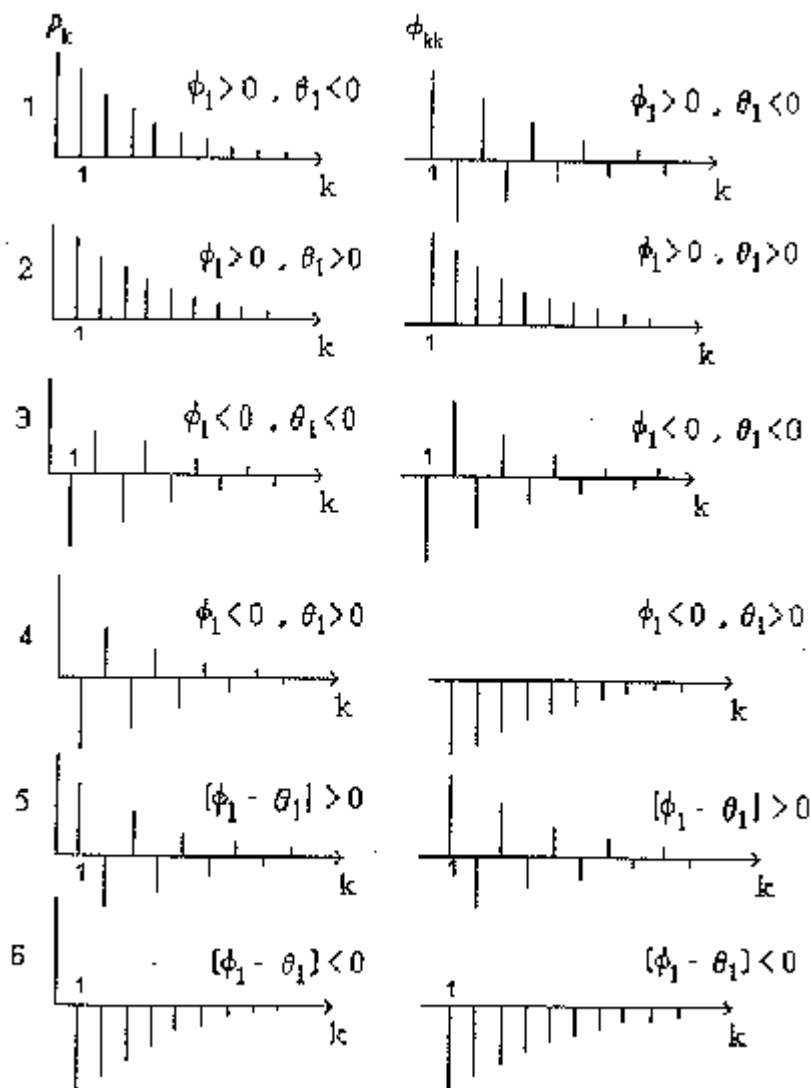


Figura 9 - ACF e PACF do modelo ARMA(1,1)

Fonte: Souza e Camargo (2004:73)

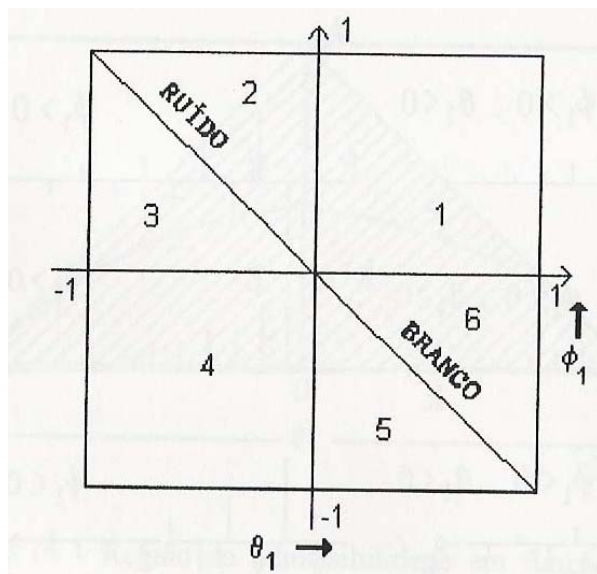


Figura 10 - Região de admissibilidade em função de θ_1 e ϕ_1 para o modelo ARMA(1, 1)

Fonte: Souza e Camargo (2004:73)

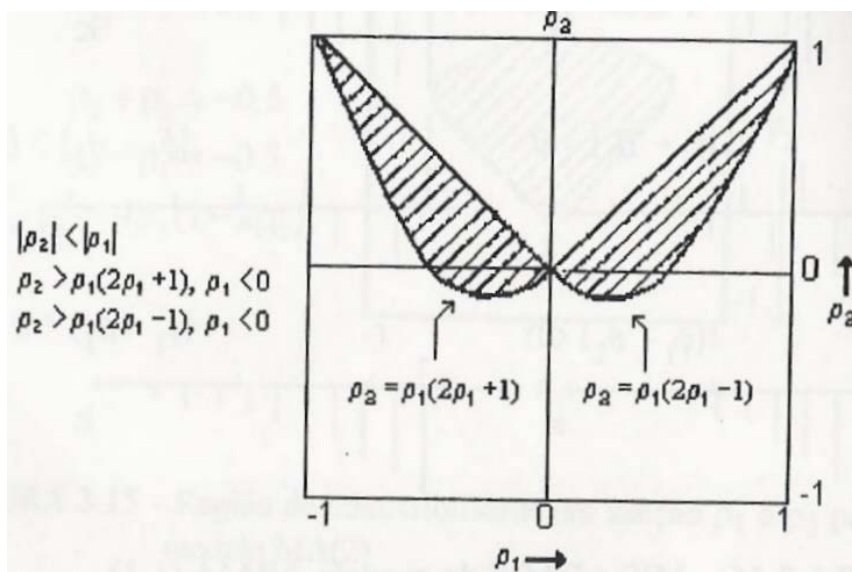


Figura 11 - Região de admissibilidade em função de ρ_1 e ρ_2 para o modelo ARMA(1, 1)

Fonte: Souza e Camargo (2004:74)

2.3 Sazonalidade

A sazonalidade representa as flutuações periódicas que ocorrem no período máximo de um ano, estando associadas as variações climáticas (estações do ano) e datas festivas (Natal, dia das crianças, Carnaval, Páscoa, etc.).

O período de sazonalidade é representado por s , com

- $s = 2$ representa séries semestrais;
- $s = 4$ representa séries trimestrais;
- $s = 12$ representa séries mensais;
- $s = 52$ representa séries semanais;
- $s = 365$ representa séries diárias.

O que se observa em séries sazonais é que ocorrem relações entre:

- Observações para meses sucessivos em um ano particular;
- As observações para o mesmo mês em anos sucessivos.

Assim, Z_t é relacionada com Z_{t-1} , Z_{t-2} , ..., mas também com Z_{t-s} , Z_{t-2s} , Isto implica que séries sazonais são caracterizadas por apresentarem correlação alta em *lags* sazonais, isto é, *lags* que são múltiplos do período s . Um ajuste sazonal, processo de retirada/filtragem do termo sazonal de uma série temporal, será tal que removerá esta correlação ou pelo menos removerá grande parte.

2.4 Possíveis Modelos Puramente Sazonais na Metodologia de Box & Jenkins

2.4.1 Modelo Sazonal Auto-Regressivo de Ordem P – SAR(P)_s (Seasonal Autoregressive)

Esse tipo de modelo permite apenas correlações entre instantes de tempo múltiplos de s .

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_P Z_{t-Ps} + a_t$$

$$(1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps}) Z_t = a_t$$

$$\Phi(B^s) Z_t = a_t$$

A ACF do modelo SAR(P)_s tem valores não nulos apenas nos *lags* múltiplos de *s*.

2.4.2

Modelo Sazonal de Médias Móveis de Ordem Q – SMA(Q)_s (Seasonal Moving Average)

$$Z_t = a_t - \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - \dots - \Theta_Q a_{t-Qs}$$

$$Z_t = (1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_Q B^{Qs}) a_t$$

$$Z_t = \Theta(B^s) a_t$$

A ACF do modelo SMA(Q)_s tem valores não nulos apenas nos *lags* múltiplos de *s*.

2.4.3

Modelo Sazonal Auto-Regressivo de Médias Móveis – SARMA(P,Q)_s (Seasonal Autoregressive Moving Average)

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_P Z_{t-Ps} a_t - \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - \dots - \Theta_Q a_{t-Qs}$$

$$\Phi(B^s) Z_t = \Theta(B^s) a_t$$

A ACF do modelo SARMA(P,Q)_s é decrescente e tem valores não nulos apenas nos *lags* múltiplos de *s*.

2.4.4

Modelo Sazonal Auto-Regressivo Integrado de Médias Móveis – SARIMA(P,D,Q)_s (Seasonal Autoregressive Integrated Moving Average)

Esse modelo se aplica a séries não estacionárias que, após D diferenças sazonais, transforma-se num processo sazonal estacionário ARMA(P,Q)_s.

$$\Phi(B^s) (1 - B^s)^D Z_t = \Theta(B^s) a_t$$

2.5

Possíveis Modelos Sazonais na Metodologia de Box & Jenkins

2.5.1

Modelo Sazonal Auto-Regressivo Integrado de Médias Móveis – SARIMA (p,d,q)x (P,D,Q)_s

O efeito sazonal implica que a observação de um determinado mês, digamos Fevereiro, está relacionada com a observação do próximo mês de Fevereiro. Supõe-se que a t-ésima observação Z_t é referente ao mês de Fevereiro; logo, a observação do mês de Fevereiro referente ao próximo ano pode ser modelada por

$$\Phi(B^s) \nabla_s^D Z_t = \Theta(B^s) \alpha_t \quad (2.2)$$

Onde:

$$s = 12;$$

$$\nabla_s = 1 - B^s;$$

$\Phi(B^s)$ e $\Theta(B^s)$ são polinômios em B^s de grau P e Q , respectivamente e satisfazem as condições de estacionariedade e invertibilidade.

Similarmente, o modelo

$$\Phi(B^s) \nabla_s^D Z_{t-1} = \Theta(B^s) \alpha_{t-1} \quad (2.3)$$

pode ser usado para analisar o comportamento do corrente mês de Janeiro com o próximo mês de Janeiro e o mesmo pode ser feito para os 12 meses.

Os componentes de erro, $\alpha_t, \alpha_{t-1}, \dots$, neste modelo geralmente não são ruído branco, isto é, existe uma correlação entre eles. Por exemplo, a venda de sorvete no mês de Fevereiro está relacionada com a venda nos anos anteriores neste mês e, poderia também, estar relacionada com a venda no mês de Janeiro. Pode-se esperar que α_t esteja relacionado com $\alpha_{t-1}, \alpha_{t-2}, \dots$. Daí, introduz-se um segundo modelo:

$$\phi(B) \nabla^d \alpha_1 = \theta(B) a_t \quad (2.4)$$

Onde:

$\phi(B) a_t$ é ruído branco;

$\nabla = \nabla_1 = 1 - B$;

$\phi(B)$ e $\theta(B)$ são polinômios em B de grau p e q , respectivamente e satisfazem as condições de estacionariedade e invertibilidade.

Substituindo (2.4) em (2.2) tem-se o modelo sazonal multiplicativo ARIMA (*multiplicative seasonal autoregressive integrated moving average model*) (Box & Jenkins, 1994), também conhecido como SARIMA (*seasonal autoregressive integrated moving average*) (Box & Jenkins, 1994), de ordem $(p,d,q) \times (P,D,Q)_s$

$$\phi_p(B) \Phi_P(B^s) \nabla^d \nabla_s^D Z_t = \theta_q(B) \Theta_Q(B^s) a_t$$

Onde:

$\phi(B)$ é o componente auto-regressivo de ordem p ;

$\theta(B)$ é o componente médias móveis de ordem q ;

$\Phi(B^s)$ é o componente sazonal auto-regressivo de ordem P ;

$\Theta(B^s)$ é o componente sazonal médias móveis de ordem Q ;

$\nabla^d = (1-B)^d$ é a diferença ordinária ;

$\nabla_s^D = (1-B^s)^D$ é a diferença sazonal;

a_t é ruído branco.

Por exemplo, um modelo SARIMA(0,1,1) \times (0,1,1)₁₂ tem a forma

$$(1-B)(1-B^{12})Z_t = (1-\theta B)(1-\Theta B^{12})a_t$$

ou
$$w_t = (1-B)(1-B^{12})Z_t = (1-\theta B - \Theta B^{12} + \theta\Theta B^{13})a_t.$$

O modelo implica que se deve tomar d diferenças simples e D diferenças sazonais da série Z_t para que o processo $w_t = \nabla^d \nabla_s^D Z_t$ seja estacionário.

A tabela a seguir apresenta as propriedades e características para a identificação teórica dos parâmetros p, q, P e Q dos modelos AR(p), MA(q), ARMA(p,q), SAR(P), SMA(Q) e SARMA(P,Q).

	AR(p)	MA(q)	ARMA(p,q)	SAR(P)	SMA(Q)	SARMA(P,Q)
Modelo expresso em termos dos w_t 's anteriores	$\phi(B)w_t = a_t$	$\theta^{-1}(B)w_t = a_t$	$\theta^{-1}(B)\phi(B)w_t = a_t$	$\Phi(B^s)w_t = a_t$	$\Theta^{-1}(B^s)w_t = a_t$	$\Theta^{-1}(B^s)\Phi(B^s)w_t = a_t$
Modelo expresso em termos dos a_t 's anteriores	$w_t = \phi^{-1}(B)a_t$	$w_t = \theta(B)a_t$	$w_t = \phi^{-1}(B)\theta(B)a_t$	$w_t = \Phi^{-1}(B^s)a_t$	$w_t = \Theta(B^s)a_t$	$w_t = \Phi^{-1}(B^s)\Theta(B^s)a_t$
Função de Autocorrelação ρ_k	Infinita (exponenciais amortecidas e/ou senóide amortecido). Não se anulam bruscamente.	Finita. Anulam-se bruscamente no lag q.	Infinita (exponenciais amortecidas e/ou senóide amortecidos para $k > q-p$). Não se anulam bruscamente.	Infinita (exponenciais amortecidas e/ou senóide amortecido). Não se anulam bruscamente.	Finita. Anulam-se bruscamente no lag Qs.	Infinita (exponenciais amortecidas e/ou senóide amortecidos para $k > Q - P$). Não se anulam bruscamente.
Função de Autocorrelação parcial ϕ_{kk}	Finita. Anulam-se bruscamente no lag p.	Infinita (dominada por exponenciais amortecidas e/ou senóide). Não se anulam bruscamente.	Infinita (dominada por exponenciais amortecidas e/ou senóide amortecidos para $k > q-p$). Não se anulam bruscamente.	Finita. Anulam-se bruscamente no lag Ps.	Infinita (domonada por exponenciais amortecidas e/ou senóide). Não se anulam bruscamente.	Infinita (exponenciais amortecidas e/ou senóide para $k > Q - P$). Não se anulam bruscamente.

Tabela 1 - Comportamento teórico dos modelos AR(p), MA(q), ARMA(p,q), SAR(P), SMA(Q) e SARMA(P,Q)

Fonte: Souza e Camargo (2004:68)

2.5.2

Exemplo de Modelo SARIMA - Modelo *Airline*

Um exemplo clássico de modelo sazonal é o modelo *Airline* – que representa o número de passageiros mensais em vôos internacionais – onde o intervalo entre as observações é de um mês e o período sazonal é $s = 12$ meses. Este modelo pode ser representado pelo modelo multiplicativo de ordem $(0,1,1) \times ((0,1,1))_{12}$ ou seja,

$$\nabla \nabla_{12} Z_t = (1 - \theta B)(1 - \Theta B^{12})a_t$$

ou de forma explícita como

$$Z_t - Z_{t-1} - Z_{t-12} + Z_{t-13} = a_t - \theta a_{t-1} - \Theta a_{t-12} - \theta \Theta a_{t-13}$$

que, para ser inversível, deve que satisfazer a condição de que as raízes de $(1 - \theta B)(1 - \Theta B^{12}) = 0$ estejam fora do círculo unitário ou seja, $-1 < \theta < 1$ e $-1 < \Theta < 1$.

Pode-se observar que o operador das médias móveis

$$(1 - \theta B)(1 - \Theta B^{12}) = 1 - \theta B - \Theta B^{12} + \theta \Theta B^{13}$$

é de ordem $q + sQ = 1 + 12 \times 1 = 13$.

2.6

Componentes Estruturais de um Modelo

Um método tradicional para representação de uma série temporal sazonal tem sido decompor a série em tendência, sazonalidade e ruído como:

$$Z_t = T_t + S_t + N_t$$

onde a tendência (T_t) e a sazonalidade (S_t) são representadas por funções determinísticas do tempo usando funções polinomiais e senóides, respectivamente, ou por funções estocásticas.

2.6.1 Sazonalidade Estocástica

Por exemplo, para os dados mensais onde a tendência pode assumir o modelo

$$(1-B)T_t = (1-\theta_t B)a_t \quad \text{ou} \quad (1-B)^2 T_t = (1-\theta_{t1}B - \theta_{t2}B^2)a_t$$

e a sazonalidade $(1-B^{12})S_t = b_t$

Onde:

a_t e b_t são ruídos branco independentes.

Estes modelos possuem um modelo ARIMA equivalente. Seja

$$Z_t = T_t + S_t + N_t$$

onde: $(1-B)T_t = (1-\theta_t B)a_t$;

$$(1-B^{12})S_t = (1-\Theta_s B^{12})b_t$$
;

$N_t = c_t$ é um ruído branco.

Daí,

$$(1-B)(1-B^{12})Z_t = (1-B^{12})(1-\theta_t B)a_t + (1-B)(1-\Theta_s B^{12})b_t + (1-B)(1-B^{12})c_t$$

Desenvolvendo o lado direito da igualdade tem-se

$$(1-B)(1-B^{12})Z_t = (1-\theta_1 B - \Theta_{12} B^{12} - \Theta_{13} B^{13})\varepsilon_t$$

onde ε_t é ruído branco e autocovariância diferente de zero nos lags 0, 1, 11, 12 e 13.

2.6.2 Sazonalidade Determinística

Algumas séries apresentam os componentes de tendência e sazonalidade determinísticos.

Quando a série temporal Z_t exibe um comportamento sazonal determinístico com período s , um modelo que pode ser útil é

$$Z_t = S_t + N_t \quad (2.5)$$

onde S_t é uma função determinística periódica, satisfazendo $S_t - S_{t-s} = 0$, ou seja,

$$(1 - B^s)S_t = 0 \quad (2.6)$$

e N_t é um processo estacionário que pode ser modelado por um modelo ARMA(p,q). Assim, N_t satisfaz a equação:

$$\phi(B)N_t = \theta(B)a_t \quad (2.7)$$

onde a_t é ruído branco e S_t tem solução geral dada por

$$S_t = \beta_0 + \sum_{j=1}^{\frac{s}{2}} \left[\beta_{1j} \cos\left(\frac{2\pi jt}{s}\right) + \beta_{2j} \text{sen}\left(\frac{2\pi jt}{s}\right) \right]$$

onde os coeficientes β são determinísticos (fixos) e

$$\frac{s}{2} = \begin{cases} \frac{1}{2}s & \text{se } s \text{ é par} \\ \frac{1}{2}(s-1) & \text{se } s \text{ é ímpar} \end{cases}.$$

Para um modelo determinístico, aplicando-se a diferença sazonal $(1 - B^s)$ à expressão (2.5), tem-se

$$(1 - B^s)Z_t = (1 - B^s)S_t + (1 - B^s)N_t$$

de (2.6), tem-se

$$(1-B^s)Z_t = (1-B^s)N_t \quad (2.8)$$

Substituindo-se (2.7) em (2.8), tem-se:

$$\phi(B)(1-B^s)Z_t = \theta(B)(1-B^s)a_t$$

ou $\phi(B)w_t = \theta(B)(1-B^s)a_t$

onde $w_t = (1-B^s)Z_t$

2.7

Testes estatísticos para verificação da validade do modelo

Para comprovar a validade do modelo selecionado, podem-se aplicar alguns testes estatísticos como, por exemplo, o teste de sobrefixação.

Segundo Souza e Camargo, “o teste de “sobrefixação” consiste, basicamente, na elaboração de um modelo com um número de parâmetros superior ao do modelo fixado, que cubra as supostas direções de discrepâncias”. (2004:119)

O erro gerado pela sobrefixação na identificação do modelo não é um problema grave pois os parâmetros extras podem ser determinados no processo de estimação. Na teoria, quando um parâmetro não é significativo, seu valor tende a zero.