

# Apostila de Análise de Séries Temporais

Autor: Prof. Manoel Ivanildo Silvestre Bezerra  
DMEC/ FCT / UNESP  
Curso de Estatística

Disciplina Semestral - 2006

# Sumário

<b>1</b>	<b>Introdução</b>	<b>4</b>
1.1	Exemplos de séries . . . . .	4
1.2	Objetivos . . . . .	5
1.3	Estacionariedade . . . . .	6
<b>2</b>	<b>Conceitos Básicos</b>	<b>7</b>
2.1	Processo Estocástico . . . . .	7
2.2	Especificação de um Processo Estocástico . . . . .	8
2.3	Médias e Covariâncias . . . . .	8
2.3.1	Propriedades da Esperança . . . . .	9
2.3.2	Propriedades da Variância . . . . .	9
2.3.3	Propriedades da Covariância . . . . .	9
2.3.4	Propriedades da Correlação . . . . .	10
2.3.5	Outras Propriedades Importantes . . . . .	10
2.4	Exemplos . . . . .	10
2.4.1	Exemplo 1: O passeio aleatório . . . . .	10
2.4.2	Exemplo 2: (Média-Móvel de ordem 1) . . . . .	12
2.5	Estacionariedade para um Processo Estocástico . . . . .	12
2.5.1	Ruído Branco . . . . .	13
2.6	A Função de Autocorrelação Amostral . . . . .	14
2.7	Exercícios . . . . .	14
<b>3</b>	<b>Modelos de Decomposição</b>	<b>17</b>
3.1	Modelo com Média constante . . . . .	17
3.2	Método de Regressão . . . . .	18
3.2.1	Tendência Linear . . . . .	18
3.2.2	Tendência Cíclica ou Sazonal . . . . .	19
3.2.3	Tendência Cosseno . . . . .	19
3.3	Sazonalidade . . . . .	19
3.3.1	Sazonalidade determinística - método de regressão . . . . .	21
3.4	Análise Residual . . . . .	22
3.5	Exercícios . . . . .	22
<b>4</b>	<b>Modelos de Suavização Exponencial</b>	<b>23</b>
4.1	Modelos para séries localmente constantes . . . . .	23
4.1.1	Médias Móveis Simples (MMS) . . . . .	23
4.1.2	Suavização Exponencial Simples (SES) . . . . .	24
4.2	Modelos para séries que apresentam Tendência . . . . .	25
4.2.1	Suavização Exponencial Biparamétrica de Holt (SEH) . . . . .	25
4.3	Modelos para séries sazonais . . . . .	25
4.3.1	Suavização Exponencial Sazonal de Holt-Winters (HW) . . . . .	25

<b>5</b>	<b>Modelos de Box-Jenkins para Séries Estacionárias</b>	<b>27</b>
5.1	Processo Linear Geral . . . . .	27
5.2	Modelo Médias-Móveis (MA(q)) . . . . .	28
5.2.1	O modelo MA(1) . . . . .	29
5.2.2	O Modelo MA(q) . . . . .	29
5.3	Modelo Autoregressivo AR(p) . . . . .	30
5.3.1	O Modelo AR(1) . . . . .	30
5.3.2	O Modelo AR(2) . . . . .	31
5.3.3	O Processo Autoregressivo geral . . . . .	33
5.4	O Modelo Autoregressivo-Médias Móveis ARMA(p,q) . . . . .	34
5.4.1	O modelo ARMA(1,1) . . . . .	34
5.5	Invertibilidade . . . . .	36
5.6	Exercícios . . . . .	37
<b>6</b>	<b>Modelos para Séries Temporais Não-Estacionárias</b>	<b>39</b>
6.1	O Modelo Autoregressivo-Integrado-Médias-Móveis ARIMA(p,d,q) . . . . .	39
6.1.1	Exemplos . . . . .	40
6.1.2	Algumas Transformações para tornar a série Estacionária . . . . .	40
6.2	Formas do Modelo ARIMA . . . . .	41
6.2.1	Forma da Equação Diferenças . . . . .	41
6.2.2	Forma de Choques Aleatórios . . . . .	41
6.2.3	Termo constante no Modelo ARIMA . . . . .	41
6.3	Construção dos Modelos ARIMA . . . . .	42
6.3.1	Função de Autocorrelação Parcial (facp) . . . . .	43
6.3.2	A Função de Autocorrelação Parcial Amostral (fapa) . . . . .	44
6.4	Identificação (ou Especificação) dos Modelos ARIMA . . . . .	44
6.4.1	Procedimento . . . . .	44
6.5	Intervalo de Confiança para a FAC Amostral . . . . .	45
6.6	Exercícios . . . . .	46
<b>7</b>	<b>Estimação dos Parâmetros</b>	<b>47</b>
7.1	Estimativas Preliminares . . . . .	47
7.1.1	Processos AR(p) . . . . .	47
7.1.2	Processos MA(q) . . . . .	47
7.1.3	Processos ARMA(p,q) . . . . .	48
7.2	O Método dos Momentos . . . . .	48
7.2.1	Modelo Autoregressivo . . . . .	48
7.2.2	Modelos Médias-Móveis . . . . .	49
7.2.3	Modelos ARMA . . . . .	49
7.2.4	Estimativas da Variância do Ruído . . . . .	49
7.3	Estimativas de Mínimos Quadrados . . . . .	50
7.3.1	Modelos Autoregressivos . . . . .	50
7.3.2	Modelos Médias Móveis . . . . .	52
7.3.3	Modelos Autoregressivos-Médias-Móveis . . . . .	53
7.4	Estimativas de Máxima Verossimilhança . . . . .	53
7.5	Mínimos Quadrados Não-Condicionais para o Modelo ARMA . . . . .	55
7.6	Propriedades das Estimativas . . . . .	55
7.7	Exercícios . . . . .	56
<b>8</b>	<b>Diagnóstico do Modelo</b>	<b>58</b>
8.1	Análise Residual . . . . .	58
8.2	Autocorrelação dos Resíduos . . . . .	59
8.3	O Teste de Box-Pierce . . . . .	59
8.4	Exercícios . . . . .	60

<b>9</b>	<b>Previsão com Modelos ARIMA</b>	<b>61</b>
9.1	Cálculo da Previsão de Erro Quadrático Médio Mínimo . . . . .	61
9.2	Tendência Determinística . . . . .	61
9.3	Previsão ARIMA . . . . .	62
9.3.1	Modelo AR(1) . . . . .	62
9.3.2	Erro de Previsão . . . . .	63
9.3.3	Modelo MA(1) . . . . .	64
9.3.4	Erro de Previsão . . . . .	64
9.3.5	O Caminho Aleatório . . . . .	65
9.3.6	Erro de Previsão . . . . .	65
9.4	Modelo ARMA Estacionário - Caso Geral . . . . .	66
9.4.1	Modelo ARMA(1,1) . . . . .	66
9.5	Modelos Não-Estacionários ARIMA . . . . .	66
9.5.1	Modelo ARIMA(1,1,1) . . . . .	67
9.5.2	Erro de Previsão . . . . .	67
9.6	Atualização das Previsões . . . . .	67
9.7	Intervalo de Confiança para Previsão . . . . .	68
9.8	Transformações e Previsões . . . . .	68
9.9	Resumo de Previsões para alguns Modelos ARIMA . . . . .	69
9.9.1	AR(1): $Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t$ . . . . .	69
9.9.2	MA(1): $Z_t = \mu + a_t - \theta a_{t-1}$ . . . . .	70
9.9.3	IMA(1,1) com termo constante: $Z_t = Z_{t-1} + \theta_0 + a_t - \theta a_{t-1}$ . . . . .	70
9.9.4	IMA(2,2): $Z_t = 2Z_{t-1} - Z_{t-2} + \theta_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$ . . . . .	70
9.10	Exercícios . . . . .	70
<b>10</b>	<b>Modelos Sazonais</b>	<b>72</b>
10.1	Sazonalidade Determinística . . . . .	72
10.2	Identificação . . . . .	73
10.3	Estimação . . . . .	73
10.4	Previsão . . . . .	73
10.5	Sazonalidade Estocástica . . . . .	74
10.6	Identificação, Estimação e Verificação . . . . .	74
10.7	Exemplos de Modelos Sazonais . . . . .	75
10.7.1	Exemplo 1 . . . . .	75
10.7.2	Exemplo 2 . . . . .	75
10.7.3	Exemplo 3 . . . . .	75
10.7.4	Exemplo 4 . . . . .	75
10.7.5	Exemplo 5 . . . . .	76
10.7.6	Exemplo 6 . . . . .	76
10.7.7	Exemplo 7 . . . . .	76
10.7.8	Exemplo 8 . . . . .	77
10.8	Exercícios . . . . .	77
<b>A</b>	<b>Esperança Condicional</b>	<b>79</b>
<b>B</b>	<b>Predição de Erro Quadrático Médio Mínimo</b>	<b>80</b>

# Capítulo 1

## Introdução

Uma série temporal é qualquer conjunto de observações ordenadas no tempo. Alguns exemplos são citados abaixo:

- a) Estimativas trimestrais do Produto Interno Bruto (PIB);
- b) Valores diários da temperatura em Presidente Prudente;
- c) Índices diários da bolsa de valores de São Paulo;
- d) Quantidade anual de chuva na cidade do Recife;
- e) Um registro de marés no porto de Santos.

Nos exemplos de a) a d) temos séries temporais discretas, enquanto que e) é um exemplo de série contínua. Podemos obter uma série temporal discreta a partir da amostragem de uma série temporal contínua considerando intervalos de tempos iguais,  $\Delta t$ . Assim para analisar a série e) será necessário amostrá-la, convertendo-a e observando-a no intervalo de tempo  $[0, T]$ , supondo uma série discreta com  $N$  pontos, onde  $N = \frac{T}{\Delta t}$  ( $T$  horas).

Existem dois enfoques utilizados na análise de séries temporais. Em ambos, o objetivo é construir modelos para estas séries. No primeiro enfoque, a análise é feita no *domínio temporal* e os modelos propostos são *modelos paramétricos* (com um número finito de parâmetros). No segundo, a análise é conduzida no *domínio de frequências* e os modelos propostos são *modelos não-paramétricos*.

Dentre os modelos paramétricos temos, por exemplo, os modelos ARIMA, que serão estudados neste curso nos próximos capítulos.

No domínio de frequências temos a *análise espectral*, que tem inúmeras aplicações em ciências físicas e engenharia, principalmente na engenharia elétrica, e que consiste em decompor a série dada em componentes de frequências e onde a existência do *espectro* é a característica fundamental. Este tipo de análise não será estudado nestas notas de aulas, para detalhes o aluno deve consultar Jenkins e Watts (1968), Koopmans (1974), Morettin (1979), Marple (1987) e Kay (1988).

### 1.1 Exemplos de séries

**Exemplo1:** Vamos supor que desejamos medir a temperatura do ar, de um local, durante 24 horas, poderíamos obter um gráfico semelhante a figura abaixo:

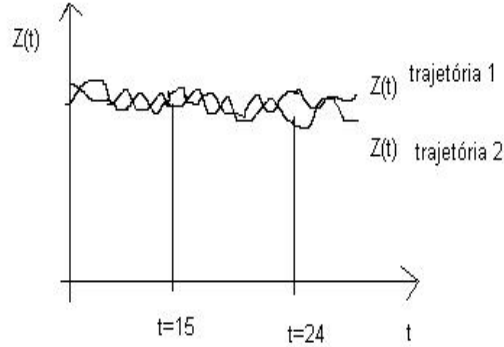


Figura 1.1: Temperatura do ar, de dado local, durante 24 horas.

Cada curva do gráfico é chamada de trajetória ou série temporal ou função amostral. No gráfico acima  $Z^{(j)}(t)$  é o valor da temperatura no instante  $t$ , para a  $j$  – ésima trajetória ( $j$ –ésimo dia de observação). Para cada  $t$  fixo, teremos os valores de uma *variável aleatória*  $Z(t)$  que terá certa *distribuição de probabilidade*. Na realidade o que chamamos de *série temporal*, é uma parte de uma trajetória, dentre muitas que poderiam ter sido observadas. O parâmetro  $t$  pode ser função de algum outro parâmetro físico como por exemplo: espaço e volume.

**Exemplo 2:**Seja

$$Z(\underline{t}) = [Z_1(\underline{t}), Z_2(\underline{t}), Z_3(\underline{t})]' \quad (1.1)$$

um vetor  $(r \times 1)$  e  $(p \times 1)$  onde as três componentes denotam, a altura, a temperatura e a pressão de um ponto no oceano, respectivamente, e  $t=(tempo, latitude, longitude)$ , ou seja, uma série multivariada ( $r=3$ ) e multidimensional ( $p=3$ ).

**Exemplo 3:** Seja  $Z(\underline{t})$  o número de acidentes ocorridos em rodovias do estado de São Paulo, por mês. Neste exemplo  $r=1$  e  $p=2$ , com  $t=(mês, rodovia)$ .

## 1.2 Objetivos

Dada uma série temporal  $\{Z(t_1), \dots, Z(t_N)\}$ , observada nos instantes  $t_1, \dots, t_N$ , podemos estar interessados em:

- i) Investigar o mecanismo gerador da série temporal;
- ii) Fazer previsões de valores futuros da série; podendo ser a curto ou longo prazo;
- iii) Descrever apenas o comportamento da série através de gráficos;
- iv) Procurar periodicidades relevantes nos dados.

Em todos estes casos podemos construir modelos probabilísticos ou estocásticos, tanto no domínio do tempo como no domínio da frequência, por exemplo: um sinal aleatório com frequência medida em Hz. Devemos construir modelos simples e com menor número de parâmetros possíveis.

### 1.3 Estacionariedade

Uma série temporal é estacionária quando ela se desenvolve aleatoriamente, no tempo, em torno de uma média constante, refletindo alguma forma de equilíbrio estável. Entretanto, a maior parte das séries que encontramos na prática apresenta alguma forma de não estacionariedade. As séries econômicas apresentam em geral tendências lineares positivas ou negativas. Podemos ter, também, uma forma de não-estacionariedade explosiva, como o crescimento de uma colônia de bactérias.

A classe dos modelos ARIMA (autoregressivos-integrados-médias-móveis), serão capaz de descrever de maneira satisfatória séries estacionárias e não-estacionárias, mas que não apresentam comportamento explosivo. Este tipo de estacionariedade é chamado homogêneo; a série pode ser estacionária, flutuando ao redor de um nível, por um certo tempo, depois mudar de nível e flutuar ao redor de um novo nível e assim por diante, ou então mudar de inclinação, ou ambas as coisas. A figura 1.2 ilustra esta forma de não-estacionariedade.

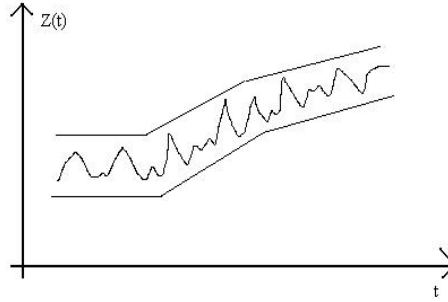


Figura 1.2: Série não estacionária quanto ao nível e inclinação.

Como a maioria dos procedimentos de análise estatística de séries temporais supõem que estas sejam estacionárias, devemos transformar os dados originais, se estes não formam uma série estacionária. A transformação mais comum consiste em tomar diferenças sucessivas da série original, até se obter uma série estacionária. A primeira diferença de  $Z_t$  é definida por

$$\Delta Z_t = Z_t - Z_{t-1}, \quad (1.2)$$

a segunda diferença é:

$$\begin{aligned} \Delta^2 Z_t &= \Delta [\Delta Z_t] = \Delta [Z_t - Z_{t-1}] \\ \Delta^2 Z_t &= \Delta Z_t - \Delta Z_{t-1} \\ \Delta^2 Z_t &= Z_t - 2Z_{t-1} + Z_{t-2} \end{aligned} \quad (1.3)$$

De modo geral, a  $n$ -ésima diferença de  $Z_t$  é  $\Delta^n Z_t = \Delta [\Delta^{n-1} Z_t]$ . Em situações normais, será suficiente tomar uma ou duas diferenças para que a série se torne estacionária.

## Capítulo 2

# Conceitos Básicos

Neste capítulo vamos descrever os conceitos básicos utilizados dentro da teoria dos modelos de séries temporais. Inicialmente vamos introduzir os conceitos de processos estocásticos, média e função de covariância, processo estacionário, e função de autocorrelação.

### 2.1 Processo Estocástico

Seja  $T$  um conjunto arbitrário. Um processo estocástico é uma família  $Z = \{Z_t, t \in T\}$  tal que, para cada  $t \in T$ ,  $Z_t$  é uma variável aleatória (v.a.) definida num espaço de probabilidades  $(\Omega, \mathcal{A}, P)$ .

O conjunto  $T$  é normalmente tomado como o conjunto dos inteiros  $\mathcal{Z} = \{0, \pm 1, \pm 2, \dots\}$  ou o conjunto dos reais  $\mathbf{R}$ . Como, para  $t \in T$ ,  $Z_t$  é uma v.a. definida sobre  $\Omega$ , na realidade  $Z_t$  é uma função de dois argumentos,  $Z(t, \omega), t \in T, \omega \in \Omega$ .

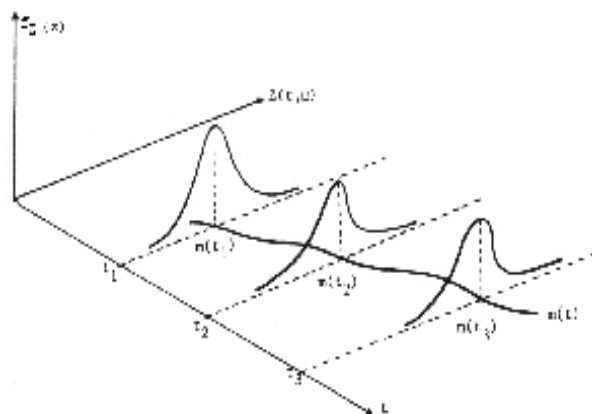


Figura 2.1: Um processo estocástico interpretado como uma família de variáveis aleatórias.



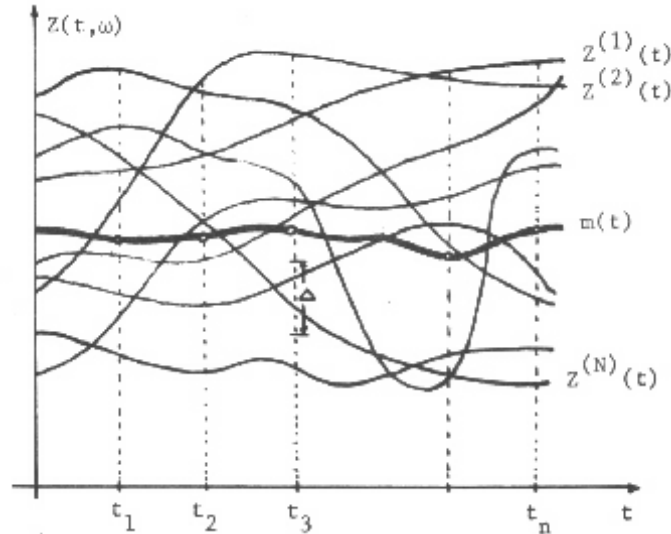


Figura 2.2: Um processo estocástico interpretado como uma família de trajetórias.

O conjunto dos valores  $\{Z_t, t \in T\}$  é chamado espaço dos estados,  $E$ , do processo estocástico, e os valores de  $Z_t$  são chamados estados. Onde  $Z_t$  e  $T$  podem ser discretos ou contínuos.

## 2.2 Especificação de um Processo Estocástico

Sejam  $t_1, t_2, \dots, t_n$  elementos quaisquer de  $T$  e consideremos

$$F(Z_1, \dots, Z_n; t_1, \dots, t_n) = P\{Z(t_1) \leq z_1, \dots, Z(t_n) \leq z_n\} \quad (2.1)$$

então, o processo estocástico  $Z = \{Z(t), t \in T\}$  estará especificado se as distribuições finito-dimensionais de (2.1), são conhecidas para todo  $n \geq 1$ .

Contudo, em termos práticos, não conhecemos todas essas distribuições finito-dimensionais. Estudaremos então certas características associadas a (2.1) e que sejam simples de calcular e interpretar.

Uma maneira de especificar o processo  $Z$  seria determinar todos os produtos dos momentos, ou seja,

$$\begin{aligned} \mu(r_1, \dots, r_n; t_1, \dots, t_n) &= E\{Z^{r_1}(t_1) \dots Z^{r_n}(t_n)\} \\ \mu(\mathbf{r}, \mathbf{t}) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} Z_1^{r_1} \dots Z_n^{r_n} f(z_1, \dots, z_n; t_1, \dots, t_n) dz_1 \dots dz_n \end{aligned} \quad (2.2)$$

onde  $f(\mathbf{Z}, \mathbf{t})$  é a função de densidade e probabilidade de  $F(\mathbf{Z}, \mathbf{t})$ . Porém o que vai nos interessar são os momentos de baixa ordem, ou seja, os chamados processos estacionários de 2ª ordem. Consideramos somente os momentos de primeira e segunda ordem, que serão apresentados a seguir.

## 2.3 Médias e Covariâncias

Para um processo estocástico  $\{Z_t : t = 0, \pm 1, \pm 2, \dots\}$  a função média (*f.m.*) é definida por

$$\mu = E(Z_t), \text{ para } t = 0, \pm 1, \pm 2, \dots \quad (2.3)$$

e a função de autocovariância (*facv*) como

$$\gamma_{t,s} = \text{Cov}(Z_t, Z_s) = E[(Z_t - \mu_t)(Z_s - \mu_s)] \text{ para } t, s = 0, \pm 1, \pm 2, \dots \quad (2.4)$$

onde  $E[(Z_t - \mu_t)(Z_s - \mu_s)] = E(Z_t Z_s) - \mu_t \mu_s$ .

A função de autocorrelação (*fac*) é dada por

$$\rho_{t,s} = \text{Corr}(Z_t, Z_s) = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t} \gamma_{s,s}}} \quad (2.5)$$

onde  $\gamma_{t,s} = \text{Cov}(Z_t, Z_s)$ ,  $\gamma_{t,t} = \text{Var}(Z_t)$  e  $\gamma_{s,s} = \text{var}(Z_s)$ .

### 2.3.1 Propriedades da Esperança

**E1.** Se  $h(x)$  é uma função tal que  $\int_{-\infty}^{\infty} |h(x)| f(x) dx < \infty$  então

$$E[h(x)] = \int_{-\infty}^{\infty} h(x) f(x) dx$$

**E2.** Se  $X$  e  $Y$  tem fdp conjunta  $f(x, y)$  e  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(x, y)| f(x, y) dx dy < \infty$ , então

$$E[h(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy$$

**E3.** Como um corolário para E2 nós temos o importante resultado

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

**E4.** Temos também que

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dx dy.$$

Temos ainda que, a variância da variável aleatória  $X$  é definida como  $\sigma_X^2 = \text{Var}(X) = E[X - E(X)]^2$ .

### 2.3.2 Propriedades da Variância

**V1.**  $\text{Var}(X) \geq 0$ .

**V2.**  $\text{Var}(a + bX) = b^2 \text{Var}(X)$ .

**V3.** Se  $X$  e  $Y$  são independentes, então  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ .

**V4.**  $\text{Var}(X) = E(X^2) - [E(X)]^2$  e  $\sqrt{\text{Var}(X)} = \sqrt{\sigma_X^2} = \sigma_X$  (desvio-padrão de  $X$ ).

### 2.3.3 Propriedades da Covariância

**CV1.** Se  $X$  e  $Y$  são independentes, então  $\text{Cov}(X, Y) = 0$ .

**CV2.**  $\text{Cov}(a + bX, c + dY) = bd \text{Cov}(X, Y)$ .

**CV3.**  $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$ .

**CV4.**  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$ .

**CV5.**  $\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$ .

**CV6.**  $\text{Cov}(X, X) = \text{Var}(X)$ .

**CV7.**  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$ .

### 2.3.4 Propriedades da Correlação

**CR1.**  $-1 \leq \text{Corr} \leq 1$ .

**CR2.**  $\text{Corr}(a + bX, c + dY) = \sin al(bd) \text{Corr}(X, Y)$  onde

$$\sin al(bd) = \begin{cases} 1 & \text{se } bd > 0 \\ 0 & \text{se } bd = 0 \\ -1 & \text{se } bd < 0 \end{cases}$$

**CR3.**  $\text{Corr}(X, Y) = \text{Cov}\left(\frac{X - \mu_X}{\sigma_X}, \frac{Y - \mu_Y}{\sigma_Y}\right)$ .

**CR4.**  $\text{Corr}(X, Y) = \pm 1$  se e somente se existem constantes  $a$  e  $b$  tal que  $P(Y = a + bX) = 1$ .

### 2.3.5 Outras Propriedades Importantes

i)  $\gamma_{t,t} = \text{Var}(Z_t)$ ,  $\rho_{t,t} = 1$

ii)  $\gamma_{t,s} = \gamma_{s,t}$ ,  $\rho_{t,s} = \rho_{s,t}$

iii)  $|\gamma_{t,s}| \leq \sqrt{\gamma_{t,t}\gamma_{s,s}}$ ,  $|\rho_{t,s}| \leq 1$ , ou  $-1 \leq \rho_{t,s} \leq 1$ .

Na correlação podemos verificar que valores próximos de  $\pm 1$  indicam forte dependência (linear) e valores próximos de 0 indicam fraca dependência (linear). Se  $\rho_{t,s} = 0$ ,  $Z_t$  e  $Z_s$  são não-correlacionadas. Agora se  $Z_t$  e  $Z_s$  são independentes, então  $\rho_{t,s} = 0$ .

Para analisar as propriedades da covariância de vários modelos de séries temporais, o seguinte resultado será utilizado: se  $c_1, c_2, \dots, c_m$  e  $d_1, d_2, \dots, d_n$  são constantes e  $t_1, t_2, \dots, t_m$  e  $s_1, s_2, \dots, s_n$  são pontos no tempo, então

$$\text{Cov}\left[\sum_{i=1}^m c_i Z(t_i), \sum_{j=1}^n d_j Z(s_j)\right] = \sum_{i=1}^m \sum_{j=1}^n c_i d_j \text{Cov}[Z(t_i), Z(s_j)] \quad (2.6)$$

podemos dizer que, a covariância entre duas combinações lineares é a soma de todas as covariâncias entre termos de suas combinações lineares. Esta expressão pode ser verificada utilizando as propriedades de esperança e covariância. Como caso especial, podemos obter o seguinte resultado

$$\text{Var}\left[\sum_{i=1}^n c_i Z(t_i)\right] = \sum_{i=1}^n c_i^2 \text{Var}[Z(t_i)] + 2 \sum_{i=2}^n \sum_{j=1}^{i-1} c_i c_j \text{Cov}[Z(t_i), Z(t_j)] \quad (2.7)$$

## 2.4 Exemplos

### 2.4.1 Exemplo 1: O passeio aleatório

Seja  $a_1, a_2, \dots$  variáveis aleatórias independentes, identicamente distribuídas, cada uma com média 0 e variância  $\sigma_a^2$ . A série temporal,  $Z_t$ , é construída da seguinte maneira:

$$\begin{aligned} Z_1 &= a_1 \\ Z_2 &= a_1 + a_2 \\ &\vdots \\ Z_t &= a_1 + a_2 + \dots + a_t \end{aligned} \quad (2.8)$$

ou

$$Z_t = Z_{t-1} + a_t \quad (2.9)$$

Obtendo a função média de (2.8) temos:

$$\begin{aligned} \mu_t &= E(Z_t) = E(a_1 + a_2 + \dots + a_t) \\ &= E(a_1) + E(a_2) + \dots + E(a_t) \\ &= 0 + 0 + \dots + 0 = 0 \end{aligned}$$

como  $E(a_t) = 0$ , nós temos:

$$\boxed{\mu_t = 0 \text{ p/ todo } t}$$

e também

$$\begin{aligned} \text{Var}(Z_t) &= \text{Var}(a_1 + a_2 + \dots + a_t) \\ &= \text{Var}(a_1) + \text{Var}(a_2) + \dots + \text{Var}(a_t) \\ &= \sigma_a^2 + \sigma_a^2 + \dots + \sigma_a^2 = t\sigma_a^2 \end{aligned}$$

ou

$$\boxed{\text{Var}(Z_t) = t\sigma_a^2}$$

Observe que a variância do processo cresce linearmente com o tempo. Suponha agora que  $1 \leq t \leq s$ , teremos então

$$\begin{aligned} \gamma_{t,s} &= \text{Cov}(Z_t, Z_s) \\ &= \text{Cov}(a_1 + a_2 + \dots + a_t, a_1 + a_2 + \dots + a_s) \\ &= \text{Cov}(a_1, a_1) + \text{Cov}(a_2, a_2) + \dots + \text{Cov}(a_t, a_t) \\ &= \sigma_a^2 + \sigma_a^2 + \dots + \sigma_a^2 = t\sigma_a^2 \end{aligned}$$

onde

$$\text{Cov}(a_t, a_s) = 0 \text{ para } t \neq s$$

temos então que a cov é dada por

$$\boxed{\gamma_{t,s} = t\sigma_a^2, \text{ para } 1 \leq t \leq s}$$

e a cov é dada por

$$\boxed{\rho_{t,s} = \sqrt{\frac{t}{s}}, \text{ para } 1 \leq t \leq s}$$

O passeio aleatório é um exemplo simples que representa diversos fenômenos como o movimento comum de preços e títulos e também a posição de pequenas partículas suspensas dentro de um fluido, chamado movimento Browniano.

### 2.4.2 Exemplo 2: (Média-Móvel de ordem 1)

Suponha

$$Z_t = a_t - 0.5a_{t-1}$$

onde  $a_t$  são v.a.i.i.d. com média zero e variância  $\sigma_a^2$ .

$$\begin{aligned}\mu_t &= E(Z_t) \\ &= E(a_t) - 0.5E(a_{t-1}) = 0\end{aligned}$$

e

$$\begin{aligned}\text{Var}(Z_t) &= \text{Var}(a_t - 0.5a_{t-1}) \\ &= \sigma_a^2 + 0.5^2\sigma_a^2 = 1.25\sigma_a^2.\end{aligned}$$

Também

$$\begin{aligned}\text{Cov}(Z_t, Z_{t-1}) &= \text{Cov}(a_t - 0.5a_{t-1}, a_{t-1} - 0.5a_{t-2}) \\ &= -0.5\text{Cov}(a_{t-1}, a_{t-1})\end{aligned}$$

ou

$$\gamma_{t,t-1} = -0.5\sigma_a^2$$

Além disso

$$\text{Cov}(Z_t, Z_{t-k}) = 0 \text{ para } k \geq 2$$

então podemos concluir que

$$\gamma_{t,s} = \begin{cases} -0.5\sigma_a^2 & \text{se } |t-s| = 1 \\ 0 & \text{se } |t-s| > 1 \end{cases}$$

Para  $\rho_{t,s}$  temos

$$\rho_{t,s} = \begin{cases} -0.4 & \text{se } |t-s| = 1 \\ 0 & \text{se } |t-s| > 1 \end{cases}$$

## 2.5 Estacionariedade para um Processo Estocástico

A estacionariedade é a suposição mais importante para um processo estocástico, com relação ao estudo que faremos. A idéia básica de estacionariedade é que as leis de probabilidade que atuam no processo não mudam com o tempo, isto é, o processo mantém o equilíbrio estatístico.

Especificamente um processo estocástico  $Z(t)$  é considerado *estritamente estacionário* se a distribuição conjunta de  $Z(t_1), Z(t_2), \dots, Z(t_n)$  é a mesma distribuição conjunta de  $Z(t_1 - k), Z(t_2 - k), \dots, Z(t_n - k)$ , para todos os tempos  $t_1, t_2, \dots, t_n$  e para todos os “lags” (posições)  $k$  (constante).

Quando  $n = 1$ , a distribuição de  $Z_t$  é igual a distribuição de  $Z_{t-k}$  para qualquer  $k$ , ou seja, se os  $Z$ 's são identicamente distribuídos,  $E(Z_t) = E(Z_{t-k})$ , para todo  $t$  e  $k$ , e as funções média,  $\mu_t$ , e variância  $\text{Var}(Z_t) = \text{Var}(Z_{t-k})$  são constantes para todo tempo  $t$ .

Quando  $n = 2$ , a distribuição de  $(Z_t, Z_s)$  é a mesma de  $(Z_{t-k}, Z_{s-k})$ , do qual temos  $Cov(Z_t, Z_s) = Cov(Z_{t-k}, Z_{s-k})$  para todo  $t, s$  e  $k$ .

Fazendo  $k = s$  temos:

$$\begin{aligned}\gamma_{t,s} &= Cov(Z_t, Z_s) = Cov(Z_{t-k}, Z_{s-k}) \\ &= Cov(Z_{t-s}, Z_{s-s}) = Cov(Z_{t-s}, Z_0) \\ &= \gamma_{t-s,0}\end{aligned}$$

e agora  $k = t$

$$\begin{aligned}\gamma_{t,s} &= Cov(Z_{t-t}, Z_{s-t}) = Cov(Z_0, Z_{s-t}) \\ &= Cov(Z_0, Z_{t-s}) \\ &= \gamma_{0,s-t}\end{aligned}$$

daí podemos concluir que

$$\gamma_{t,s} = \gamma_{0,|t-s|}, \text{ onde } |t-s| = \begin{cases} t-s & \text{p/ } t > s \\ s-t & \text{p/ } s > t \end{cases}$$

A covariância entre  $Z_t$  e  $Z_s$  depende somente da diferença temporal  $|t-s|$  e não dos tempos  $t$  e  $s$ . Além disso, para um *processo estacionário* simplificando a notação temos

$$\begin{aligned}\gamma_k &= Cov(Z_t, Z_{t-k}) \text{ e} \\ \rho_k &= Corr(Z_t, Z_{t-k}) = \frac{\gamma_k}{\gamma_0}\end{aligned} \tag{2.10}$$

As propriedades gerais para um *processo estacionário* são:

$$\begin{aligned}i) \gamma_0 &= Var(Z_t), \rho_0 = 1 \\ ii) \gamma_k &= \gamma_{-k}, \rho_k = \rho_{-k} \\ iii) |\gamma_k| &\leq \gamma_0, |\rho_k| \leq 1.\end{aligned} \tag{2.11}$$

Se um processo é estritamente estacionário e tem variância finita, então a facv depende somente de um certo lag  $k$ .

Uma definição que é semelhante a estritamente estacionária mas é matematicamente mais fraca, é a seguinte: um processo estocástico  $Z_t$  é dito ser *fracamente (ou de segunda-ordem) estacionário* se:

- i) a função média é constante p/ todo tempo  $t$ ,
- ii)  $\gamma_{t,t-k} = \gamma_{0,k}$  p/ todo tempo  $t$  e de “lag”  $k$ .

### 2.5.1 Ruído Branco

Um importante exemplo de processo estacionário é o ruído branco, o qual é definido como uma sequência de variáveis aleatórias independente, identicamente distribuídas  $\{a_t\}$ . Muitos processos podem ser construídos a partir do ruído branco. Pode-se verificar facilmente que a sequência  $\{a_t\}$  é estritamente estacionária

$$\begin{aligned}P[a(t_1) \leq x_1, a(t_2) \leq x_2, \dots, a(t_n) \leq x_n] \\ = P[a(t_1) \leq x_1] P[a(t_2) \leq x_2] \dots P[a(t_n) \leq x_n]\end{aligned}$$

$$\begin{aligned}
 & \text{são variáveis aleatórias independentes} \\
 & = P[a(t_1 - k) \leq x_1] P[a(t_2 - k) \leq x_2] \dots P[a(t_n - k) \leq x_n] \\
 & \text{identicamente distribuídas} \\
 & = P[a(t_1 - k) \leq x_1, a(t_2 - k) \leq x_2, \dots, a(t_n - k) \leq x_n].
 \end{aligned}$$

Temos também que  $\mu_t = E(a_t)$  é constante com *facv* dada por

$$\gamma_{t,s} = \begin{cases} \text{Var}(a_t) & \text{se } t = s \\ 0 & \text{se } t \neq s \end{cases}$$

e *fac* dada por

$$\rho_k = \begin{cases} 1 & \text{para } k = 0 \\ 0 & \text{para } k \neq 0 \end{cases}$$

O termo ruído branco resulta do fato que em uma análise de frequência do modelo, podemos mostrar que todas as frequências são iguais. Geralmente assumimos que o ruído branco tem média zero e variância  $\sigma_a^2$ .

## 2.6 A Função de Autocorrelação Amostral

Considere uma sequência de valores  $Z_1, Z_2, \dots, Z_t$ . Assumindo que esta série seja estacionária, vamos estimar a função de autocorrelação  $\rho_k$  (*fac*) para  $k = 1, 2, \dots, N-1$ . O caminho mais indicado é calcular as correlações amostrais entre os pares  $(Z_1, Z_{k+1}), (Z_2, Z_{k+2}), \dots, (Z_{t-k}, Z_t)$ . A *fac* amostral  $r_k$  é definida como:

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, 2, \dots, N-1$$

onde

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z}), \quad k = 0, 1, \dots, N-1.$$

então

$$\hat{\rho}_k = r_k = \frac{\sum_{t=1}^{N-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2}, \quad k = 0, 1, \dots, N-1.$$

sendo  $\bar{Z}$  a média amostral, temos ainda que  $c_{-k} = c_k$  e  $r_{-k} = r_k$ . O gráfico de  $r_k \times k$  é chamado de *correlograma*.

## 2.7 Exercícios

1) Suponha  $E(X) = 2$ ,  $\text{Var}(X) = 9$ ,  $E(Y) = 0$ ,  $\text{Var}(Y) = 4$ , e  $\text{Corr}(X, Y) = 1/4$ . Encontre:

- $\text{Var}(X + Y)$
- $\text{Cov}(X, X + Y)$
- $\text{Corr}(X + Y, X - Y)$

2) Se  $X$  e  $Y$  são dependentes mas  $\text{Var}(X) = \text{Var}(Y)$ , encontre  $\text{Cov}(X + Y, X - Y)$ .

3) Seja  $X$  uma variável aleatória com média  $\mu$  e variância  $\sigma^2$ , e seja  $Z_t = X$  para todo  $t$ .

- Mostrar que  $Z_t$  é estritamente e fracamente estacionária.

- b) Encontre a função de autocovariância de  $Z_t$ .  
 c) Fazer o gráfico de  $Z_t$  em função de  $t$ .  
 4) Mostre que a função de autocorrelação do exemplo 1 é:

$$\rho_{t,s} = \sqrt{\frac{t}{s}}, \text{ para } 1 \leq t \leq s$$

5) Seja  $\{a_t\}$  um ruído branco com média zero. Encontre a função de autocorrelação para os seguintes processos:

- a)  $Z_t = a_t + 1/3a_{t-1}$ .  
 b)  $Z_t = a_t + 2a_{t-1} + 0.5$ .

6) Suponha  $Z_t = 5 + 2t + X_t$ , onde  $X_t$  é uma série estacionária com média zero e facv  $\gamma_k$ . Encontre:

- a) A função média para  $Z_t$ .  
 b) A facv para  $Z_t$ .  
 c)  $Z_t$  é uma série estacionária?

7) Suponha  $Z_t$  uma série estacionária com função de autocovariância  $\gamma_k$ .

a) Mostrar que  $W_t = \Delta Z_t = Z_t - Z_{t-1}$  é estacionária e encontre a média e a facv para  $W_t$ .

b) Mostrar que  $U_t = \Delta W_t = \Delta^2 Z_t$  é estacionária (não precisa encontrar necessariamente a facv).

8) Seja  $X_t$  uma série estacionária com  $\mu_X = 0$ ,  $\sigma^2 = 1$  e  $\rho_k = \rho$  (média, variância e fac, respectivamente). Suponha que  $\mu_t$  e  $\sigma_t > 0$ , funções não-constantes. A série observada é formada como:

$$Z_t = \mu_t + \sigma_t X_t$$

Para a série  $Z_t$  determine: a) A média, a variância e a facv.

b) Mostrar que a fac de  $Z_t$  depende somente de um “lag”  $k$ .  $Z_t$  é estacionária?

9) Suponha  $Cov(X_t, X_{t-k}) = \gamma_k$  que não depende de  $t$ , mas  $E(X_t) = 3t$ .

a)  $X_t$  é estacionária?

b) Seja  $Z_t = 7 - 3t + X_t$ .  $Z_t$  é estacionária?

10) Suponha  $Z_t = A + X_t$ , onde  $X_t$  é estacionária e  $A$  é uma variável aleatória, com média  $\mu_A$  e variância  $\sigma_A^2$ , mas independente de  $X_t$ . Encontre:

- a) A média de  $Z_t$ , em termos das médias de  $A$  e  $X_t$ .  
 b) A fac de  $Z_t$ , em termos da variância de  $A$  e facv de  $X_t$ .

11) Suponha  $Z_t = \beta_0 + \beta_1 t + X_t$  onde  $X_t$  é estacionária.

a) Mostrar que  $Z_t$  não é estacionária, mas  $\Delta Z_t$  é estacionária.

b) Generalizando, mostrar que se  $Z_t = \mu_t + X_t$  é não-estacionária, onde  $\mu_t$  é um polinômio em  $t$  de grau  $d$ , então  $\Delta^m Z_t$  é estacionária para  $m \geq d$  e não estacionária para  $0 \leq m < d$ .

12) Seja  $Z_t$  uma série estacionária com facv  $\gamma_k$ . Seja  $\bar{Z} = \frac{1}{n} \sum_{t=1}^n Z_t$

Mostrar que

$$\begin{aligned} Var(\bar{Z}) &= \frac{1}{n} \sum_{k=-n+1}^{n-1} \left(1 - \frac{|k|}{n}\right) \gamma_k \\ &= \frac{\gamma_0}{n} + \frac{2}{n} \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \gamma_k \end{aligned}$$

13) Calcular a média e a facv para cada um dos seguintes processos. Dentro de cada caso determine se o processo é estacionário.

- a)  $Z_t = \theta_0 + ta_t$ .  
 b)  $W_t = \Delta Z_t$ , onde  $Z_t$  é definido como em a)  
 c)  $Z_t = a_t \cdot a_{t-1}$ .



**14)** Considere a série temporal abaixo:

ANO	Produto Interno Bruto
1964	27.614
1965	44.073
1966	63.746
1967	86.171
1968	122.430
1969	161.900
1970	208.300
1971	276.807
1972	363.167
1973	498.307
1974	719.519

Dados em milhões de Cruzeiros.

FONTE: Boletim do Banco do Brasil, dezembro de 1976

- Faça o gráfico da série; ela é estacionária?
- Obtenha a primeira diferença da série e faça o gráfico correspondente; a diferença é estacionária?

c) Mesmas questões de b) para a segunda diferença.

**15)** Considere a série abaixo:

Exportação de suco concentrado de laranja

ANO	1970	1971	1972	1973	1974	1975
Valor	15	35	41	64	60	82

ANO	1976	1977	1978	1979	1980
Valor	101	177	333	432	460

Dados em US\$ 1.000.000

FONTE: Revista Veja, nº 468, Fev. de 1981.

- A série é estacionária? Tem tendência?
- Considere a diferença  $\Delta Z_t$ ; é estacionária?
- Tome agora  $\log Z_t$ ; a série é estacionária?
- Investigue se a série  $\Delta \log Z_t$  é estacionária ou não.

## Capítulo 3

# Modelos de Decomposição

Nas séries temporais, em geral, a função média é totalmente aleatória em função do tempo, e nós vimos que numa série estacionária a função média é constante para todo o tempo. Frequentemente nós precisamos modelar séries que exibem comportamentos com alguma tendência. Neste capítulo apresentaremos alguns métodos de modelagem para tendências determinísticas e estocásticas.

### 3.1 Modelo com Média constante

Considere o modelo

$$Z_t = \mu + X_t \quad (3.1)$$

onde  $E(X_t) = 0$  para todo  $t$ . Desejamos estimar  $\mu$  utilizando a série  $Z_1, Z_2, \dots, Z_n$ . A estimativa mais comum de  $\mu$  é a média amostral:

$$\hat{\mu} = \bar{Z} = \frac{1}{n} \sum_{t=1}^n Z_t \quad (3.2)$$

O  $\hat{\mu}$  é um estimador não-viciado de  $\mu$ , ou seja,  $E(\bar{Z}) = \mu$ . Para analisar a precisão de  $\bar{Z}$  como estimativa de  $\mu$ , devemos fazer algumas suposições sobre  $X_t$ . Vamos supor  $\{X_t\}$  uma série estacionária com fac  $\rho_k$ . Então a mesma fac aplica-se a  $\{Z_t\}$ , ou seja, se  $\{Z_t\}$  é estacionária então:

$$\begin{aligned} Var(\bar{Z}) &= \frac{\gamma_0}{n} \sum_{k=-n+1}^{n-1} \left(1 - \frac{|k|}{n}\right) \rho_k \\ Var(\bar{Z}) &= \frac{\gamma_0}{n} \left[1 + 2 \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right) \rho_k\right]. \end{aligned} \quad (3.3)$$

Se a série  $\{X_t\}$  é um ruído branco, então  $\rho_k = 0$  para  $k \geq 1$  e  $Var(\bar{Z}) = \frac{\gamma_0}{n} = \frac{\sigma^2}{n}$ . Para muitos processos estacionários, a fac decai rapidamente quando o  $k$  cresce, além disso

$$\sum_{k=0}^{\infty} \rho_k < \infty. \quad (3.4)$$

Com exceção da variável

$$Z_t = \cos \left[ 2\pi \left( \frac{t}{12} + \Phi \right) \right], t = 0, \pm 1, \pm 2, \dots$$

onde  $\Phi \sim U[0, 1]$ . Sobre a suposição (3.4) acima e dado que o tamanho da amostra  $n$  é grande, temos a seguinte aproximação para (3.3)

$$\text{Var}(\bar{Z}) \approx \frac{\gamma_0}{n} \sum_{k=-\infty}^{\infty} \rho_k, \text{ para } n \text{ grande.} \quad (3.5)$$

Note que a variância é inversamente proporcional ao tamanho da amostra  $n$ .

## 3.2 Método de Regressão

O método clássico de análise de regressão pode ser utilizado para estimar os parâmetros do modelo assim como suas tendências. Nós vamos considerar precisamente as tendências: linear, quadrática, sazonais e cosseno.

### 3.2.1 Tendência Linear

Considere o modelo com tendência determinística linear

$$Z_t = \beta_0 + \beta_1 t + X_t \quad (3.6)$$

onde  $\beta_0$  e  $\beta_1$  são o intercepto e a inclinação, respectivamente, e os parâmetros desconhecidos. O método clássico de mínimos quadrados (ou regressão) é utilizado para estimar os parâmetros  $\beta_0$  e  $\beta_1$ , que minimizam

$$Q(\beta_0, \beta_1) = \sum_{t=1}^n (Z_t - \beta_0 - \beta_1 t)^2. \quad (3.7)$$

A solução pode ser obtida calculando as derivadas parciais com relação a  $\beta_0$  e  $\beta_1$  igualando-as a zero, encontrando as equações normais

$$\begin{aligned} N\beta_0 + \beta_1 \sum_{t=1}^n t &= \sum_{t=1}^n Z_t \\ \beta_0 \sum_{t=1}^n t + \beta_1 \sum_{t=1}^n t^2 &= \sum_{t=1}^n t Z_t. \end{aligned} \quad (3.8)$$

Daí encontraremos os seguintes estimadores de mínimos quadrados:

$$\begin{aligned} \hat{\beta}_0 &= \bar{Z} - \hat{\beta}_1 \bar{t} \\ \hat{\beta}_1 &= \frac{\sum_{t=1}^n (Z_t - \bar{Z})(t - \bar{t})}{\sum_{t=1}^n (t - \bar{t})^2} \end{aligned} \quad (3.9)$$

onde  $\bar{t} = (n+1)/2$ ,  $t = 1, 2, \dots, n$ . Uma forma matricial de representar o problema seria:

$$\mathbf{Z} = \mathbf{Y}\boldsymbol{\beta} + \mathbf{X} \quad (3.10)$$

onde

$$\mathbf{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_t \end{bmatrix}, \mathbf{Y} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & t \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \text{ e } \mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_t \end{bmatrix} \quad (3.11)$$

com  $t = 1, 2, \dots, n$ .

teremos então o seguinte estimador de mínimos quadrados

$$\hat{\boldsymbol{\beta}} = (\mathbf{Y}^T \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{Z}. \quad (3.12)$$

### 3.2.2 Tendência Ciclica ou Sazonal

Vamos considerar o seguinte modelo para a tendência sazonal

$$Z_t = \mu_t + X_t \quad (3.13)$$

onde  $\mu_t = \mu_{t+12}$ , e  $E(X_t) = 0$  para todo  $t$ . A suposição mais geral para  $\mu_t$  é que temos 12 constantes (parâmetros)  $\beta_1, \beta_2, \dots, \beta_{12}$ , que poderemos escrever como:

$$\mu_t = \begin{cases} \beta_1, & \text{para } t = 1, 13, 25, \dots \\ \beta_2, & \text{para } t = 2, 14, 26, \dots \\ \vdots & \\ \beta_{12}, & \text{para } t = 12, 24, 36, \dots \end{cases} \quad (3.14)$$

Sendo este o modelo de médias sazonais.

### 3.2.3 Tendência Cosseno

O modelo de médias sazonais contém 12 parâmetros independentes, mas algumas vezes não é adequado para estimar tendências sazonais quando, por exemplo, alguns meses do ano são semelhantes. Em alguns casos a tendência sazonal pode ser modelada economicamente com curvas cossenos que incorpora uma esperada, e suave, mudança de tempos em tempos entretanto preservando a sazonalidade.

Considere a curva cosseno

$$\mu_t = \beta \cos(2\pi ft + \Phi)$$

onde  $\beta$  é a amplitude,  $f$  a frequência, e  $\Phi$  a fase da curva. Como  $t$  varia, a curva oscila entre um máximo de  $\beta$  e um mínimo de  $-\beta$ . Desde que a curva se repete, exatamente a cada  $1/f$  unidade de tempos,  $1/f$  é o período. Para dados mensais, a mais importante frequência é  $f = \frac{1}{12}$ , na qual a curva cosseno se repetirá a cada 12 meses.

## 3.3 Sazonalidade

Nesta seção será ajustada uma série para a componente sazonal, ou seja, estima-se a série e subtrai-se de  $Z_t$  no modelo. A seguir tem-se um modelo com tendência e sazonalidade

$$Z_t = T_t + S_t + a_t, \quad t = 1, \dots, N. \quad (3.15)$$

Um procedimento de ajuste sazonal consiste:

- a) Obter estimativas  $\hat{S}_t$  de  $S_t$ .

Tabela 3.1: Observações mensais de uma série temporal com  $p$  anos

Anos	Jan	Fev	Mar	...	Dez	Médias
	1	2	3	...	12	
1	$Z_{11}$	$Z_{12}$	$Z_{13}$	...	$Z_{112}$	$\bar{Z}_{1.}$
2	$Z_{21}$	$Z_{22}$	$Z_{23}$	...	$Z_{212}$	$\bar{Z}_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$p$	$Z_{p1}$	$Z_{p2}$	$Z_{p3}$	...	$Z_{p12}$	$\bar{Z}_{p.}$
Médias	$\bar{Z}_{.1}$	$\bar{Z}_{.2}$	$\bar{Z}_{.3}$	...	$\bar{Z}_{.12}$	$\bar{Z}$

b) Calcular

$$Z_t^{SA} = Z_t - \hat{S}_t \quad (3.16)$$

Se o modelo for multiplicativo, da forma

$$Z_t = T_t S_t a_t \quad (3.17)$$

a série sazonalmente ajustada será

$$Z_t^{SA} = Z_t / \hat{S}_t \quad (3.18)$$

O modelo (3) geralmente é adequado para séries econômicas, que apresentam um crescimento exponencial. Estimando  $S_t$  comete-se um erro de ajustamento sazonal, dado por:

$$\xi_t = S_t - \hat{S}_t$$

o procedimento de ajustamento sazonal é ótimo se minimizar  $E(\xi_t^2)$ .

Sem perda de generalidade considere o caso que tem-se dados mensais e o número total de observações,  $N$ , é múltiplo de 12, isto é,  $N = 12p$ ,  $p$  = número de anos, de modo que os dados podem ser representados como na tabela 3. Considere o modelo abaixo

$$Z_{ij} = T_{ij} + S_j + a_{ij}, \quad i = 1, \dots, p; \quad j = 1, \dots, 12.$$

neste modelo a sazonalidade é considerada constante. Para sazonalidade não-constante, ou seja, o padrão sazonal muda de ano para ano, deve-se considerar o modelo

$$Z_{ij} = T_{ij} + S_{ij} + a_{ij}, \quad i = 1, \dots, p; \quad j = 1, \dots, 12.$$

A notação da tabela 3. é padrão com:

$$\bar{Z}_{i.} = \frac{1}{12} \sum_{j=1}^{12} Z_{ij}, \quad i = 1, \dots, p.$$

$$\bar{Z}_{.j} = \frac{1}{p} \sum_{i=1}^p Z_{ij}, \quad j = 1, \dots, 12.$$

$$\bar{Z} = \frac{1}{12p} \sum_{j=1}^{12} \sum_{i=1}^p Z_{ij} = \frac{1}{N} \sum_{t=1}^N Z_t$$

### 3.3.1 Sazonalidade determinística - método de regressão

Os métodos de regressão são ótimos para séries que apresentam sazonalidade determinística, ou seja, esta pode ser prevista perfeitamente a partir de meses anteriores. Então, no modelo 3. tem-se que

$$T_t = \sum_{j=0}^m \beta_j t^j;$$

$$S_t = \sum \alpha_j d_{jt}$$

onde  $d_{jt}$  são variáveis periódicas e  $a_t$  é ruído branco, com média zero e variância  $\sigma_a^2$ .

Supondo a sazonalidade constante,  $\alpha_j$  não depende de  $T$ . Pode-se ter, por exemplo,

$$d_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, j = 1, \dots, 12. \\ 0, & \text{caso contrário.} \end{cases}$$

Neste caso,

$$d_{1t} + d_{2t} + \dots + d_{12,t} = 1, \quad t = 1, \dots, N.$$

de modo que a matriz de regressão não é de posto completo, mas de posto  $m + 12$  (temos  $m + 13$  parâmetros). Impondo-se a restrição adicional:

$$\sum_{j=1}^{12} \alpha_j = 0$$

obtém-se um modelo de posto completo:

$$Z_t = \sum_{j=0}^m \beta_j t^j + \sum_{j=1}^{11} \alpha_j D_{jt} + a_t$$

onde agora

$$D_{jt} = \begin{cases} 1, & \text{se o período } t \text{ corresponde ao mês } j, \\ -1, & \text{se o período } t \text{ corresponde ao mês } 12 \\ 0, & \text{caso contrário.} \end{cases}$$

Utilizando o método de mínimos quadrados pode-se obter os estimadores de  $\alpha_j$  e  $\beta_j$ , ou seja, a partir de uma amostra  $Z_1, \dots, Z_N$  obtém-se o modelo matricial:

$$\mathbf{Z} = \mathbf{C}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\alpha} + \mathbf{a} \quad (3.19)$$

onde

$$\mathbf{Z} = \begin{bmatrix} Z_1 \\ \vdots \\ Z_N \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & 2^m \\ \vdots & \vdots & \ddots & \vdots \\ 1 & N & \dots & N^m \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} D_{11} & D_{12} & \dots & D_{1,11} \\ D_{21} & D_{22} & \dots & D_{2,11} \\ \vdots & \vdots & \ddots & \vdots \\ D_{N1} & D_{N2} & \dots & D_{N,11} \end{bmatrix}, \quad \boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{11} \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{bmatrix}$$

A equação (3.) pode ser escrita na forma:

$$\mathbf{Z} = \mathbf{X}\boldsymbol{\gamma} + \mathbf{a}$$

onde

$$\mathbf{X} = [\mathbf{C} : \mathbf{D}] \text{ e } \boldsymbol{\gamma} = \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\alpha} \end{bmatrix}$$

de modo que:

$$\hat{\boldsymbol{\gamma}} = [\mathbf{X}'\mathbf{X}]^{-1} \mathbf{X}'\mathbf{Z}$$

são os estimadores de mínimos quadrados.

### 3.4 Análise Residual

Podemos notar que, o componente estocástico não observado  $\{X_t\}$  pode ser estimado, ou predito, por

$$\hat{X}_t = Z_t - \hat{\mu}_t$$

Predito é um termo realmente melhor. Nós utilizamos o termo *estimativa* para um parâmetro desconhecido, e o termo predito para uma estimativa de uma variável aleatória não observada. Nós chamamos  $\hat{X}_t$  o *t-ésima* estimativa residual. Se a tendência do modelo é razoavelmente correta, então os resíduos comportam-se mais ou menos igual ao verdadeiro componente estocástico, e várias suposições sobre o componente estocástico podem ser confirmadas observando os resíduos. Se o componente estocástico é ruído branco, então os resíduos comportam-se semelhante a variáveis aleatórias independentes (normal) com média zero e desvio padrão  $S$ . Desde que o ajuste de mínimos quadrados de alguma tendência contenha um termo constante, automaticamente podemos produzir resíduos com média zero, se considerarmos os *resíduos padronizado*  $\hat{X}_t/S$ , e será possível checar a independência e a normalidade.

### 3.5 Exercícios

1) Verifique as equações (3.9) para as estimativas de mínimos quadrados quando

$$Z_t = \beta_0 + \beta_1 t + X_t.$$

2) Suponha  $Z_t = \mu + a_t - a_{t-1}$ . Encontre a  $Var(\bar{Z})$  onde

$$\bar{Z} = \frac{1}{n} \sum_{t=1}^n Z_t$$

3) Suponha  $Z_t = \mu + \phi Z_{t-1} + a_t$ . Verifique que

$$Var(\bar{Z}) \approx \left( \frac{1+\phi}{1-\phi} \right) \left( \frac{\gamma_0}{n} \right)$$

## Capítulo 4

# Modelos de Suavização Exponencial

A análise de regressão teve durante muito tempo razoável aceitação como método de ajustar modelos auto-regressivos, com o objetivo de calcular previsões de série temporais. Entretanto, quando o número de observações é muito pequeno, tal análise não é apropriada, pois a hipótese básica de independência dos resíduos é quase sempre violada, produzindo estimadores inconsistentes, nos impossibilitando de testar hipóteses e estabelecer intervalos de confiança para os parâmetros (Morettin e Toloi, 1985).

### 4.1 Modelos para séries localmente constantes

Considere o caso da série temporal  $Z_1, \dots, Z_N$ , localmente composta de seu nível mais um ruído aleatório,

$$Z_t = \mu_t + a_t, \quad t = 1, \dots, N.$$

onde  $E(a_t) = 0$ ,  $Var(a_t) = \sigma_a^2$  e  $\mu_t$  é um parâmetro desconhecido, que pode variar lentamente com o tempo.

#### 4.1.1 Médias Móveis Simples (MMS)

**Procedimento** - A técnica de média móvel consiste em calcular a média aritmética das  $r$  observações mais recentes, isto é,

$$M_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-r+1}}{r}$$

ou

$$M_t = M_{t-1} + \frac{Z_t - Z_{t-r}}{r}$$

Sendo assim,  $M_t$  é uma estimativa do nível  $\mu_t$  que não leva em conta as observações mais antigas, o que é razoável devido ao fato do parâmetro variar suavemente com o tempo.

**Determinação de  $r$**  - Um valor grande de  $r$  faz com que a previsão acompanhe lentamente as mudanças do parâmetro  $\mu$ ; um valor pequeno implica numa reação mais rápida.

**Casos Extremos** - *i*) Se  $r = 1$ , então o valor mais recente da série é utilizado como previsão de todos valores futuros; *ii*) Se  $r = N$ , então a previsão será igual a



média aritmética de todos os dados observados. Este caso só é indicado quando a série é altamente aleatória.

Assim, o valor de  $r$  deve ser proporcional à aleatoriedade de  $a_t$ . Deveria-se escolher o valor de  $r$  de tal maneira que a previsão tivesse o menor EQMMP (erro quadrático médio mínimo de previsão).

**Vantagens** - *i)* Simples aplicação; *ii)* Para número pequeno de observações; *iii)* Permite uma flexibilidade grande devido à variação de  $r$  de acordo com o padrão da série.

**Desvantagens** - *i)* Somente para prever séries estacionárias, pois os pesos atribuídos às  $r$  observações são todos iguais; *ii)* Necessidade de armazenar pelo menos  $(r - 1)$  observações; e *iii)* dificuldade em determinar o valor de  $r$ .

Na prática, o método MMS não é utilizado frequentemente, pois o método AES possui as vantagens anteriores e mais algumas.

### 4.1.2 Suavização Exponencial Simples (SES)

#### Procedimento

A SES pode ser descrita matematicamente por:

$$\bar{Z}_t = \alpha Z_t + (1 - \alpha) \bar{Z}_{t-1}, \quad \bar{Z}_0 = Z_1, \quad t = 1, \dots, N. \quad (4.1)$$

$$\bar{Z}_t = \alpha \sum_{k=0}^{t-1} (1 - \alpha)^k Z_{t-k} + (1 - \alpha)^t \bar{Z}_0, \quad t = 1, \dots, N. \quad (4.2)$$

onde  $\bar{Z}_t$  é denominado valor exponencialmente suavizado e  $\alpha$  é a cte de suavização,  $0 \leq \alpha \leq 1$ . A equação (3.) pode ser obtida de

$$M_t = M_{t-1} + \frac{Z_t - Z_{t-r}}{r}$$

substituindo  $Z_{t-r}$  por  $\bar{Z}_{t-1}$  e  $\frac{1}{r}$  por  $\alpha$ . Fazendo a expansão de (3.) temos que

$$\bar{Z}_t = \alpha Z_t + \alpha (1 - \alpha) Z_{t-1} + \alpha (1 - \alpha)^2 Z_{t-2} + \dots$$

significa que o SES é uma média ponderada que dá pesos maiores às observações mais recentes, eliminando uma das desvantagens do método de MMS.

#### Determinação da cte $\alpha$

Quanto menor for o valor  $\alpha$  mais estáveis serão as previsões finais, uma vez que a utilização de baixo valor de  $\alpha$  implica que pesos maiores são dados às observações passadas e, conseqüentemente, qualquer flutuação aleatória, no presente, exercerá um peso menor no cálculo da previsão. Quanto mais aleatória a série estudada, menores serão os valores da cte de suavização. O efeito de  $\alpha$  grande ou pequeno é completamente análogo (em direção oposta) ao efeito do parâmetro  $r$  no método MMS.

**Vantagens** - *i)* fácil entendimento; *ii)* aplicação não dispendiosa; *iii)* grande flexibilidade permitida pela variação da cte de suavização  $\alpha$ ; *iv)* necessidade de armazenar  $Z_t, \bar{Z}_t$  e  $\alpha$ ; *v)* o valor de  $\alpha = \frac{2}{r-1}$  fornece previsões semelhantes ao método MMS com parâmetro  $r$ .

A principal desvantagem é a dificuldade em determinar o valor mais apropriado da cte de suavização, que pode ser superada através da utilização da suavização exponencial adaptativo de Trigg e Leach (Moretti e Toloi, 1985).

## 4.2 Modelos para séries que apresentam Tendência

As técnicas apresentadas anteriormente não são adequadas para analisar séries temporais que apresentem tendência. Considere então a série temporal não sazonal

$$Z_t = \mu_t + T_t + a_t, \quad t = 1, \dots, N$$

onde  $\mu_t$  é o nível,  $T_t$  é a tendência e  $a_t$  é o ruído branco com média zero e variância constante.

### 4.2.1 Suavização Exponencial Biparamétrica de Holt (SEH)

#### Procedimento

O método SES quando aplicado a uma série que apresenta tendência linear positiva (ou negativa), fornece previsões que subestimam continuamente os valores reais. Para evitar esse erro sistemático, um dos métodos aplicáveis é o SEH. Esse método é similar, em princípio, ao SES, a diferença é que ao invés de suavizar só o nível, ele utiliza uma nova constante de alisamento para "modelar" a tendência da série.

Os valores do nível e da tendência da série, no instante  $t$ , serão estimados por

$$\bar{Z}_t = AZ_t + (1 - A) (\bar{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \text{ e } t = 2, \dots, N \quad (4.3)$$

$$\hat{T}_t = C (\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C) \hat{T}_{t-1}, \quad 0 < C < 1 \text{ e } t = 2, \dots, N \quad (4.4)$$

as constantes  $A$  e  $C$  são denominadas ctes de suavização.

#### Determinação das Constantes

O procedimento é análogo ao de determinação da constante de suavização de um SES, só que ao invés de escolher o valor de  $\alpha$  que torna a soma de EQMP mínimo, escolhe-se o valor do vetor  $(A, C)$  tal que isto ocorra (Morettin e Tolo, 1985).

**Vantagens e Desvantagens** - As vantagens são semelhantes às do método anterior. A desvantagem principal é dificuldade em determinar os valores mais apropriados das ctes de suavização,  $A$  e  $C$ .

## 4.3 Modelos para séries sazonais

Para séries sazonais que apresentam um padrão de comportamento mais complexo, existem outras formas de suavização tais como os métodos de Holt-Winters e o método de suavização exponencial geral.

### 4.3.1 Suavização Exponencial Sazonal de Holt-Winters (HW)

#### Procedimento

Existem dois tipos de procedimento cuja utilização depende das características da série considerada. Tais procedimentos são baseados em três equações com constantes de suavização diferentes, que são associadas a cada uma das componentes do padrão da série: nível, tendência e sazonalidade.

Considere então a série sazonal com tendência aditiva

$$Z_t = \mu_t + T_t + S_t + a_t, \quad t = 1, \dots, N$$

onde  $\mu_t$  é o nível,  $T_t$  é a tendência,  $S_t$  a sazonalidade e  $a_t$  é o ruído branco com média zero e variância constante. As estimativas do fator sazonal, nível e tendência

da série são dados por:

$$\begin{aligned}\hat{S}_t &= D(Z_t - \bar{Z}_t) + (1 - D)\hat{S}_{t-s}, \quad 0 < D < 1 \\ \bar{Z}_t &= A(Z_t - \hat{S}_{t-s}) + (1 - A)(\bar{Z}_{t-1} + \hat{S}_{t-1}), \quad 0 < A < 1 \\ \hat{T}_t &= C(\bar{Z}_t - \bar{Z}_{t-1}) + (1 - C)\hat{T}_{t-1}, \quad 0 < C < 1.\end{aligned}$$

respectivamente A, C e D são as constantes de suavização. Existem também a série sazonal multiplicativa.

**Vantagens e Desvantagens** - As vantagens são semelhantes às da utilização do método de Holt, sendo que os métodos de HW são adequados à análise de série com padrão de comportamento mais geral. As desvantagens são as dificuldades em determinar os valores mais apropriados das constantes de suavização e a impossibilidade e/ou dificuldade de estudar as propriedades estatísticas, tais como média e variância de previsão e, conseqüentemente, construção de um intervalo de confiança. Um procedimento alternativo que não possui tais desvantagens é a Suavização Exponencial Geral de Brown.

As constantes de suavização (A,C,D) devem ser determinadas de tal forma que a soma de quadrado dos erros de suavização seja mínima.

Para maiores detalhes, principalmente sobre previsão, consultar Morettin e Toloi (2004) e outros autores lá citados.

## Capítulo 5

# Modelos de Box-Jenkins para Séries Estacionárias

Apresentaremos neste capítulo os principais modelos de Box-Jenkins para estimação e previsão de séries temporais. Sendo estes modelos pertencentes a família dos autoregressivos-médias-móveis (ARMA), subdividindo em dois outros modelos: o autoregressivo (AR) e médias-móveis (MA).

### 5.1 Processo Linear Geral

Seja  $Z_t$  uma série temporal observada,  $a_t$  o ruído branco, ou seja, uma sequência de variáveis aleatórias independentes identicamente distribuídas, com  $E(a_t) = 0$ , e variância  $\sigma_a^2$  para todo  $t$ . Podemos representar um modelo linear geral como uma combinação linear ponderada, do termo atual, mais os termos passados do ruído branco:

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (5.1)$$

onde

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

e o coeficiente de  $a_t$  é igual a 1, sem perda de generalidade, ou seja,  $\psi_0 = 1$ .

Um caso importante é quando os  $\psi_i$ s são uma sequência que decai exponencialmente, ou seja:

$$\psi_j = \phi^j, \quad -1 < \phi < 1.$$

Então

$$Z_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots$$

Para este exemplo vamos ter:

$$\begin{aligned} E(Z_t) &= E(a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots) \\ &= E(a_t) + \phi E(a_{t-1}) + \phi^2 E(a_{t-2}) + \dots \\ &= 0 \text{ (média constante).} \end{aligned}$$

e variância

$$\begin{aligned}
 Var(Z_t) &= Var(a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots) \\
 &= Var(a_t) + \phi^2 Var(a_{t-1}) + \phi^4 Var(a_{t-2}) + \dots \\
 &\quad (\text{desde que os } a_t\text{'s sejam independentes}) \\
 &= \sigma_a^2 (1 + \phi^2 + \phi^4 + \dots) \\
 &= \sigma_a^2 \frac{1}{1 - \phi^2} \quad (\text{soma de uma série geométrica})
 \end{aligned}$$

Também

$$\begin{aligned}
 Cov(Z_t, Z_{t-1}) &= Cov(a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots, a_{t-1} + \phi a_{t-2} + \phi^2 a_{t-3} + \dots) \\
 &= Cov(\phi a_{t-1}, a_{t-1}) + Cov(\phi^2 a_{t-2}, \phi a_{t-2}) + \dots \\
 &= \sigma_a^2 (\phi + \phi^3 + \phi^5 + \dots) \\
 &= \sigma_a^2 \frac{\phi}{1 - \phi^2}
 \end{aligned}$$

Além disso

$$Corr(Z_t, Z_{t-1}) = \frac{\sigma_a^2 \frac{\phi}{1 - \phi^2}}{\sigma_a^2 \frac{1}{1 - \phi^2}} = \phi$$

De maneira semelhante podemos calcular:

$$Cov(Z_t, Z_{t-k}) = \sigma_a^2 \frac{\phi^k}{1 - \phi^2}$$

e

$$Corr(Z_t, Z_{t-k}) = \phi^k, \quad k = 0, 1, 2, \dots$$

É importante observar que o processo definido desta maneira é estacionário. A estrutura de autocovariância depende somente do "lag" na posição  $k$ .

Para o processo linear geral definido em (4.1) vamos ter:

$$\begin{aligned}
 E(Z_t) &= 0 \\
 \gamma_k &= Cov(Z_t, Z_{t-k}) = \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k}, \quad k \geq 0
 \end{aligned}$$

com  $\psi_0 = 1$ .

## 5.2 Modelo Médias-Móveis (MA(q))

Considere a série  $Z_t$ , fazendo os  $\psi'_i = -\theta'_i$  no processo linear geral e tomando a série como finita, chamamos de *médias-móveis* de ordem  $q$  o modelo:

$$\begin{aligned}
 Z_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \\
 &\quad \text{ou abreviadamente } MA(q)
 \end{aligned} \tag{5.2}$$

Esta terminologia vem do fato que  $Z_t$  é obtido aplicando os pesos  $1, -\theta_1, -\theta_2, \dots, -\theta_q$ , as variáveis  $a_t, a_{t-1}, a_{t-2}, \dots, a_{t-q}$  e então movendo os mesmos pesos 1 unidade do tempo a frente e aplicando-lhes a  $a_{t+1}, a_{t+2}, \dots, a_{t+q+1}$  para obter  $Z_{t+1}$ .

### 5.2.1 O modelo MA(1)

Considere o seguinte modelo:

$$Z_t = a_t - \theta_1 a_{t-1}$$

onde

$$E(Z_t) = 0$$

e a variância é igual a:

$$\begin{aligned} \gamma_0 &= \text{Var}(Z_t) = \text{Var}(a_t - \theta_1 a_{t-1}) \\ &= \sigma_a^2 + \theta_1^2 \sigma_a^2 = \sigma_a^2 (1 + \theta_1^2) \end{aligned}$$

temos ainda que a facv é:

$$\begin{aligned} \gamma_1 &= \text{Cov}(Z_t, Z_{t-1}) \\ &= \text{Cov}(a_t - \theta_1 a_{t-1}, a_{t-1} - \theta_1 a_{t-2}) \\ &= -\theta_1 \text{Cov}(a_{t-1}, a_{t-1}) = -\theta_1 \sigma_a^2 \end{aligned}$$

e para  $k \geq 2$  teremos

$$\gamma_k = \text{Cov}(Z_t, Z_{t-k}) = 0$$

E a fac será dada por:

$$\rho_k = \begin{cases} 1, & k = 0 \\ \frac{-\theta_1 \sigma_a^2}{\sigma_a^2 (1 + \theta_1^2)} = \frac{-\theta_1}{(1 + \theta_1^2)}, & k = 1 \\ 0, & k \geq 2 \end{cases}$$

### 5.2.2 O Modelo MA(q)

Considere o modelo dado em (4.2)

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

onde

$$E(Z_t) = 0$$

e a variância é

$$\begin{aligned} \gamma_0 &= \text{Var}(Z_t) = \text{Var}(a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}) \\ &= (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2 \end{aligned}$$

a facv é dada por

$$\begin{aligned} \gamma_1 &= \text{Cov}(Z_t, Z_{t-1}) \\ &= \text{Cov}(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, a_{t-1} - \theta_1 a_{t-2} - \dots - \theta_q a_{t-q-1}) \\ &= -\theta_1 \sigma_a^2 + \theta_1 \theta_2 \sigma_a^2 + \dots + \theta_{q-1} \theta_q \sigma_a^2 \\ &= (-\theta_1 + \theta_1 \theta_2 + \dots + \theta_{q-1} \theta_q) \sigma_a^2, \text{ para } k = 1 \end{aligned}$$

e

$$\gamma_2 = (-\theta_2 + \theta_1\theta_3 + \dots + \theta_{q-2}\theta_q) \sigma_a^2, \text{ para } k = 2$$

e para  $k \geq q + 1$  vamos ter  $\gamma_k = 0$ . Enquanto que a fac será dada por

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

### 5.3 Modelo Autoregressivo AR(p)

Chamamos de autoregressivo de ordem  $p$  o modelo:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (5.3)$$

ou simplesmente  $AR(p)$

onde os  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$  são independentes de  $a_t$ . Os valores da série  $Z_t$  são uma combinação linear dos  $p$  valores passados mais um termo  $a_t$ , no qual incorpora coisas na série até o tempo  $t$  que não é explicado pelos valores passados.

#### 5.3.1 O Modelo AR(1)

Considere o seguinte modelo:

$$Z_t = \phi Z_{t-1} + a_t$$

onde

$$E(Z_t) = 0, \text{ para todo } t$$

e a variância

$$\begin{aligned} \gamma_0 &= \text{Var}(Z_t) = \text{Var}(\phi Z_{t-1} + a_t) \\ &= \phi^2 \text{Var}(Z_{t-1}) + \sigma_a^2 = \phi^2 \gamma_0 + \sigma_a^2 \\ &= \frac{\sigma_a^2}{1 - \phi^2}, \quad |\phi| < 1. \end{aligned}$$

A facv para  $k = 1$  é

$$\begin{aligned} \gamma_1 &= \text{Cov}(Z_t, Z_{t-1}) = E(Z_t Z_{t-1}) \\ &= E[(\phi Z_{t-1} + a_t)(\phi Z_{t-2} + a_{t-1})] \\ &= E[\phi Z_{t-1} \phi Z_{t-2} + \phi Z_{t-1} a_{t-1} + a_t \phi Z_{t-2} + a_t a_{t-1}] \\ &= \phi^2 E[Z_{t-1} Z_{t-2}] + \phi E[Z_{t-1} a_{t-1}] + \phi E[a_t Z_{t-2}] + E[a_t a_{t-1}] \end{aligned}$$

onde

$$\begin{aligned} E(a_t Z_{t-j}) &= 0, \quad j > 0 \\ &\neq 0, \quad j < 0 \\ &= \sigma_a^2, \quad j = 0 \end{aligned}$$

sendo assim temos

$$\begin{aligned}\gamma_1 &= \phi^2 \gamma_1 + \phi \sigma_a^2 \\ &= \frac{\phi}{1 - \phi^2} \sigma_a^2\end{aligned}$$

para  $k = 2$

$$\gamma_2 = \phi \gamma_1.$$

No geral a facv será

$$\gamma_k = \frac{\phi^k}{1 - \phi^2} \sigma_a^2, \text{ para } k = 0, 1, 2, \dots$$

e a fac para o caso geral será

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \phi^k, \text{ para } k = 0, 1, 2, \dots$$

Desde que  $|\phi| < 1$ , a fac é uma exponencial decrescente, a medida que  $k$  cresce. Se  $0 < \phi < 1$ , todas as correlações são positivas; se  $-1 < \phi < 0$  a autocorrelação de “lag ” 1 é negativa ( $\rho_1 = \phi$ ) e os sinais das autocorrelações são alternadamente positivo e negativo.

Considere o modelo AR(1):

$$Z_t = \phi Z_{t-1} + a_t$$

substituindo  $t$  por  $t - 1$ , temos:

$$Z_{t-1} = \phi Z_{t-2} + a_{t-1}$$

agora substituindo  $Z_{t-1}$  em  $Z_t$ , temos:

$$\begin{aligned}Z_t &= \phi(\phi Z_{t-2} + a_{t-1}) + a_t \\ &= \phi^2 Z_{t-2} + \phi a_{t-1} + a_t\end{aligned}$$

repetindo este processo  $k - 1$  vezes, vamos ter:

$$Z_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots + \phi^{k-1} a_{t-(k-1)} + \phi^k Z_{t-k}$$

para  $k$  grande vamos ter:

$$Z_t = a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots = \psi_0 a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (5.4)$$

onde  $|\phi| < 1$  e  $\phi^j = \psi_j$ . Temos ainda que os  $a_t$  são independente de  $Z_{t-1}, Z_{t-2}, \dots$  e  $\sigma_a^2 > 0$ , a solução de  $Z_t = \phi Z_{t-1} + a_t$  será estacionária se e somente se  $|\phi| < 1$ .

### 5.3.2 O Modelo AR(2)

Considere o seguinte modelo:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t \quad (5.5)$$

onde  $a_t$  é independente de  $Z_{t-1}$  e  $Z_{t-2}$ . Para analisar a estacionariedade deste processo AR(2), vamos introduzir o polinômio característico:



$$\phi(x) = 1 - \phi_1 x - \phi_2 x^2$$

e a equação característica

$$1 - \phi_1 x - \phi_2 x^2 = 0$$

A equação quadrática, acima tem duas raízes (possivelmente complexas). Suponha  $a_t$  independente de  $Z_{t-1}, Z_{t-2}, \dots$ , uma solução estacionária para a equação (4.5) existirá se e somente se a raiz da equação característica AR excede a unidade em valor absoluto. Este procedimento pode ser generalizado para o modelo AR(p). Para o caso AR(1) a equação característica é  $1 - \phi x = 0$  com raiz  $1/\phi$ , que exceda 1 em valor absoluto se e somente se  $|\phi| < 1$ . Para a equação característica do AR(2) temos a seguinte solução:

$$x = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}$$

as raízes da equação acima excede a 1 em módulo se e somente se, simultaneamente

$$\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1 \quad \text{e} \quad |\phi_2| < 1.$$

Essas são as condições de estacionariedade para o AR(2).

Para encontrar a fac para o AR(2), multiplicamos ambos os lados da equação (4.5) por  $Z_{t-k}$ ,  $k = 1, 2, \dots$ , e tomamos as esperanças

$$\begin{aligned} \gamma_k &= E(Z_t Z_{t-k}) = E[(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t) Z_{t-k}] \\ &= E(\phi_1 Z_{t-1} Z_{t-k}) + E(\phi_2 Z_{t-2} Z_{t-k}) + E(a_t Z_{t-k}). \end{aligned}$$

Assumindo estacionariedade, médias zero, e que  $a_t$  é independente de  $Z_{t-k}$  nós temos

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2}, \quad k = 1, 2, \dots \quad (5.6)$$

dividindo por  $\gamma_0$

$$\begin{aligned} \frac{\gamma_k}{\gamma_0} &= \phi_1 \frac{\gamma_{k-1}}{\gamma_0} + \phi_2 \frac{\gamma_{k-2}}{\gamma_0}, \quad k = 1, 2, \dots \\ \rho_k &= \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k = 1, 2, \dots \end{aligned} \quad (5.7)$$

As equações (4.6) e/ou (4.7) são chamadas de **equações de Yule-Walker**. Para  $k = 1$ ,  $\rho_0 = 1$ ,  $\rho_{-1} = \rho_1$  temos

$$\begin{aligned} \rho_1 &= \phi_1 \rho_0 + \phi_2 \rho_1 \\ \rho_1 &= \frac{\phi_1}{1 - \phi_2} \end{aligned}$$

e para  $k = 2$

$$\begin{aligned} \rho_2 &= \phi_1 \rho_1 + \phi_2 \rho_0 \\ &= \frac{\phi_1^2 + \phi_2(1 - \phi_2)}{1 - \phi_2}. \end{aligned}$$

Podemos calcular sucessivamente os valores de  $\rho_k$  através das equações de Yule-Walker.

A variância do processo AR(2) é obtida através da equação (4.5)

$$\begin{aligned}\gamma_0 &= \text{Var}(Z_t) = \text{Var}(\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + a_t) \\ &= \phi_1^2 \gamma_0 + \phi_2^2 \gamma_0 + 2\phi_1 \phi_2 \gamma_1 + \sigma_a^2 \\ &= (\phi_1^2 + \phi_2^2) \gamma_0 + 2\phi_1 \phi_2 \gamma_1 + \sigma_a^2\end{aligned}$$

utilizando as equações de Yule-Walker vamos ter

$$\gamma_0 = \frac{(1 - \phi_2) \sigma_a^2}{(1 - \phi_2)(1 - \phi_1^2 - \phi_2^2) - 2\phi_1^2 \phi_2}$$

Os coeficientes  $\psi_j$  no processo linear geral (4.1) para um AR(2) são mais complexos do que o AR(1). Entretanto, podemos substituir a representação (4.1) e definir a equação do AR(2) para  $Z_t, Z_{t-1}$ , e  $Z_{t-2}$  e obter os coeficientes de  $a_j$ . Encontrando

$$\begin{aligned}\psi_0 &= 1 \\ \psi_1 - \phi_1 \psi_0 &= 0\end{aligned}$$

e

$$\psi_j - \phi_1 \psi_{j-1} - \phi_2 \psi_{j-2} = 0, \quad j = 2, 3, \dots$$

Estas equações podem ser resolvidas recursivamente obtendo  $\psi_0 = 1$ ,  $\psi_1 = \phi_1$ ,  $\psi_2 = \phi_1^2 + \phi_2$ , e assim por diante.

### 5.3.3 O Processo Autoregressivo geral

Considere agora o modelo autoregressivo de ordem  $p$

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (5.8)$$

com polinômio característico

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

e a correspondente equação característica AR

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0.$$

Considerando que  $a_t$  é independente de  $Z_{t-1}, Z_{t-2}, \dots$ , uma solução estacionária para a equação (4.8) existe se e somente se as  $p$  raízes da equação característica AR é maior do que a unidade em módulo.

Assumindo que a equação (4.8) é estacionária e multiplicando-a por  $Z_{t-k}$ , dividindo pela variância e tomando as esperanças temos:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k \geq 1. \quad (5.9)$$

Fazendo  $k = 1, 2, \dots, p$  na equação (4.9) e utilizando  $\rho_0 = 1$  e  $\rho_{-k} = \rho_k$ , obtemos as equações de **Yule-Walker**

$$\begin{aligned}
 \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\
 \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\
 &\vdots \\
 \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p
 \end{aligned} \tag{5.10}$$

Dados os valores  $\phi_1, \phi_2, \dots, \phi_p$ , estas equações podem ser resolvidas para obtermos:  $\rho_1, \rho_2, \dots, \rho_p$ . Para obter a variância multiplicamos a equação (4.8) por  $Z_t$ , e tomamos as esperanças encontrando:

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_a^2$$

utilizando  $\rho_k = \gamma_k / \gamma_0$  podemos escrever

$$\text{Var}(Z_t) = \gamma_0 = \frac{\sigma_a^2}{1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p} \tag{5.11}$$

observando que

$$\begin{aligned}
 E(a_t Z_t) &= E[a_t (\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t)] \\
 &= E(a_t^2) = \sigma_a^2
 \end{aligned}$$

e a variância do processo é expressa em termos dos parâmetros  $\phi_1, \phi_2, \dots, \phi_p, \sigma_a^2$ , e os desconhecidos valores de  $\rho_1, \rho_2, \dots, \rho_p$ .

Assumindo estacionariedade, o processo pode ser expressado na forma linear geral da equação (4.1), mas os pesos  $\psi$  são funções complicadas dos parâmetros  $\phi_1, \phi_2, \dots, \phi_p$ , mas podem ser encontrados numericamente.

## 5.4 O Modelo Autoregressivo-Médias Móveis AR-MA(p,q)

Se considerarmos uma série formada pelas partes autoregressiva e médias-móveis, vamos ter um modelo mais geral de séries temporais, ou seja, se

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \tag{5.12}$$

nós dizemos que  $Z_t$  é um processo *autoregressivo médias-móveis* de ordens  $p$  e  $q$ , respectivamente, e parâmetros  $\phi's$  e  $\theta's$ , ou abreviadamente  $ARMA(p,q)$ .

### 5.4.1 O modelo ARMA(1,1)

Considere o seguinte modelo

$$Z_t = \phi Z_{t-1} + a_t - \theta a_{t-1} \tag{5.13}$$

onde  $E(Z_t) = 0$ , para todo  $t$ ,  $E(a_t) = 0$  e  $\text{Var}(a_t) = \sigma_a^2$ . Vamos obter as equações de Yule-Walker, antes porém, observamos que:

$$E(a_t Z_t) = E[a_t (\phi Z_{t-1} + a_t - \theta a_{t-1})] = \sigma_a^2$$

e

$$\begin{aligned} E(a_{t-1}Z_t) &= E[a_{t-1}(\phi Z_{t-1} + a_t - \theta a_{t-1})] \\ &= \phi \sigma_a^2 - \theta \sigma_a^2 = (\phi - \theta) \sigma_a^2 \end{aligned}$$

Se multiplicamos a equação (4.13) por  $Z_{t-k}$  e tomamos as esperanças, temos:

$$\begin{aligned} \gamma_k &= E(Z_t Z_{t-k}) = E[(\phi Z_{t-1} + a_t - \theta a_{t-1}) Z_{t-k}] \\ &= E(\phi Z_{t-1} Z_{t-k}) + E(a_t Z_{t-k}) - E(\theta a_{t-1} Z_{t-k}) \end{aligned}$$

variando o  $k$

$$\begin{aligned} \gamma_0 &= \phi \gamma_1 + \sigma_a^2 - \theta (\phi - \theta) \sigma_a^2, \quad k = 0 \\ \gamma_1 &= \phi \gamma_0 - \theta \sigma_a^2, \quad k = 1 \end{aligned} \tag{5.14}$$

e

$$\gamma_k = \phi \gamma_{k-1}, \quad k \geq 2 \tag{5.15}$$

resolvendo as equações (4.14) e (4.15) vamos ter a variância:

$$\gamma_0 = \frac{(1 - 2\theta\phi + \theta^2)}{1 - \phi^2}$$

a facv

$$\gamma_k = \frac{(1 - \theta\phi)(\phi - \theta)}{1 - \phi^2} \phi^{k-1} \sigma_a^2, \quad \text{para } k \geq 1$$

e a fac

$$\rho_k = \frac{(1 - \theta\phi)(\phi - \theta)}{1 - 2\theta\phi + \theta^2} \phi^{k-1}, \quad \text{para } k \geq 1$$

A forma linear geral pode ser obtida da mesma maneira que a equação (4.5), definimos

$$Z_t = a_t + (\phi - \theta) \sum_{j=1}^{\infty} \phi^{j-1} a_{t-j}$$

onde  $\psi_j = (\phi - \theta) \phi^{j-1}$ , para  $j \geq 1$ . Com a condição de estacionariedade  $|\phi| < 1$ , ou a raiz da equação  $1 - \phi x = 0$  é maior do que 1 em valor absoluto.

O modelo ARMA(1,1) é equivalente a:

$$AR(\infty) : \frac{(1 - \phi x)}{(1 - \theta x)} Z_t = (1 - \theta x + \theta^2 x - \dots) (1 - \phi x) Z_t = a_t$$

e

$$MA(\infty) : Z_t = \frac{(1 - \theta x)}{(1 - \phi x)} a_t = (1 + \phi x + \phi^2 x^2 + \dots) (1 - \theta x) a_t$$

A função de autocorrelação do modelo ARMA(1,1) é semelhante a do AR(1) (exponenciais amortecidas), o primeiro comporta-se como exponenciais amortecidas para  $k > 0$ , enquanto que o segundo o comportamento é de exponenciais amortecidas para  $k \geq 0$ .

Para o modelo geral  $ARMA(p, q)$ , nós temos que dado que  $a_t$  é independente de  $Z_{t-1}, Z_{t-2}, \dots$ , uma solução estacionária para  $Z_t$ , satisfazendo a equação (4.13) existe, se e somente se, as raízes da equação característica AR,  $\phi(x) = 0$ , excede a unidade em valor absoluto.

Se as condições de estacionariedade são satisfeitas, então o modelo pode ser escrito como um processo linear geral com os pesos  $\psi_j$  determinados da seguinte forma:

$$\begin{aligned}\psi_0 &= 1 \\ \psi_1 &= -\theta_1 + \phi_1 \\ \psi_2 &= -\theta_2 + \phi_2 + \phi_1 \psi_1 \\ &\vdots \\ \psi_j &= -\theta_j \phi_p \psi_{j-p} + \dots + \phi_1 \psi_{j-1}\end{aligned}$$

onde  $\psi_j = 0$  para  $j > 0$  e  $\theta_j = 0$  para  $j > q$ .

Agora assumindo estacionariedade pode-se mostrar que a fac satisfaz:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > q \quad (5.16)$$

Podemos também desenvolver para  $k = 0, 1, \dots, q$  que envolve  $\theta_1, \theta_2, \dots, \theta_q$ .

**Teorema:** Se  $X_t \sim ARMA(p_1, q_1)$  e  $Y_t \sim ARMA(p_2, q_2)$ , sendo  $X_t$  e  $Y_t$  independentes, seja  $Z_t = X_t + Y_t$  então  $Z_t \sim ARMA(p, q)$  onde  $p \leq p_1 + p_2$ ,  $q \leq \max(p_1 + q_2, p_2 + q_1)$ .

**Prova:** Seja  $\phi_X(B) X_t = \theta_X(B) \varepsilon_t$  e  $\phi_Y(B) Y_t = \theta_Y(B) a_t$  onde  $\phi_X, \phi_Y, \theta_X$  e  $\theta_Y$  são polinômios em  $B$  de ordem  $p_1, p_2, q_1$  e  $q_2$ , respectivamente. Onde  $\varepsilon_t$  e  $a_t$  são ruídos brancos independentes. Como  $Z_t = X_t + Y_t$ , então multiplicando  $Z_t$  por  $\phi_X(B) \phi_Y(B)$  temos:

$$\phi_X(B) \phi_Y(B) Z_t = \phi_X(B) \phi_Y(B) X_t + \phi_X(B) \phi_Y(B) Y_t$$

como  $\phi_X(B) X_t = \theta_X(B) \varepsilon_t$  é  $ARMA(p_1, q_1)$  e  $\phi_Y(B) Y_t = \theta_Y(B) a_t$  é  $ARMA(p_2, q_2)$  vamos ter:

$$\begin{aligned}\phi_X(B) \phi_Y(B) Z_t &= \theta_X(B) \phi_Y(B) \varepsilon_t + \phi_X(B) \theta_Y(B) a_t \\ AR(p_1 + p_2) &= MA(q_1 + p_2) + MA(p_1 + q_2)\end{aligned}$$

Utilizando o fato que a soma de dois processos de médias móveis independentes também é um processo MA de ordem igual ou menor ao max das ordens, temos que  $Z_t$  é um processo  $ARMA(p, q)$  onde  $p \leq p_1 + p_2$ ,  $q \leq \max(p_1 + q_2, p_2 + q_1)$ .

## 5.5 Invertibilidade

Nós vimos que um processo AR pode ser reescrito como um processo MA de ordem infinita através de pesos  $\psi'$ s. Além disso podemos escrever um processo MA como um autoregressivo. Considere o modelo abaixo:

$$Z_t = a_t - \theta a_{t-1} \quad (5.17)$$

reescrevendo a equação acima como  $a_t = Z_t + \theta a_{t-1}$ , e então substituindo  $t$  por  $t-1$  e  $a_{t-1}$  na equação modificada, temos:

$$\begin{aligned} a_t &= Z_t + \theta(Z_{t-1} + \theta a_{t-2}) \\ &= Z_t + \theta Z_{t-1} + \theta^2 a_{t-2} \end{aligned}$$

Se  $|\theta| < 1$ , podemos continuar a substituição e obter:

$$\begin{aligned} a_t &= Z_t + \theta Z_{t-1} + \theta^2 Z_{t-2} + \dots \\ Z_t &= (-\theta Z_{t-1} - \theta^2 Z_{t-2} - \dots) + a_t \end{aligned}$$

então se  $|\theta| < 1$ , nós vimos que o MA(1) pode ser “*invertido*” (transformado) para um AR( $\infty$ ), então dizemos que o modelo MA(1) é *invertível*.

Para um modelo geral MA( $q$ ), definimos o **polinômio característico MA** como:

$$\theta(x) = 1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q$$

e a correspondente **equação característica**

$$1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q = 0$$

Pode-se então, demonstrar que o modelo MA( $q$ ) é invertível, e existirão constantes  $\pi_j$ , tal que:

$$Z_t = \sum_{j=1}^{\infty} \pi_j Z_{t-j} + a_t$$

se e somente se as raízes da equação característica MA excede a unidade em valor absoluto.

**Proposição:** Um processo linear geral será estacionário se a série  $\psi(x)$  converge para  $|x| \leq 1$ ; será invertível se  $\pi(x)$  converge para  $|x| \leq 1$ .

Podemos tirar algumas conclusões sobre *estacionariedade e invertibilidade*, são elas:

- 1) Para o modelo AR( $p$ ) não há restrições sobre os parâmetros  $\phi_j$  para assegurar a *invertibilidade*.
- 2) Para o modelo MA( $q$ ) não há restrições sobre os parâmetros  $\theta_j$  para que o processo seja *estacionário*.
- 3) Para o modelo ARMA( $p, q$ ) o processo é *estacionário* se as raízes da equação característica  $\phi(x) = 0$  excede a unidade em valor absoluto e o processo é *invertível* se todas as raízes  $\theta(x) = 0$  excede a unidade em valor absoluto, ou seja, fora do círculo unitário.

## 5.6 Exercícios

1) Esboçe a fac para cada um dos seguintes modelos ARMA:

- a) AR(2) com  $\phi_1 = 1.2$  e  $\phi_2 = -0.7$
- b) MA(2) com  $\theta_1 = -1$  e  $\theta_2 = -0.6$
- c) ARMA(1,1) com  $\phi = 0.7$  e  $\theta = -0.4$
- 2) Suponha  $Z_t$  um processo AR(1) com  $-1 < \phi < 1$ .
  - a) Encontre a facv para  $W_t = Z_t - Z_{t-1}$  em termos de  $\phi$  e  $\sigma_a^2$ .
  - b) Em particular, mostrar que  $Var(W_t) = 2\sigma_a^2 / (1 + \phi)$ .
- 3) Encontre a fac para o processo definido por

$$Z_t = 5 + a_t - 0.5a_{t-1} + 0.25a_{t-2}$$

- 4) Descreva as características importantes da fac dos seguintes modelos: a) MA(1), b) MA(2), c) AR(1), d) AR(2) e e) ARMA(1,1).  
 5) Para o modelo ARMA(2,1)

$$Z_t = 0.8Z_{t-1} + a_t + 0.7a_{t-1} + 0.6a_{t-2}$$

mostrar que

- a)  $\rho_k = 0.8\rho_{k-1}$  para  $k \geq 3$  e  
 b)  $\rho_2 = 0.8\rho_1 + 0.6\sigma_a^2/\gamma_0$ .  
 6) Considere dois processos MA(2), um com  $\theta_1 = \theta_2 = 1/6$  e outro com  $\theta_1 = -1$  e  $\theta_2 = 6$ . Mostrar que estes processos tem exatamente a mesma fac. Como são as raízes dos correspondentes polinômios característicos, compare-as.  
 7) Considere o não-estacionário modelo AR(1)

$$Z_t = 3Z_{t-1} + a_t$$

- a) Mostrar que  $Z_t = -\sum_{j=1}^{\infty} (1/3)^j a_{t+j}$  satisfaz a equação AR(1) e é realmente estacionária.  
 b) Quando está solução não é satisfatória.  
 8) Considere o modelo

$$Z_t = a_{t-1} - a_{t-2} + 0.5a_{t-3}$$

- a) Encontre a facv para  $Z_t$ .  
 b) Mostrar que  $Z_t$  é um modelo estacionário ARMA(p,q). Identifique  $p, q$  e os  $\theta'$ s e  $\phi'$ s.  
 9) Considere os modelos:  
 i)  $Z_t = a_t + 0.8a_{t-1}$ .  
 ii)  $Z_t - 0.4Z_{t-1} = a_t - 0.3a_{t-1} + 0.8a_{t-2}$   
 iii)  $Z_t = 0.3Z_{t-1} + 0.6Z_{t-2} + a_t$ .  
 Pede-se:  
 a) Escreva-os utilizando o operador  $X$ ;  
 b) Identifique cada um dos modelos abaixo, assim como os seus parâmetros;  
 c) Verifique se cada um deles são *estacionários e/ou invertíveis*.

## Capítulo 6

# Modelos para Séries Temporais Não-Estacionárias

Os modelos apresentados até o momento são adequados para séries estacionárias, ou seja, aquelas onde a média é constante por todo tempo, mas em geral, na prática, as séries são não-estacionárias. Como, por exemplo, as séries econômicas. Para torna a série estacionária deve-se tomar diferenças quantas vezes for necessário, até atingir estacionariedade. O procedimento é o seguinte:

$$\begin{aligned} W_t &= Z_t - Z_{t-1} \\ &= (1 - X) Z_t \\ &= \Delta Z_t \end{aligned}$$

Os modelos que apresentaremos a partir de agora, serão para séries cujo comportamento são não-estacionário.

### 6.1 O Modelo Autoregressivo-Integrado-Médias-Móveis ARIMA(p,d,q)

As séries  $Z_t$ , tais que, tomando-se um número finito de diferenças,  $d$ , tornam-se estacionárias, são chamadas *não-estacionárias homogêneas*. Se

$$W_t = \Delta^d Z_t$$

é estacionária, podemos representar  $W_t$  por um modelo  $ARMA(p,q)$ , ou seja,

$$\phi(X) W_t = \theta(X) a_t \quad (6.1)$$

Se  $W_t$  é uma diferença de  $Z_t$ , então  $Z_t$  é uma integral de  $W_t$ , daí dizemos que  $Z_t$  segue um modelo *autoregressivo-integrado-médias-móveis*, ou modelo *ARIMA(p,d,q)*,

$$\phi(X) \Delta^d Z_t = \theta(X) a_t \quad (6.2)$$

de ordem  $(p,d,q)$ , se  $p$  e  $q$  são as ordens de  $\phi(X)$  e  $\theta(X)$ , respectivamente.

No modelo (5.1) todas as raízes de  $\phi(X)$  estão fora do círculo unitário. Escrever (5.2) é equivalente a escrever

$$\xi(X) Z_t = \theta(X) a_t \quad (6.3)$$



onde  $\xi(X)$  é um operador autoregressivo não-estacionário, de ordem  $p+d$ , com  $d$  raízes iguais a 1 (sobre o círculo unitário) e as restantes  $p$  estão fora do círculo unitário, ou seja,

$$\xi(X) = \phi(X) \Delta^d = \phi(X) (1 - X)^d \quad (6.4)$$

Portanto o modelo (5.2) supõe que a  $d$ -ésima diferença da série  $Z_t$  pode ser representada por um modelo *ARMA*, estacionário e invertível. Na maioria dos casos usuais,  $d=1$  ou  $d=2$ , que correspondem a dois casos interessantes e comuns de não-estacionariedade homogênea:

1) Séries não-estacionárias quanto ao nível: oscilam ao redor de um nível médio durante algum tempo e depois saltam para outro nível temporário. Para torná-las estacionárias é suficiente tomar uma diferença, este é o caso típico de séries econômicas.

2) Séries não-estacionárias quanto a inclinação: oscilam numa direção por algum tempo e depois mudam para outra direção temporária. Para torná-las estacionárias é necessário tomar a segunda diferença.

### 6.1.1 Exemplos

Alguns casos particulares do modelo *ARIMA*:

- i) *ARIMA*(0,1,1):  $\Delta Z_t = (1 - \theta X) a_t$
- ii) *ARIMA*(1,1,1):  $(1 - \phi X) \Delta Z_t = (1 - \theta X) a_t$
- iii) *ARIMA*( $p,0,0$ ): *AR*( $p$ ); *ARIMA*(0,0, $q$ ): *MA*( $q$ ); *ARIMA*( $p,0,q$ ): *ARMA*( $p,q$ ).

Um modelo que é considerado importante é o caso i) *IMA*(1,1): *Integrado-Médias-Móveis*. Utilizado especialmente na área de economia.

### 6.1.2 Algumas Transformações para tornar a série Estacionária

Tomar diferenças pode não ser suficiente para se alcançar estacionariedade, principalmente, no caso das séries econômicas. Uma transformação, não linear, utilizada para série  $Z_t$  é  $Z_t^* = \ln Z_t$ , que será suficiente para obter a homogeneidade. Outro procedimento usualmente, utilizado em séries temporais econômicas é:

$$\Delta \ln Z_t = \ln Z_t - \ln Z_{t-1}$$

A principal razão para se fazer uma transformação é tentar estabilizar a variância. Uma transformação, também adequada, seria: utilizar um gráfico que traz no eixo das abscissas médias de subconjuntos de observações da série original e no eixo das ordenadas a amplitude de cada um destes subconjuntos. Seja  $Z_{t_1}, \dots, Z_{t_k}$  um subconjunto com  $k$  observações, então calculamos:

$$\bar{Z} = \frac{1}{k} \sum_{i=1}^k Z_{t_i}$$

(medida de posição)

e

$$W = \max(Z_{t_i}) - \min(Z_{t_i})$$

(medida de variabilidade)

o par  $(\bar{Z}, W)$  será um ponto do gráfico. O número de elementos em cada subsérie pode ser igual ao período, no caso de séries sazonais. Se  $W$  é independente de  $\bar{Z}$ , obteremos pontos espalhados ao redor de uma reta paralela ao eixo das abscissas e neste caso não haverá necessidade de transformação. Se  $W$  for diretamente proporcional a  $\bar{Z}$ , a transformação logarítmica é apropriada.

Uma classe geral de transformações que podem ser utilizadas é a de Box-Cox, definida por:

$$Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^\lambda - c}{\lambda}, & \lambda \neq 0 \\ \ln Z_t, & \lambda = 0 \end{cases}$$

onde  $\lambda$  e  $c$  são parâmetros a serem estimados. Para  $c = 1$  temos:

$$\frac{Z_t^\lambda - 1}{\lambda} \longrightarrow \ln Z_t \text{ se } \lambda \rightarrow 0$$

## 6.2 Formas do Modelo ARIMA

O modelo ARIMA dado em (5.2) pode ser representado de três formas:

- a) Em termos de valores prévios de  $Z_t$  e do valor atual e prévios de  $a_t$ ;
- b) Em termos de valor atual e prévios de  $a_t$ ;
- c) Em termos de valores prévios de  $Z_t$  e do valor atual de  $a_t$ .

### 6.2.1 Forma da Equação Diferenças

Esta é a forma usual do modelo, útil para calcular previsões:

$$Z_t = \xi_1 Z_{t-1} + \xi_2 Z_{t-2} + \dots + \xi_{p+d} Z_{t-p-d} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (6.5)$$

onde  $\xi(X) = 1 - \xi_1 X - \xi_2 X^2 - \dots - \xi_{p+d} X^{p+d}$ .

### 6.2.2 Forma de Choques Aleatórios

Uma forma conveniente para se calcular a variância dos erros de previsão é:

$$\begin{aligned} Z_t &= a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \\ &= \psi(X) a_t \end{aligned} \quad (6.6)$$

Desta equação obtemos

$$\xi(X) Z_t = \xi(X) \psi(X) a_t \quad (6.7)$$

e utilizando (5.3) segue-se que

$$\xi(X) \psi(X) = \theta(X). \quad (6.8)$$

Logo, os pesos  $\psi_j$  da forma (5.6) podem ser obtidos de (5.8) identificando-se coeficientes de  $X, X^2$ , etc.:

$$(1 - \xi_1 X - \dots - \xi_{p+d} X^{p+d}) (1 + \psi_1 X + \psi_2 X^2 + \dots) = 1 - \theta_1 X - \dots - \theta_q X^q$$

### 6.2.3 Termo constante no Modelo ARIMA

No modelo  $ARIMA(p, d, q)$

$$\begin{aligned} \phi(X) W_t &= \theta(X) a_t \\ \text{onde } W_t &= \Delta^d Z_t. \end{aligned} \quad (6.9)$$

Se um termo constante for omitido, então  $E(W_t) = \mu_W = 0$ . O modelo acima pode descrever o que chamaríamos de *tendências estocásticas*, no sentido que o processo não é estacionário e muda de nível e/ou inclinação, no decorrer do tempo. A tendência (ou não-estacionariedade) estocástica é caracterizada pela existência de

zeros de  $\xi(X) = 1 - X$  sobre o círculo unitário. Além da não-estacionariedade estocástica, muitas séries podem apresentar uma *tendência determinística*. Podemos ter  $Z_t$  como a soma de um polinômio mais um processo  $ARIMA(p, d, q)$ :

$$\begin{aligned} Z_t &= \sum_{j=0}^m \beta_j t^j + \frac{\theta(X)}{\Delta^d \phi(X)} a_t \\ &= T_t + Y_t \end{aligned} \quad (6.10)$$

onde  $T_t$  é uma tendência determinística e  $Y_t$  é um processo  $ARIMA(p, d, q)$ . Segue-se que  $Z_t$  é não-estacionário se  $m > 0$  e/ou  $d > 0$ .

Tomando  $d$  diferenças temos:

$$\begin{aligned} \Delta^d Z_t &= \theta_0 + \frac{\theta(X)}{\phi(X)} a_t, \text{ se } m = d \\ \Delta^d Z_t &= \frac{\theta(X)}{\phi(X)} a_t, \text{ se } m < d \end{aligned} \quad (6.11)$$

onde  $\theta_0 = \beta_d d!$ , obtendo-se uma série estacionária. Significando que podemos incluir uma tendência polinomial determinística de grau  $d$  no modelo, bastando acrescentar uma constante  $\theta_0$ :

$$\xi(X) Z_t = \theta_0 + \theta(X) a_t \quad (6.12)$$

Se  $m > d$ , podemos obter um modelo não-estacionário, tomando  $d$  diferenças, devido a tendência determinística, e tomando  $m$  diferenças, obteremos um processo estacionário, mas não invertível.

Se  $\theta_0 \neq 0$ ,

$$W_t = \phi W_{t-1} + \dots + \phi_p W_{t-p} + \theta_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (6.13)$$

e obtemos a média

$$\begin{aligned} \mu_W &= E(W_t) = \phi_1 \mu_W + \dots + \phi_p \mu_W \\ &= \frac{\theta_0}{1 - \phi_1 - \dots - \phi_p} \end{aligned} \quad (6.14)$$

ou

$$\theta_0 = \mu_W (1 - \phi_1 - \dots - \phi_p) \quad (6.15)$$

e se

$$\widetilde{W}_t = W_t - E(W_t)$$

teremos:

$$\phi(X) \widetilde{W}_t = \theta(X) a_t$$

No que segue, quando  $d > 0$ , suporemos  $\mu_W = 0$  e portanto  $\theta_0 = 0$ .

### 6.3 Construção dos Modelos ARIMA

Nesta seção vamos apresentar os estágios do ciclo iterativo do método de Box e Jenkins, para construção dos modelos ARIMA, que são: 1) *identificação*; 2) *estimação* e 3) *verificação*. Dentre estes estágios o mais difícil é fase de identificação do modelo ARIMA, que será utilizado para ajustar os dados. Esta escolha é feita a partir das autocorrelações e autocorrelações parciais estimadas, dentre as quais esperamos que representem adequadamente as verdadeiras quantidades teóricas, que são desconhecidas. Anteriormente definimos a *facv* e *fac*, agora vamos definir a *função de autocorrelação parcial (facp)*.

### 6.3.1 Função de Autocorrelação Parcial (facp)

Nós podemos definir a correlação entre  $Z_t$  e  $Z_{t-k}$  removendo o efeito das variáveis  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$ . Esta medida, para séries estacionárias, é chamada a *auto-correlação parcial* até a posição  $k$  e será denotada por  $\phi_{kk}$ , se  $Z_t$  é uma série normalmente distribuídos, ou seja,

$$\phi_{kk} = \text{Corr}(Z_t, Z_{t-k} / Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}) \quad (6.16)$$

onde  $\phi_{kk}$  é o coeficiente de correlação da distribuição de  $Z_t, Z_{t-k}$  condicional a  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$ . Um método geral para encontrar a *facp* para um processo estacionário com *fac*  $\rho_k$  é o seguinte : para um dado  $k$ , mostra-se que  $\phi_{kk}$  satisfaz as equações de Yule-Walker:

$$\rho_j = \phi_{k1}\rho_{j-1} + \phi_{k2}\rho_{j-2} + \dots + \phi_{kk}\rho_{j-k}, \quad j = 1, 2, \dots, k \quad (6.17)$$

Mais explicitamente

$$\begin{aligned} \rho_1 &= \phi_{k1} + \phi_{k2}\rho_1 + \dots + \phi_{kk}\rho_{j-1} \\ \rho_2 &= \phi_{k1}\rho_1 + \phi_{k2} + \dots + \phi_{kk}\rho_{k-2} \\ &\vdots \\ \rho_k &= \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \dots + \phi_{kk} \end{aligned} \quad (6.18)$$

Estas equações podem ser resolvidas sucessivamente para  $k = 1, 2, \dots$ , e obtendo-se  $\phi_{kk}$ , da seguinte maneira:

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \end{aligned} \quad (6.19)$$

e para  $k = 3$

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \quad (6.20)$$

em geral

$$\phi_{kk} = \frac{|P_k^*|}{|P_k|} \quad (6.21)$$

onde  $P_k$  é a matriz de autocorrelação, e  $P_k^*$  é a matriz  $P_k$  com a última coluna substituída pelo vetor de autocorrelação.

Pode-se demonstrar que, para os processos estudados temos:

- (i) um processo  $AR(p)$  tem *facp*  $\phi_{kk} \neq 0$ , para  $k \leq p$  e  $\phi_{kk} = 0$ , para  $k > p$ ;
- (ii) Um processo  $MA(q)$  tem *facp* que se comporta de maneira similar a *fac* de um processo  $AR(p)$ , para o  $MA(1)$  temos:

$$\phi_{22} = -\frac{\theta^2}{1 + \theta^2 + \theta^4}$$

e

$$\phi_{kk} = -\frac{\theta^k (1 - \theta^2)}{1 - \theta^{2(k+1)}}, \quad k \geq 1$$

- (iii) Um processo  $ARMA(p, q)$  tem *facp* que se comporta como a *facp* de um processo  $MA$ .

### 6.3.2 A Função de Autocorrelação Parcial Amostral (fapa)

Para série temporal observada vamos precisar calcular a estimativa da fapa. Nós chamamos esta função estimada de *função de autocorrelação parcial amostral* e denotamos por  $\hat{\phi}_{kk}$ .

Levinson (1974) e Durbin (1960) descobriram um método eficiente para obter as soluções para as equações (5.18), considerando as facp ou fapa. Eles demonstraram que independentemente as equações (5.18) podem ser resolvidas recursivamente como segue:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j} \rho_j} \quad (6.22)$$

onde

$$\phi_{kj} = \phi_{k-1,j} - \phi_{kk} \phi_{k-1,k-j}, \text{ para } j = 1, 2, \dots, k-1. \quad (6.23)$$

Por exemplo, utilizando  $\phi_{11} = \rho_1$ , nós temos

$$\phi_{22} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

Nas estimativas vamos substituir  $\hat{\rho}_k$  por  $r_k$ , e obtemos fapa  $\hat{\phi}_{kk}$ .

## 6.4 Identificação (ou Especificação) dos Modelos ARIMA

O objetivo da identificação é determinar os valores de  $p, d$  e  $q$  do modelo  $ARIMA(p, d, q)$ , além de obter estimativas preliminares dos parâmetros a serem utilizados no estágio de estimação.

### 6.4.1 Procedimento

I) Inicialmente diferenciamos a série  $Z_t$ , tantas vezes quantas necessárias, para se obter uma série estacionária, de modo que o processo  $\Delta^d Z_t$  seja reduzido a um  $ARMA(p, q)$ . O número de diferenças,  $d$ , necessárias para que o processo se torne estacionário, é alcançado quando a fac amostral de  $W_t = \Delta^d Z_t$  decresce rapidamente para zero;

II) Identificamos o processo  $ARMA(p, q)$  resultante, através da análise das fac e facp estimadas, cujo comportamentos devem ser semelhantes aqueles das respectivas quantidades teóricas ( $AR$ ,  $MA$  e  $ARMA$ ).

- Geralmente, na prática  $d=0, 1$ , ou  $2$ , e será suficiente inspecionar as primeiras 15 ou 20 autocorrelações da série e de suas diferenças.

- Convém testar se  $E(W_t) = \mu_W$  é zero, comparando  $\overline{W}$  com o seu desvio-padrão estimado. A tabela abaixo fornece as variâncias de  $\overline{W}$  para alguns modelos usuais. Se  $d=0$ ,  $\overline{W} = \overline{Z}$ .

Tabela - Variâncias Aproximadas p/  $\overline{W}$ ,  $n = N - d$

$AR(1)$	$MA(1)$	$ARMA(1, 1)$
$\frac{c_0(1+r_1)}{n(1-r_1)}$	$\frac{c_0(1+2r_1)}{n}$	$\frac{c_0}{n} \left[ 1 + \frac{2r_1^2}{r_1-r_2} \right]$
$AR(2)$	$MA(2)$	
$\frac{c_0(1+r_1)(1-2r_1^2+r_2)}{n(1-r_1)(1-r_2)}$	$\frac{c_0(1+2r_1+2r_2)}{n}$	

## 6.5 Intervalo de Confiança para a FAC Amostral

Apresentaremos a seguir o intervalo de confiança para a faca, antes porém, sabemos que a faca é definida como:

$$\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0} = r_j = \frac{c_j}{c_0}, \quad j = 0, 1, \dots, N-1 \quad (6.24)$$

onde

$$c_j = \frac{1}{N} \sum_{t=1}^{N-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})], \quad j = 0, 1, \dots, N-1 \quad (6.25)$$

sendo

$$\bar{Z} = \frac{1}{N} \sum_{t=1}^N Z_t \text{ e } r_{-j} = r_j \quad (6.26)$$

temos que a variância de  $r_j$  é

$$\text{Var}(r_j) \cong \frac{1}{N} \sum_{k=-\infty}^{\infty} [\rho_k^2 + \rho_{k+j}\rho_{k-j} - 4\rho_j\rho_k\rho_{k-j} + 2\rho_k^2\rho_j^2] \quad (6.27)$$

para um processo estacionário normal (Gaussiano).

Para um processo em que as autocorrelações são nulas para  $k > q$ , obtem-se:

$$\text{Var}(r_j) \cong \frac{1}{N} \left[ 1 + 2 \sum_{k=1}^q \rho_k^2 \right], \quad j > q \quad (6.28)$$

substituindo  $\rho_k$  por  $r_k$ , temos:

$$\sigma^2(r_j) \cong \frac{1}{N} \left[ 1 + 2 \sum_{k=1}^q r_k^2 \right], \quad j > q \quad (6.29)$$

Para  $N$  suficientemente grande, sob a hipótese  $H_0 : \rho_j = 0, j > q$  a distribuição é  $N(0, \sigma^2(r_j))$ . Assim, pode-se construir um intervalo de confiança aproximado para as autocorrelações:

$$r_j \pm t_\alpha \sigma(r_j) \quad (6.30)$$

onde  $t_\alpha$  é o valor da estatística t-Student com  $N-1$  graus de liberdade, tal que  $P(-t_\alpha < t < t_\alpha) = \alpha$ . Na prática utiliza-se  $t_\alpha = 2$ , de modo que podemos considerar  $\rho_j$  como significativamente diferente de zero se

$$|r_j| > 2\sigma(r_j), \quad j > q. \quad (6.31)$$

Para *facp* sob a hipótese que o processo é AR(p),

$$\text{Var}(\hat{\phi}_{jj}) \simeq \frac{1}{N}, \quad j > p. \quad (6.32)$$

De modo que

$$\sigma(\hat{\phi}_{jj}) \simeq \frac{1}{\sqrt{N}}, \quad j > p. \quad (6.33)$$

Além disso, para  $N$  grande,  $\hat{\phi}_{jj}$  terá distribuição aproximadamente normal, com média zero e variância (5.32), de modo que consideraremos  $\phi_{jj}$  significativamente diferente de zero se

$$|\hat{\phi}_{jj}| > \frac{2}{\sqrt{N}}, \quad j > p. \quad (6.34)$$

## 6.6 Exercícios

1) Considere os dois modelos:

$$A : Z_t = 0.9Z_{t-1} + 0.09Z_{t-2} + a_t$$

$$B : Z_t = Z_{t-1} + a_t - 0.1a_{t-1}$$

- a) Identifique e especifique os modelos ARIMA.
  - b) Qual a semelhança entre eles? (Compare os pesos  $\psi's$  e  $\pi's$ ).
- 2) Identifique e especifique cada modelo ARIMA:
- a)  $Z_t = Z_{t-1} - 0.25Z_{t-2} + a_t + 0.5a_{t-1}$
  - b)  $Z_t = 2Z_{t-1} - Z_{t-2} + a_t$
  - c)  $Z_t = 0.5Z_{t-1} + 0.5Z_{t-2} + a_t - 0.5a_{t-1} + 0.25a_{t-2}$ .
- 3) Suponha que  $\{Z_t\}$  é gerada como

$$Z_t = a_t + ca_{t-1} + ca_{t-2} + \dots + ca_0, \quad t > 0$$

- a) Encontre a função média e a facv para  $Z_t$ .  $Z_t$  é estacionária?
  - b) Encontre a função média e facv para  $\Delta Z_t$ . Esta série é estacionária?
  - c) Identifique e especifique  $Z_t$  como um modelo ARIMA.
- 4) Suponha que  $Z_t = A + Bt + X_t$  onde  $A$  e  $B$  são variáveis aleatórias independentes do passeio aleatório  $X_t$ .
- a) A série  $Z_t$  é estacionária?
  - b) A série  $\Delta Z_t$  é estacionária?
- 5) De uma série de 100 observações, obteve-se:  $r_1 = -0.49$ ,  $r_2 = 0.31$ ,  $r_3 = -0.21$ ,  $r_4 = 0.11$ , e  $|r_k| \leq 0.09$  para  $k > 4$ . Com base somente nestas informações, que modelo ARIMA nós poderíamos especificar para esta série?
- 6) Um série estacionária de tamanho  $N = 121$  produz as seguintes autocorrelações parciais estimadas:  $\hat{\phi}_{11} = 0.8$ ,  $\hat{\phi}_{22} = -0.6$ ,  $\hat{\phi}_{33} = 0.08$ , e  $\hat{\phi}_{44} = 0.00$ . Baseado somente nestas informações, que modelo poderíamos especificar para esta série?
- 7) Para uma série de 169 observações, encontramos que:  $r_1 = 0.41$ ,  $r_2 = 0.42$ ,  $r_3 = 0.26$ ,  $r_4 = 0.21$ , e  $r_5 = 0.16$ . Qual modelo se ajustaria a este padrão de autocorrelações?
- 8) As autocorrelações amostrais para a primeira diferença são dadas na tabela abaixo ( $N = 100$ ).

FACA	1	2	3	4	5	6
$Z_t$	0.97	0.97	0.93	0.85	0.80	0.71
$\Delta Z_t$	-0.42	0.18	-0.02	0.07	-0.10	-0.09

Baseado nas informações, quais modelos ARIMA poderíamos identificar para a série?

9) Para a série de tamanho 64, as facp amostrais são dadas por:

1	2	3	4	5	6
0.47	-0.34	0.20	0.02	0.15	-0.06

Quais modelos poderíamos considerar neste caso?

10) Suponha  $X_t$  um processo estacionário  $AR(1)$  com parâmetro  $\phi$ , mas que podemos somente observar

$$Z_t = X_t + N_t$$

onde  $N_t$  é um ruído branco independente de  $X_t$ .

- a) Encontre a fac para o processo  $Z_t$  em termos de  $\phi$ ,  $\sigma_X^2$ , e  $\sigma_N^2$ .
- b) Qual modelo ARMA podemos identificar para  $Z_t$ ?

## Capítulo 7

# Estimação dos Parâmetros

Neste capítulo vamos trabalhar com a estimação dos parâmetros para o modelo ARIMA, considerando a série temporal observada  $Z_1, Z_2, \dots, Z_n$ . Assumiremos que o modelo já foi especificado, isto é, nós já especificamos os valores para  $p, d$  e  $q$  utilizando os métodos do capítulo 5. Com relação a séries não-estacionárias, faremos a  $d$ -ésima diferença da série observada até torna-lá um processo estacionário ARMA( $p, q$ ). Na prática, nós trabalhamos com a  $d$ -ésima diferença da série temporal original, estimando os parâmetros, a partir, do modelo completo. Para simplificar o estudo sobre estimação, definimos  $Z_1, Z_2, \dots, Z_n$  como um processo estacionário observado da série original, depois de *diferenciada* adequadamente. Primeiramente apresentaremos as estimativas preliminares, em seguida apresentaremos os estimadores do método dos momentos, de mínimos quadrados e o estimador de máxima verossimilhança.

### 7.1 Estimativas Preliminares

É na identificação do modelo que são obtidas as estimativas preliminares, que serão utilizadas como valores iniciais, para as estimativas finais dos parâmetros. Estas estimativas são obtidas através das  $r_j$  da série  $W_t = \Delta^d Z_t$ .

#### 7.1.1 Processos AR(p)

Para esses processos, resolvemos as equações de Yule-Walker, com  $\hat{\rho}_j$  substituído por  $r_j$ . A estimativa da variância do ruído branco é:

$$\hat{\sigma}_a^2 = \hat{\gamma}_0 (1 - \phi_1 \rho_1 - \dots - \phi_p \rho_p)$$

com  $\hat{\gamma}_0$  substituído por  $c_0$ , e os  $\phi_j$  por suas estimativas  $\hat{\phi}_j$ .

#### 7.1.2 Processos MA(q)

Para esses processos, vamos utilizar a equação

$$\begin{aligned} \rho_j &= \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, \quad j = 0, 1, \dots, q \\ &= 0, \quad j > q. \end{aligned}$$

substituindo  $\rho_j$  por  $r_j$  e  $\theta_j$  por  $\hat{\theta}_j$  e a variância do ruído é estimada como

$$\hat{\sigma}_a^2 = \frac{\hat{\gamma}_0}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2}$$



onde  $\hat{\gamma}_0 = c_0$ .

### 7.1.3 Processos ARMA(p,q)

Para esses processos, obtemos as estimativas iniciais para  $\phi_1, \dots, \phi_p$ , resolvendo as equações:

$$\hat{\rho}_j = \hat{\phi}_1 \hat{\rho}_{j-1} + \dots + \hat{\phi}_p \hat{\rho}_{j-p}, \quad j = q+1, \dots, q+p$$

substituindo  $\hat{\rho}_j$  por  $r_j$ . Depois a partir das relações entre as autocorrelações  $\rho_1, \dots, \rho_q$ ,  $\phi_1, \dots, \phi_p$  e  $\theta_1, \dots, \theta_q$ , obtemos  $\hat{\theta}_1, \dots, \hat{\theta}_q$  e  $\hat{\sigma}_a^2$ .

Obs.: Na tabela 9.2 do Morettin (1987) temos estimativas iniciais para os parâmetros dos modelos mais utilizados na prática.

## 7.2 O Método dos Momentos

O método dos momentos é geralmente um dos mais fáceis, dos métodos para obter estimativas dos parâmetros. O método consiste de equacionar momentos amostrais com momentos teóricos e resolver as equações resultante para obter estimativas dos parâmetros desconhecidos.

### 7.2.1 Modelo Autoregressivo

Considere o modelo AR(1). Para este modelo nós temos:  $\rho_1 = \phi$ . Então podemos estimar  $\phi$  simplesmente por:

$$\hat{\phi} = r_1$$

Agora consideramos o caso AR(2). A relação entre os parâmetros  $\phi_1$  e  $\phi_2$  e os vários momentos é dado pelas equações de Yule-Walker:

$$\begin{aligned} \rho_1 &= \phi_1 + \rho_1 \phi_2 \\ \rho_2 &= \rho_1 \phi_1 + \phi_2 \end{aligned}$$

O método dos momentos substitui  $\rho_1$  por  $r_1$  e  $\rho_2$  por  $r_2$  para obter:

$$\begin{aligned} r_1 &= \phi_1 + r_1 \phi_2 \\ r_2 &= r_1 \phi_1 + \phi_2 \end{aligned}$$

resolvendo estas equações obtemos:

$$\hat{\phi}_1 = \frac{r_1(1-r_2)}{1-r_1^2} \quad (7.1)$$

e

$$\hat{\phi}_2 = \frac{r_2 - r_1^2}{1-r_1^2} \quad (7.2)$$

Para o caso geral AR(p) o procedimento é semelhante: substituímos  $\rho_k$  por  $r_k$  nas equações de Yule-Walker para obter:

$$\begin{aligned} r_1 &= \phi_1 + r_1 \phi_2 + \dots + r_{p-1} \phi_p \\ r_2 &= r_1 \phi_1 + \phi_2 + \dots + r_{p-2} \phi_p \\ &\vdots \\ r_p &= r_{p-1} \phi_1 + r_{p-2} \phi_2 + \dots + \phi_p \end{aligned} \quad (7.3)$$

Estas equações lineares serão resolvidas para obter  $\hat{\phi}_1, \dots, \hat{\phi}_p$ , em termos de  $r_1, \dots, r_p$ . O método recursivo de Durbin-Levinson (equações 5.22) é um algoritmo adequado para resolver estas equações, mas estará sujeito a erros, principalmente se solução está no limite de estacionariedade (na circunferência de raio unitário).

### 7.2.2 Modelos Médias-Móveis

Para o caso do modelo MA o método dos momentos, não é tão fácil. Vamos considerar o processo MA(1). Da fac do MA temos:

$$\rho_1 = \frac{-\theta}{1 + \theta^2}, \quad |\theta| < 1. \quad (7.4)$$

Podemos resolver a equação acima com relação a  $\theta$ . Se  $|r_1| < 0.5$ , então as duas raízes são dadas por:

$$-\frac{1}{2r_1} \pm \left[ \frac{1}{4r_1^2} - 1 \right]^{0.5}$$

Dentre as soluções só uma satisfaz a condição de invertibilidade, que pode ser escrita como:

$$\hat{\theta} = \frac{-1 + [1 - 4r_1^2]^{0.5}}{2r_1} \quad (7.5)$$

Se  $r_1 = \pm 0.5$ , a solução real é única, ou seja,  $\pm 1$ , mas por outro lado, não é invertível. Se  $|r_1| > 0.5$ , não existe soluções reais, e o método dos momentos não produz um estimador adequado para  $\theta$ , além disso a especificação do modelo torna-se duvidosa.

Para modelos MA(q), de ordem grande, o método dos momentos torna-se bastante complicado. As equações resultantes em função dos  $\theta's$ , não são lineares, entretanto, e suas soluções precisam ser numéricas. Porém haverá várias soluções múltiplas, das quais somente uma será invertível.

### 7.2.3 Modelos ARMA

Para esse modelo vamos considerar somente o ARMA(1,1). Da equação abaixo (capítulo 4):

$$\rho_k = \frac{(1 - \theta\phi)(\phi - \theta)}{1 - 2\theta\phi + \theta^2} \phi^{k-1}, \quad \text{para } k \geq 1 \quad (7.6)$$

Notando que  $\rho_2/\rho_1 = \phi$ , nós podemos primeiro estimar  $\phi$  como

$$\hat{\phi} = \frac{r_2}{r_1} \quad (7.7)$$

Então podemos substituir (6.7) em (6.6) e obter

$$r_1 = \frac{(1 - \theta\hat{\phi})(\hat{\phi} - \theta)}{1 - 2\hat{\phi}\theta + \theta^2} \quad (7.8)$$

par obter a estimativa de  $\hat{\theta}$ , resolvemos a equação em função de  $\theta$ , e considerando somente a solução invertível.

### 7.2.4 Estimativas da Variância do Ruído

O parâmetro final a ser estimado é a variância  $\sigma_a^2$ . Dentre todos os casos, nós podemos primeiro estimar  $\gamma_0 = Var(Z_t)$  pela variância amostral

$$S^2 = \frac{\sum_{t=1}^n (Z_t - \bar{Z})^2}{n-1} \quad (7.9)$$

e daí utilizamos a relação (capítulo 4) que existe entre  $\gamma_0$ ,  $\sigma_a^2$  e os  $\theta's$ , e  $\phi's$  para estimar  $\sigma_a^2$ .

Para os modelos AR(p), temos:

$$\hat{\sigma}_a^2 = \left(1 - \hat{\phi}_1 r_1 - \dots - \hat{\phi}_p r_p\right) S^2 \quad (7.10)$$

Em particular, para um processo AR(1),

$$\hat{\sigma}_a^2 = (1 - r_1^2) S^2 \quad (7.11)$$

desde que  $\hat{\phi}_1 = r_1$ .

Para o caso MA(q), nós temos:

$$\hat{\sigma}_a^2 = \frac{S^2}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2} \quad (7.12)$$

Para o processo ARMA(1,1), temos:

$$\hat{\sigma}_a^2 = \frac{(1 - \hat{\phi}^2)}{1 - 2\hat{\theta}\hat{\phi} + \hat{\theta}^2} S^2 \quad (7.13)$$

## 7.3 Estimativas de Mínimos Quadrados

Como já sabemos, o método dos momentos não é satisfatório para modelos com termos médias móveis, precisaríamos aplicar outros métodos, e uma das alternativas é o método de mínimos quadrados.

### 7.3.1 Modelos Autoregressivos

Considere o caso AR(1), onde

$$Z_t - \mu = \phi (Z_{t-1} - \mu) + a_t \quad (7.14)$$

Nós podemos vê-lo como um modelo de regressão com variável preditora  $Z_{t-1}$  e variável resposta  $Z_t$ . A estimação de mínimos quadrados consiste em minimizar a soma de quadrados das diferenças  $(Z_t - \mu) - \phi (Z_{t-1} - \mu)$ . Desde que somente  $Z_1, Z_2, \dots, Z_n$  são observados, nós podemos somente somar de  $t = 2$  até  $t = n$ . Seja

$$S_*(\phi, \mu) = \sum_{t=2}^n [(Z_t - \mu) - \phi (Z_{t-1} - \mu)]^2 \quad (7.15)$$

Onde  $S_*$  é geralmente chamada de *função da soma de quadrados condicional*. De acordo com o princípio de mínimos quadrados, nós estimamos  $\phi$  e  $\mu$  com respeito aos valores que minimizam  $S_*(\phi, \mu)$ , dado os valores observados da série  $Z_1, Z_2, \dots, Z_n$ .

Vamos considerar as derivadas  $\partial S_*/\partial \mu = 0$  e  $\partial S_*/\partial \phi = 0$ . Ou seja,

$$\frac{\partial S_*}{\partial \mu} = \sum_{t=2}^n 2[(Z_t - \mu) - \phi (Z_{t-1} - \mu)](-1 + \phi) = 0$$

simplificando e resolvendo para  $\mu$ , temos

$$\mu = \frac{\sum_{t=2}^n Z_t - \phi \sum_{t=2}^n Z_{t-1}}{(n-1)(1-\phi)} \quad (7.16)$$

Para  $n$  grande, temos

$$\sum_{t=2}^n \frac{Z_t}{n-1} \approx \sum_{t=2}^n \frac{Z_{t-1}}{n-1} \approx \bar{Z}$$

Além disso, temos que a equação (6.) reduz-se a

$$\hat{\mu} \approx \frac{\bar{Z} - \phi \bar{Z}}{1 - \phi} = \bar{Z} \quad (7.17)$$

Podemos dizer então que  $\hat{\mu} = \bar{Z}$ .

Considere a minimização de  $S_*(\phi, \bar{Z})$  com relação a  $\phi$

$$\frac{\partial S_*}{\partial \phi} = \sum_{t=2}^n 2 [(Z_t - \bar{Z}) - \phi (Z_{t-1} - \bar{Z})] (Z_{t-1} - \bar{Z}) = 0$$

simplificando e resolvendo para  $\phi$ , temos

$$\hat{\phi} = \frac{\sum_{t=2}^n (Z_t - \bar{Z}) (Z_{t-1} - \bar{Z})}{\sum_{t=2}^n (Z_{t-1} - \bar{Z})^2} \quad (7.18)$$

Exeto para o termo  $(Z_n - \bar{Z})^2$ ,  $\hat{\phi} = r_1$ ; além disso os estimadores de mínimos quadrados e método dos momentos são quase idênticos, especialmente para grandes amostras.

Para o processo geral  $AR(p)$ , os métodos utilizados para obter as equações (6.16) e (6.17) podem facilmente ser estendidos para produzir o mesmo resultado, ou seja,  $\hat{\mu} = \bar{Z}$ .

Para generalizar as estimativas dos  $\phi'$ s, nós podemos considerar o modelo de segunda ordem,  $AR(2)$ . Substituindo  $\hat{\mu}$  por  $\bar{Z}$  na função da soma de quadrados condicional

$$S_*(\phi_1, \phi_2) = \sum_{t=3}^n [(Z_t - \bar{Z}) - \phi_1 (Z_{t-1} - \bar{Z}) - \phi_2 (Z_{t-2} - \bar{Z})]^2 \quad (7.19)$$

Fazendo  $\partial S_*/\partial \phi_1 = 0$ , temos

$$-2 \sum_{t=3}^n [(Z_t - \bar{Z}) - \phi_1 (Z_{t-1} - \bar{Z}) - \phi_2 (Z_{t-2} - \bar{Z})] (Z_{t-1} - \bar{Z}) = 0 \quad (7.20)$$

que podemos escrever como

$$\sum_{t=3}^n (Z_t - \bar{Z}) (Z_{t-1} - \bar{Z}) = \sum_{t=3}^n (Z_{t-1} - \bar{Z})^2 \phi_1 + \sum_{t=3}^n (Z_{t-2} - \bar{Z}) (Z_{t-1} - \bar{Z}) \phi_2 \quad (7.21)$$

Agora se dividimos ambos os lados por  $\sum_{t=3}^n (Z_t - \bar{Z})^2$ , obtemos:

$$r_1 = \phi_1 + r_1 \phi_2 \quad (7.22)$$

Fazendo o mesmo para  $\partial S_*/\partial \phi_2 = 0$ , vamos ter

$$r_2 = r_1 \phi_1 + \phi_2 \quad (7.23)$$

Podemos verificar que trata-se das equações de Yule-Walker amostrais para o modelo  $AR(2)$ .

Inteiramente análogo ao resultado obtido, segue-se para o caso do modelo geral  $AR(p)$ , as estimativas de mínimos quadrados dos  $\phi'$ s, podem ser obtidas resolvendo-se as equações amostrais de Yule-Walker, e com uma boa aproximação.

### 7.3.2 Modelos Médias Móveis

Considere agora a estimação de mínimos quadrados de  $\theta$  no modelo MA(1):

$$Z_t = a_t - \theta a_{t-1} \quad (7.24)$$

Vamos supor que o modelo MA(1) é invertível, ou seja,  $|\theta| < 1$ , então podemos expressá-lo como

$$Z_t = -\theta Z_{t-1} - \theta^2 Z_{t-2} - \dots + a_t \quad (7.25)$$

que é um modelo AR( $\infty$ ). Além disso, podemos obter o estimador de mínimos quadrados, escolhendo o valor do parâmetro que minimiza

$$S_*(\theta) = \sum a_t^2 \quad (7.26)$$

onde, implicitamente,  $a_t = a_t(\theta)$  é uma função da série observada e o parâmetro  $\theta$ .

Da equação (6.25), claramente podemos observar que o problema de mínimos quadrados é não-linear nos parâmetros. Além disso, para o caso simples MA(1),  $S_*$  não pode ser minimizado analiticamente, e precisamos recorrer a técnicas de otimização numérica.

Para uma da série observada  $Z_1, Z_2, \dots, Z_n$  e um particular valor de  $\theta$ , vamos reescrever a equação (6.24) como

$$a_t = Z_t + \theta a_{t-1} \quad (7.27)$$

Utilizando a equação (6.27), os valores  $a_1, \dots, a_n$  podem ser calculados recursivamente dado um o valor inicial  $a_0$ . Um condição inicial, que se utiliza é  $a_0 = 0$ , obtendo

$$\begin{aligned} a_1 &= Z_1 \\ a_2 &= Z_2 + \theta a_1 \\ &\vdots \\ a_n &= Z_n + \theta a_{n-1} \end{aligned} \quad (7.28)$$

e além disso calculamos  $S_*(\theta) = \sum_{t=1}^n a_t^2$ , condicional a  $a_0 = 0$  para o particular valor de  $\theta$ .

Para o caso simples de um parâmetro, podemos variar  $\theta$  no intervalo  $(-1, 1)$ , que garante invertibilidade do modelo, e encontrar a soma de mínimos quadrados. Para os modelos MA(q), um algoritmo de otimização numérica, como o Gauss-Newton, é preferível. A aproximação de Gauss-Newton consiste em aproximar  $a_t = a_t(\theta)$  para uma função linear de  $\theta$  em torno da estimativa inicial  $\hat{\theta}$ . Que é

$$a_t(\theta) \approx a_t(\hat{\theta}) + (\theta - \hat{\theta}) \frac{da_t(\hat{\theta})}{d\theta} \quad (7.29)$$

Nota-se que  $da_t(\theta)/d\theta$  pode ser calculado recursivamente diferenciando ambos os lados da equação (6.27) para obter

$$\frac{da_t(\theta)}{d\theta} = \frac{\theta da_{t-1}(\theta)}{d\theta} + a_{t-1}(\theta) \quad (7.30)$$

com valor inicial  $da_t(\theta)/d\theta = 0$ .

Desde que a aproximação em (6.28) é linear em  $\theta$ , a soma de quadrados calculada pode ser minimizada analiticamente levando a uma nova estimativa de  $\theta$ . Este processo pode ser repetido com  $\hat{\theta}$  substituído pela nova estimativa, calculando as mudanças em ambos, nas estimativas de  $\theta$  e na soma de quadrados, quando a soma

torná-se pequena e o parâmetro atingir a convergência. O método dos momentos pode ser utilizado para obter as primeiras estimativas de  $\theta$ , mas em muitos casos o procedimento converge para a soma de mínimos quadrados de uma estimativa arbitrária, tal como  $\theta = 0.1$ .

Para os modelos MA de altas ordens, as idéias são análogas. Nós calculamos  $a_t = a_t(\theta_1, \dots, \theta_q)$  recursivamente de

$$a_t = Z_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (7.31)$$

como  $a_0 = a_{-1} = \dots = a_{1-q} = 0$ . A soma de quadrados é minimizada conjuntamente com relação a  $\theta_1, \theta_2, \dots, \theta_q$  utilizando o algoritmo multivariado de Gauss-Newton (Box e Jenkins, 1976).

### 7.3.3 Modelos Autoregressivos-Médias-Móveis

Vamos considerar somente o modelo ARMA(1,1):

$$Z_t = \phi Z_{t-1} + a_t - \theta a_{t-1} \quad (7.32)$$

Como no caso MA, consideramos  $a_t = a_t(\phi, \theta)$  e desejamos minimizar  $S_*(\phi, \theta) = \sum_{t=1}^n a_t^2$ . Podemos reescrever a equação (6.32) como

$$a_t = Z_t - \phi Z_{t-1} + \theta a_{t-1} \quad (7.33)$$

Para obter  $a_1$ , nós temos que supor  $Z_0 = 0$  ( $\mu = 0$ ) ou  $Z_0 = \bar{Z}$  ( $\mu = \bar{Z}$ ). Entretanto, uma melhor aproximação, seria simplesmente minimizar  $\sum_{t=1}^n a_t^2$ . As derivadas necessárias para o algoritmo de Gauss-Newton podem de novo serem obtidas recursivamente da equação (6.33).

Para o modelo geral ARMA(p,q), calculamos  $a_t = a_t(\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$  para  $t = p + 1, \dots, n$  de

$$a_t = Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (7.34)$$

com  $a_p = a_{p-1} = \dots = a_{1-q} = 0$  e então numericamente minimizamos  $\sum_{t=1}^n a_t^2$  para obter as estimativas de mínimos quadrados condicional de  $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$ . Um termo constante pode ser incluído no modelo.

Para o conjunto de parâmetros  $\theta_1, \dots, \theta_q$ , que corresponde a invertibilidade do modelo, os valores  $a_p = a_{p-1} = \dots = a_{1-q} = 0$  terão muito pouca influência nas estimativas finais dos parâmetros para grandes amostras

## 7.4 Estimativas de Máxima Verossimilhança

A vantagem dos métodos de máxima verossimilhança é que todas as informações nos dados são utilizadas, ao invés de utilizar somente os primeiros momentos, como é o caso do mínimos quadrados. Outra vantagem é que, sobre certas condições gerais, muitos resultados já são conhecidos, para o caso de grandes amostras. Entretanto, uma desvantagem é que para os primeiros valores de  $t$ , devemos trabalhar especificamente com a *função de densidade de probabilidade (f.d.p.)* conjunta.

Para um dado conjunto de observações  $Z_1, Z_2, \dots, Z_n$ , a *função de verossimilhança*  $L$  é definida como a *f.d.p.* conjunta dos dados observados, em função dos parâmetros do modelo. Para os modelos ARIMA,  $L$  será uma função de  $\phi's$ ,  $\theta's$ ,  $\mu$  e  $\sigma_a^2$  dadas as observações  $Z_1, Z_2, \dots, Z_n$ . Os estimadores de máxima verossimilhança são definidos como aqueles valores dos parâmetros, com relação aos dados observados, que são os mais verossímeis (prováveis), e maximizam a função de verossimilhança.

Inicialmente analisaremos os modelos AR(1). A suposição mais comum é que o ruído branco seja uma variável aleatória normal independente, identicamente distribuída com média 0 e variância  $\sigma_a^2$ , a f.d.p. para cada uma é

$$f(a) = (2\pi\sigma_a^2)^{-0.5} \exp\left(-\frac{a^2}{2\sigma_a^2}\right), \quad -\infty < a < \infty \quad (7.35)$$

e, como são v.a.i.i.d. a f.d.p. conjunta para  $a_2, \dots, a_n$  é

$$f(\mathbf{a}) = (2\pi\sigma_a^2)^{-(n-1)/2} \exp\left(-\frac{1}{2\sigma_a^2} \sum_{t=2}^n a_t^2\right), \quad (7.36)$$

Agora considerando

$$\begin{aligned} Z_2 - \mu &= \phi(Z_1 - \mu) + a_2 \\ Z_3 - \mu &= \phi(Z_2 - \mu) + a_3 \\ &\vdots \\ Z_n - \mu &= \phi(Z_{n-1} - \mu) + a_n \end{aligned} \quad (7.37)$$

Seja  $Z_1 = z_1$ , vamos utilizar uma transformação linear de  $a_2, \dots, a_n$  e  $Z_2, \dots, Z_n$  (com jacobiano igual a 1). Além disso, a f.d.p. conjunta de  $Z_2, \dots, Z_n/Z_1 = z_1$  pode ser obtida utilizando (6.37) pela substituição de  $a$ 's em termos dos  $z$ 's em (6.36)

$$f(z_2, \dots, z_n/z_1) = (2\pi\sigma_a^2)^{-(n-1)/2} \exp\left\{\left(-\frac{1}{2\sigma_a^2}\right) \sum_{t=2}^n [(Z_t - \mu) - \phi(Z_{t-1} - \mu)]^2\right\} \quad (7.38)$$

Agora consideramos a f.d.p. marginal de  $Z_1$ , que também é normal com média  $\mu$  e variância  $\gamma_0 = \sigma_a^2/(1 - \phi^2)$ . Multiplicando (6.38) por  $f(z_1)$ , temos

$$f(z_2, \dots, z_n/z_1) f(z_1) = \frac{f(z_2, \dots, z_n, z_1) f(z_1)}{f(z_1)} \quad (7.39)$$

que a f.d.p. conjunta de  $Z_1, Z_2, \dots, Z_n$ . Interpretada como uma função de parâmetros  $\phi, \mu$ , e  $\sigma_a^2$ , e a *função de verossimilhança* do AR(1) é dada por:

$$L(\phi, \mu, \sigma_a^2) = (2\pi\sigma_a^2)^{-n/2} (1 - \phi^2)^{0.5} \exp\left[-\frac{1}{2\sigma_a^2} S(\phi, \mu)\right] \quad (7.40)$$

onde

$$S(\phi, \mu) = \sum_{t=2}^n [(Z_t - \mu) - \phi(Z_{t-1} - \mu)]^2 + (1 - \phi^2)(Z_1 - \mu)^2 \quad (7.41)$$

A soma  $S(\phi, \mu)$  é chamada a *soma de quadrados não-condicional*.

Como regra geral o logaritmo da função de verossimilhança matematicamente, é mais conveniente, do que a anterior. Para o caso AR(1), a função log-verossimilhança, é dada por:

$$\lambda(\phi, \mu, \sigma_a^2) = \frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma_a^2 + \frac{1}{2} \log(1 - \phi^2) - \frac{1}{2\sigma_a^2} S(\phi, \mu) \quad (7.42)$$

Dado os valores de  $\phi$  e  $\mu$ ,  $\lambda(\phi, \mu, \sigma_a^2)$  pode ser maximizada analiticamente com relação a  $\sigma_a^2$ :

$$\frac{\partial \lambda}{\partial \sigma_a^2} = -\frac{n}{2} \frac{1}{\sigma_a^2} + \frac{1}{2} (\sigma_a^2)^{-2} S(\phi, \mu) = 0 \quad (7.43)$$

determinamos então o estimador de  $\sigma_a^2$ , em termos de  $\phi$  e  $\mu$  :

$$\widehat{\sigma_a^2} = \frac{S(\phi, \mu)}{n} \quad (7.44)$$

ou

$$\widehat{\sigma_a^2} = \frac{S(\phi, \mu)}{n-2} \quad (7.45)$$

na primeira expressão o vício é menor.

Considere agora a estimação de  $\phi$  e  $\mu$ . Uma comparação entre as duas soma de quadrados, condicional,  $S_*(\phi, \mu)$ , e não-condicional,  $S(\phi, \mu)$ , vamos ter:

$$S(\phi, \mu) = S_*(\phi, \mu) + (1 - \phi^2)(Z_1 - \mu)^2 \quad (7.46)$$

os valores de  $\phi$  e  $\mu$  que minimizam  $S_*$  ou  $S$  são semelhantes e

$$S_*(\phi, \mu) \approx S(\phi, \mu) \quad (7.47)$$

## 7.5 Mínimos Quadrados Não-Condicionais para o Modelo ARMA

A definição da função de verossimilhança é associada a soma de quadrados não-condicional. Para o modelo ARMA é consideravelmente mais complicada, então citaremos a expressão da forma da função da soma de quadrados não-condicional. Segundo Box e Jenkins (1976), temos que para o modelo geral ARMA( $p, q$ )

$$S(\phi, \theta, \theta_0) = \sum_{t=-\infty}^n a_t^2 \quad (7.48)$$

onde

$$\hat{a}_t = E(a_t / Z_1, \dots, Z_n) \quad (7.49)$$

Note que, para  $t \leq 0$ , o  $\hat{a}_t$  será visto como “previsão” *backward no tempo* dos termos de  $a_t$ , dado  $Z_1, \dots, Z_n$ . Por esta razão, eles frequentemente tem sido chamados *back forecasts*, mas o termo mais apropriado é *backcasts*.

## 7.6 Propriedades das Estimativas

As propriedades dos estimadores de máxima verossimilhança e mínimos quadrados (condicional e não-condicional), são iguais para o caso de grandes amostras, e podem ser obtidas modificando-se a teoria de máxima verossimilhança nesses casos. Para mais detalhes consultar Box e Jenkins (1976) e Feller (1976).

Para  $n$  suficientemente grande, os estimadores são não-viciados e seguem aproximadamente uma distribuição normal. As variâncias e correlações, para o AR, são as seguintes,

$$\begin{aligned} AR(1) : \quad Var(\hat{\phi}) &\approx \frac{1 - \phi^2}{n} \\ AR(2) : \quad Var(\hat{\phi}_1) &\approx Var(\hat{\phi}_2) \approx \frac{1 - \phi_2^2}{n} \\ &: \quad Corr(\hat{\phi}_1, \hat{\phi}_2) \approx -\frac{\phi_1}{1 - \phi_2} = -\rho_1 \end{aligned} \quad (7.50)$$



para o MA

$$\begin{aligned} MA(1) &: Var(\hat{\theta}) \approx \frac{1 - \theta^2}{n} \\ MA(2) &: Var(\hat{\theta}_1) \approx Var(\hat{\theta}_2) \approx \frac{1 - \theta_2^2}{n} \\ &: Corr(\hat{\theta}_1, \hat{\theta}_2) \approx \frac{-\theta_1}{1 - \theta_2} \end{aligned} \quad (7.51)$$

e para o modelo ARMA

$$\begin{aligned} ARMA(1,1) &: Var(\hat{\phi}) \approx \frac{1 - \phi^2}{n} \left( \frac{1 - \phi\theta}{\phi - \theta} \right)^2 \\ &: Var(\hat{\theta}) \approx \frac{1 - \theta^2}{n} \left( \frac{1 - \phi\theta}{\phi - \theta} \right)^2 \\ &: Corr(\hat{\phi}, \hat{\theta}) \approx \frac{[(1 - \phi^2)(1 - \theta^2)]^{0.5}}{1 - \phi\theta} \end{aligned} \quad (7.52)$$

## 7.7 Exercícios

1) Para uma série de tamanho 100, foi calculado  $r_1 = 0.8$ ,  $r_2 = 0.5$ ,  $r_3 = 0.4$ ,  $\bar{Z} = 2$ , e uma variância amostral de 5. Se assumimos que o modelo AR(2) com um termo constante é adequado, como podemos obter as estimativas (simples) de  $\phi_1$ ,  $\phi_2$ ,  $\theta_0$ , e  $\sigma_a^2$ ?

2) Assumindo que os seguintes dados sejam de um processo estacionário, calcule as estimativas de  $\mu$ ,  $\gamma_0$ , e  $\rho_1$ :

$$6, \quad 5, \quad 4, \quad 6, \quad 4$$

3) Se  $\{Z_t\}$  satisfaz um modelo AR(1) com  $\phi \simeq 0.7$ , de que modo podemos estimar o verdadeiro valor de  $\phi = \rho_1$  com uma confiança 95% que o nosso erro  $\leq \pm 0.1$ ?

4) Considere um processo MA(1) para o qual sabemos que a média é zero. Baseado num série de tamanho 3, nós observamos  $Z_1 = 0$ ,  $Z_2 = -1$ , e  $Z_3 = 0.5$ .

a) Mostrar que a estimativa de mínimos quadrados condicional de  $\theta = 0.5$ .

b) Encontrar uma estimativa da variância do ruído  $\sigma_a^2$ .

5) Dado os dados  $Z_0 = 10$ ,  $Z_1 = 10$ ,  $Z_2 = 9$ , e  $Z_3 = 9.5$ , desejamos ajustar um modelo IMA(1,1) com um termo constante.

a) Encontrar a estimativa de mínimos quadrados condicional de  $\theta$ .

b) Estime  $\sigma_a^2$ .

6) Considere duas parametrizações do modelo AR(1):

$$I. Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t$$

$$II. Z_t = \phi Z_{t-1} + \theta_0 + a_t$$

onde  $\theta_0 = \mu(1 - \phi)$ . Estimar  $\phi$  e  $\mu$  ou  $\phi$  e  $\theta_0$  utilizando o mínimos quadrados condicionado a  $Z_1$ .

Mostrar que com o modelo I nós somos levado a resolver equações não-lineares para obter as estimativas, enquanto que no modelo II precisamos somente resolver equações lineares.

7) Suponha que para um modelo ARMA(1,1), com  $N=152$ , foi obtido:  $\phi = 0.85$ ,  $\theta = -0.6$  e  $\sigma_a^2 = 0.086$ . Obtenha intervalos de confiança para  $\phi$  e  $\theta$ , com nível de 95% de confiança.

8) Uma série com 400 observações apresentou os seguintes resultados:

$j$	1	2	3	4	5	6	7
$\phi_{jj}$	0.8	-0.5	0.07	-0.02	-0.01	0.05	0.04

- Explique por que podemos ajustar à série um modelo AR(2);
- Obtenha as estimativas de  $\phi_1$  e  $\phi_2$  do modelo AR(2) utilizando as equações de Yule-Walker, e as estimativas do termo  $\theta_0$  e a  $Var(a_t)$ ;
- Verifique se o modelo ajustado satisfaz as condições de estacionariedade;
- Utilizando  $\hat{\phi}_1$  e  $\hat{\phi}_2$  como sendo os verdadeiros valores de  $\phi_1$  e  $\phi_2$  do processo AR(2), e determine os valores de  $\hat{\rho}_1, \hat{\rho}_2$  e  $\hat{\rho}_3$ .

9) Suponha que um programa de identificação forneceu os seguintes resultados:

$j$	1	2	3	4	5	6
$r_j$	-0.82	0.41	-0.12	0.08	-0.09	0.05
$\phi_{jj}$	-0.82	-0.43	-0.05	0.25	0.20	0.12

$N = 100, \bar{Z} = 0.08, S_Z^2 = 2.4$ .

Identifique um modelo para  $Z_t$  e obtenha as estimativas preliminares dos parâmetros.

## Capítulo 8

# Diagnóstico do Modelo

Nesta seção faremos o diagnóstico do modelo, começando com a análise residual, e depois análise de modelos sobre-parametrizada, que são, modelos mais gerais do que o modelo especificado mas contém o modelo especificado como um caso especial.

### 8.1 Análise Residual

No capítulo 3 apresentamos algumas idéias de análise residual para checar a adequacidade de uma tendência determinística ajustada. No modelo autoregressivo definimos facilmente os resíduos. Considere em particular um modelo AR(2) com um termo constante:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \theta_0 + a_t \quad (8.1)$$

tendo estimado  $\phi_1, \phi_2$ , e  $\theta_0$ , os resíduos são definidos como:

$$\hat{a}_t = Z_t - \hat{\phi}_1 Z_{t-1} - \hat{\phi}_2 Z_{t-2} - \hat{\theta}_0 \quad (8.2)$$

Os resíduos iniciais  $\hat{a}_1$  e  $\hat{a}_2$  podem ser obtidos do procedimento de estimação utilizando valores passados para  $Z_0$  e  $Z_{-1}$ .

Para o modelo geral ARMA, precisamos colocar na forma autoregressiva de ordem infinita, para definir os resíduos, apresentada no capítulo 4, teremos então a seguinte equação

$$Z_t = \sum_{j=1}^{\infty} \pi_j Z_{t-j} + a_t \quad (8.3)$$

os resíduos serão definidos como

$$\hat{a}_t = Z_t - \sum_{j=1}^{\infty} \hat{\pi}_j Z_{t-j} \quad (8.4)$$

Aqui os  $\pi_j$  não são estimados diretamente a expressão acima, mas de funções implícitas de  $\phi's$  e  $\theta's$ , utilizando a equação (6.34). No capítulo 8 veremos que a equação

$$\hat{Z}_t = \sum_{j=1}^{\infty} \hat{\pi}_j Z_{t-j} \quad (8.5)$$

é a melhor previsão de  $Z_t$  baseada em  $Z_{t-1}, Z_{t-2}, \dots$ . Além disso a equação (7.4) pode ser escrita como

$$\text{Resíduos} = \text{Real} - \text{Predito(estimado)}$$

ou seja,

$$\hat{a}_t = Z_t - \hat{Z}_t \quad (8.6)$$

## 8.2 Autocorrelação dos Resíduos

Para checar a independência do ruído (erro) no modelo, vamos considerar inicialmente a função de autocorrelação residual amostral,  $\hat{r}_k$ . Para um  $n$  suficientemente grande, verifica-se que os  $\hat{r}_k$  são normalmente distribuídos com média zero,

$$Var(\hat{r}_k) \approx \frac{1}{n} \quad (8.7)$$

e

$$Corr(\hat{r}_k, \hat{r}_j) \approx 0, \text{ para } k \neq j \quad (8.8)$$

Sempre sob a suposição que o modelo ajustado é apropriado. As autocorrelações  $\hat{r}_k$  são calculadas por

$$\hat{r}_k = \frac{\sum_{t=k+1}^n \hat{a}_t \hat{a}_{t-k}}{\sum_{t=1}^n \hat{a}_t^2} \quad (8.9)$$

Se considerarmos um modelo AR(1) especificado e estimado corretamente, mostra-se que para um  $n$  grande

$$Var(\hat{r}_1) \approx \frac{\phi^2}{n}, \quad (8.10)$$

$$Var(\hat{r}_k) \approx \frac{1 - (1 - \phi^2) \phi^{2k-2}}{n}, \quad k > 1,$$

e

$$Corr(\hat{r}_1, \hat{r}_k) \approx -\sin(\phi) \frac{(1 - \phi^2) \phi^{2k-2}}{1 - (1 - \phi^2) \phi^{2k-2}}, \quad k > 1 \quad (8.11)$$

onde

$$\sin(\phi) = \begin{cases} 1, & \text{se } \phi > 0 \\ 0, & \text{se } \phi = 0 \\ -1, & \text{se } \phi < 0 \end{cases}$$

Para um modelo AR(2) mostra-se que:

$$Var(\hat{r}_1) \approx \frac{\phi_2^2}{n} \quad (8.12)$$

e

$$Var(\hat{r}_2) \approx \frac{\phi_2^2 + \phi_1^2 (1 + \phi_2)^2}{n} \quad (8.13)$$

Se os parâmetros  $\phi_1$  e  $\phi_2$  não estiverem na região de estacionariedade, então

$$Var(\hat{r}_k) \approx \frac{1}{n}, \text{ para } k \geq 3 \quad (8.14)$$

## 8.3 O Teste de Box-Pierce

Box e Pierce (1970) propuseram um teste para as autocorrelações dos resíduos estimados, que, apesar de não detectar quebras específicas no comportamento de ruído branco, pode indicar se esses valores são muito altos. Se o modelo for adequado, a estatística:

$$Q = n(n+2) \sum_{k=1}^K \frac{\hat{r}_k^2}{(n-k)} \quad (8.15)$$

tem uma distribuição Qui-Quadrado,  $\chi^2$ , com  $K - p - q$  graus de liberdade (Ljung e Box, 1978). A hipótese de ruído branco para os resíduos é rejeitada para valores grandes de  $Q$ . Em geral basta tomar as primeiras 20 ou 25 primeiras  $\hat{r}_k$ .

## 8.4 Exercícios

1) Suponha que os resíduos obtidos ajustando-se o modelo  $\Delta Z_t = (1 - 0.6X) a_t$  a uma série com N=127 observações, forneceram as seguintes autocorrelações residuais:

$k$	1	2	3	4	5	6	7	8	9	10
$r_k(a)$	-0.4	0.02	-0.07	-0.01	-0.07	-0.02	-0.15	-0.07	0.04	0.02

a) Verifique se há valores discrepantes;

b) Use o teste de Box-Pierce para verificar se o modelo é adequado.

2) Suponha que os resíduos obtidos, ajustando-se o modelo AR(2),  $\hat{\phi}_1 = 1.56$  e  $\hat{\phi}_2 = 0.66$ , de uma série com N=120 observações, forneceu as seguintes autocorrelações residuais:

$j$	1	2	3	4	5	6	7	8	9	10	11	12
$\hat{r}_a(j)$	0.18	-0.06	-0.04	-0.11	-0.04	0.13	0.19	-0.14	0.07	0.09	0.11	-0.09

a) Verifique através do teste de Box-Pierce se o modelo é adequado.

b) Há indicações de valores discrepantes?

3) Suponha que os resíduos  $\hat{a}_t$  do modelo  $(1 - X) Z_t = (1 + 0.6X) a_t$  ajustados a uma série com N=80 observações, forneceu as seguintes autocorrelações residuais:

$k$	1	2	3	4	5	6	7	8	9	10
$\hat{r}_a(k)$	0.39	0.2	0.09	0.04	0.09	-0.13	-0.05	0.06	0.11	-0.02

a) Verifique se o modelo é adequado utilizando a estatística de Box-Pierce.

b) Há indicações de valores discrepantes?

## Capítulo 9

# Previsão com Modelos ARIMA

Um dos principais objetivos da análise de séries temporais é fazer previsões futuras, a partir do modelo identificado, estimado e adequado. Neste capítulo vamos considerar os cálculos para as previsões, e apresentaremos, também, as suas propriedades para os modelos com tendência determinística e para modelos ARIMA.

### 9.1 Cálculo da Previsão de Erro Quadrático Médio Mínimo

Seja a série  $Z_t, Z_{t-1}, \dots, Z_1$ , desejamos prever um valor  $Z_{t+h}$ , ou seja, a origem  $t$  e o horizonte  $h$ . A previsão de *erro quadrático médio mínimo* (EQMM), denotada por  $\hat{Z}_t(h)$ , é dada por

$$\hat{Z}_t(h) = E(Z_{t+h}/Z_t, Z_{t-1}, \dots, Z_1) \quad (9.1)$$

(Rever as propriedades de esperança condicional e EQMM de Predição nos apêndices F e G).

### 9.2 Tendência Determinística

Considere o modelo com tendência determinística

$$Z_t = \mu_t + X_t \quad (9.2)$$

onde o componente estocástico,  $X_t$ , tem média zero e variância  $\sigma_a^2$ .

Para o modelo na equação (8.2) temos

$$\begin{aligned} \hat{Z}_t(h) &= E(\mu_{t+h} + X_{t+h}/Z_t, Z_{t-1}, \dots, Z_1) \\ &= E(\mu_{t+h}/Z_t, Z_{t-1}, \dots, Z_1) + E(X_{t+h}/Z_t, Z_{t-1}, \dots, Z_1) \\ &= \mu_{t+h} + E(X_{t+h}) \end{aligned}$$

ou

$$\hat{Z}_t(h) = \mu_{t+h}, \quad h \geq 1 \quad (9.3)$$

desde que para  $h \geq 1$ ,  $X_{t+h}$  seja independente de  $Z_t, Z_{t-1}, \dots, Z_1$ .

Para o caso da tendência linear,  $\mu_t = \beta_0 + \beta_1 t$ , a previsão é

$$\hat{Z}_t(h) = \beta_0 + \beta_1(t+h) \quad (9.4)$$

Para uma tendência sazonal de 12 meses,  $\mu_t = \mu_{t+12}$ , nossa previsão será  $\hat{Z}_t(h) = \mu_{t+h} = \mu_{t+12+h} = \hat{Z}_t(h+12)$ . Além disso, a previsão será periódica. O erro de previsão,  $e_t(h)$ , é dado por

$$\begin{aligned} e_t(h) &= Z_{t+h} - \hat{Z}_t(h) \\ &= \mu_{t+h} + X_{t+h} - \mu_{t+h} \\ &= X_{t+h} \end{aligned}$$

então

$$E(e_t(h)) = E(X_{t+h}) = 0 \quad (9.5)$$

estas previsões são não-viciadas, e

$$Var[e_t(h)] = Var(X_{t+h}) = \gamma_0 \quad (9.6)$$

e a variância do erro de previsão.

### 9.3 Previsão ARIMA

Para os modelos ARIMA as previsões podem ser obtidas por diferentes caminhos. Cada expressão contribui para o nosso entendimento dos procedimentos de previsão com respeito a cálculo, atualização, precisão, ou comportamento da previsão a longo prazo.

#### 9.3.1 Modelo AR(1)

Considere o processo AR(1) com média diferente de zero

$$Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t \quad (9.7)$$

Considere a previsão de 1 unidade no tempo futuro (horizonte:  $h = 1$ ). Substituindo  $t$  por  $t + 1$  na equação (8.7), temos

$$Z_{t+1} - \mu = \phi(Z_t - \mu) + a_{t+1} \quad (9.8)$$

Dado que  $Z_t, Z_{t-1}, \dots, Z_1$ , tomamos a esperança condicional de ambos os lados da equação acima e obtemos:

$$\hat{Z}_t(1) - \mu = \phi E[(Z_t/Z_t, Z_{t-1}, \dots, Z_1) - \mu] + E[a_{t+1}/Z_t, Z_{t-1}, \dots, Z_1] \quad (9.9)$$

Agora das propriedades de esperança condicional sabemos que  $E(H(X)/X = x) = H(x)$ , temos

$$E(Z_t/Z_t, Z_{t-1}, \dots, Z_1) = Z_t \quad (9.10)$$

Além disso  $a_{t+1}$  é independente de  $Z_t, Z_{t-1}, \dots, Z_1$ , obtemos então a seguinte equação

$$E[a_{t+1}/Z_t, Z_{t-1}, \dots, Z_1] = E(a_{t+1}) = 0 \quad (9.11)$$

Sendo assim, podemos escrever

$$\hat{Z}_t(1) = \mu + \phi(Z_t - \mu) \quad (9.12)$$

Vimos que, uma proporção  $\phi$  multiplica o desvio  $(Z_t - \mu)$  somado com a média teremos o próximo valor futuro (horizonte).

Agora para o caso geral para um tempo futuro  $h$ . Substituindo  $t$  por  $t + h$  na equação (8.7), e tomando as esperanças condicionais de ambos os lados, obtemos

$$\hat{Z}_t(h) = \mu + \phi(\hat{Z}_t(h-1) - \mu), \quad h \geq 1 \quad (9.13)$$

desde que  $E(Z_{t+h-1}/Z_t, Z_{t-1}, \dots, Z_1) = \hat{Z}_t(h-1)$  e, para  $h \geq 1$ ,  $a_{t+h}$  é independente de  $Z_t, Z_{t-1}, \dots, Z_1$ . A equação (8.13) é recursiva em  $h$ , e é chamada de *equação de diferença de previsão*. A partir da recursão obtemos a equação

$$\begin{aligned}\hat{Z}_t(h) &= \phi(\hat{Z}_t(h-1) - \mu) + \mu \\ &= \phi(\phi[\hat{Z}_t(h-2) - \mu]) + \mu \\ &\vdots \\ &= \phi^{h-1}[\hat{Z}_t(1) - \mu] + \mu\end{aligned}$$

ou

$$\hat{Z}_t(h) = \mu + \phi^h(Z_t - \mu), \quad h \geq 1 \quad (9.14)$$

No caso geral, desde que  $|\phi| < 1$ , temos simplesmente

$$\hat{Z}_t(h) \approx \mu, \quad \text{para } h \text{ grande} \quad (9.15)$$

### 9.3.2 Erro de Previsão

O erro de previsão a 1 passo,  $e_t(1)$ , é dado como

$$\begin{aligned}e_t(1) &= Z_{t+1} - \hat{Z}_t(1) \\ &= \phi(Z_t - \mu) + \mu + a_{t+1} - [\mu + \phi(Z_t - \mu)]\end{aligned}$$

ou

$$e_t(1) = a_{t+1} \quad (9.16)$$

Da equação (8.16) veremos que a variância do erro de previsão a 1 passo é dada por

$$\text{Var}[e_t(1)] = \sigma_a^2 \quad (9.17)$$

Para investigar as propriedades do erro de previsão a longo prazo, é conveniente expressar o modelo AR(1) na forma de um processo linear geral, ou MA( $\infty$ )

$$Z_t = \mu + a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots \quad (9.18)$$

Temos então

$$\begin{aligned}e_t(h) &= Z_{t+h} - \mu - \phi^h(Z_t - \mu) \\ &= a_{t+h} + \phi a_{t+h-1} + \dots + \phi^{h-1} a_{t+1} + \phi^h a_t + \dots - \phi^h(a_t + \phi a_{t-1} + \dots)\end{aligned}$$

vamos ter

$$e_t(h) = a_{t+h} + \phi a_{t+h-1} + \dots + \phi^{h-1} a_{t+1} \quad (9.19)$$

que pode ser escrito como

$$e_t(h) = \sum_{j=0}^{h-1} \psi_j a_{t+h-j} \quad (9.20)$$

esta equação pode ser utilizada para o caso geral dos modelos ARIMA. Note que  $E[e_t(h)] = 0$ , assim as previsões são não-viciadas. Portanto da equação (8.20) temos

$$\text{Var}[e_t(h)] = \sigma_a^2 \sum_{j=0}^{h-1} \psi_j^2 \quad (9.21)$$

Nós vimos que a variância do erro de previsão cresce quando  $h$  aumenta.



Em particular para o caso AR(1)

$$\text{Var}[e_t(h)] = \sigma_a^2 \frac{1 - \phi^{2h}}{1 - \phi^2}, \quad h > 1 \quad (9.22)$$

Para o tempo a longo prazo temos

$$\text{Var}[e_t(h)] \approx \sigma_a^2 \frac{1}{1 - \phi^2}, \quad \text{para } h \text{ grande} \quad (9.23)$$

ou

$$\text{Var}[e_t(h)] \approx \text{Var}(Z_t) = \gamma_0, \quad \text{para } h \text{ grande} \quad (9.24)$$

A equação (8.24) será válida para todos os processos ARMA estacionários.

### 9.3.3 Modelo MA(1)

Vamos considerar o caso MA(1) com a média diferente de zero:

$$Z_t = \mu + a_t - \theta a_{t-1} \quad (9.25)$$

De novo substituindo  $t$  por  $t + 1$  na equação (8.25) e tomando as esperanças de ambos os lados, temos

$$\hat{Z}_t(1) = \mu - \theta E(a_t/Z_t, Z_{t-1}, \dots, Z_1) \quad (9.26)$$

Entretanto, vimos no capítulo 4 que, para um modelo invertível,  $a_t$  é uma função de  $Z_t, Z_{t-1}, \dots$ . Assim utilizando as propriedades de esperança condicional temos

$$E(a_t/Z_t, Z_{t-1}, \dots, Z_1) = a_t \quad (9.27)$$

Utilizando (8.27) e (8.26), temos que a previsão a 1 passo para o modelo MA(1) invertível é:

$$\hat{Z}_t(1) = \mu - \theta a_t \quad (9.28)$$

### 9.3.4 Erro de Previsão

O erro de previsão a 1 passo é dado por

$$\begin{aligned} e_t(1) &= Z_{t+1} - \hat{Z}_t(1) \\ &= \mu + a_{t+1} - \theta a_t - (\mu - \theta a_t) \\ e_t(1) &= a_{t+1} \end{aligned}$$

do processo linear geral temos ainda que

$$e_t(h) = \sum_{j=0}^{h-1} \psi_j a_{t+h-j}$$

e

$$\text{Var}[e_t(h)] = \sigma_a^2 \sum_{j=0}^{h-1} \psi_j^2$$

com  $\psi_1 = -\theta$  e  $\psi_j = 0$  para  $j > 1$ .

Para um tempo a longo prazo temos

$$\hat{Z}_t(h) = \mu + E(a_{t+h}/Z_t, Z_{t-1}, \dots, Z_1) - \theta E(a_{t+h-1}/Z_t, Z_{t-1}, \dots, Z_1)$$

Mas para  $h > 1$  ambos  $a_{t+h}$  e  $a_{t+h-1}$  são independentes de  $Z_t, Z_{t-1}, \dots, Z_1$ . Consequentemente, estes valores esperados condicionais são zero, e temos

$$\hat{Z}_t(h) = \mu, \quad \text{para } h > 1 \quad (9.29)$$

### 9.3.5 O Caminho Aleatório

Para ilustrar a previsão com séries não-estacionárias, considere o caminho aleatório definido por

$$Z_t = Z_{t-1} + \theta_0 + a_t, \text{ para } t \geq -m \quad (9.30)$$

Aqui

$$\hat{Z}_t(1) = E(Z_t/Z_t, Z_{t-1}, \dots, Z_1) + \theta_0 + \theta E(a_{t+1}/Z_t, Z_{t-1}, \dots, Z_1)$$

e

$$\hat{Z}_t(1) = Z_t + \theta_0 \quad (9.31)$$

Similarmente, a forma da equação diferença para o  $h$  a longo prazo é

$$\hat{Z}_t(h) = \hat{Z}_t(h-1) + \theta_0, \quad h \geq 1 \quad (9.32)$$

Recursivamente podemos obter a expressão

$$\hat{Z}_t(h) = Z_t + h\theta_0, \quad h \geq 1$$

Se  $\theta_0 \neq 0$ , a previsão não converge a longo prazo, mas segue uma linha reta com inclinação  $\theta_0$  para todo  $h$ .

Note que a presença ou falta do termo constante  $\theta_0$  altera significativamente a previsão. Portanto, os termos constantes não são incluídos nos modelos ARIMA a menos que, haja evidência que a média na série diferenciada seja significativamente diferente de zero.

### 9.3.6 Erro de Previsão

Vimos que nos modelos AR(1) e MA(1), o erro de previsão a 1 passo é

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_{t+1}$$

Também aqui temos

$$\begin{aligned} e_t(h) &= Z_{t+h} - \hat{Z}_t(h) \\ &= (Z_t + h\theta_0 + a_{t+1} + \dots + a_{t+h}) - (Z_t + h\theta_0) \\ &= \sum_{j=0}^{h-1} a_{t+h-j} \end{aligned}$$

de acordo com a equação (8.20), desde que no modelo  $\psi_j = 1$  para todo  $j$ . Da equação (8.21) nós temos

$$Var[e_t(h)] = \sigma_a^2 \sum_{j=0}^{h-1} 1^2$$

que é,

$$Var[e_t(h)] = h\sigma_a^2, \quad h \geq 1 \quad (9.33)$$

Em contraste com o caso estacionário, aqui  $Var[e_t(h)]$  cresce sem limite a medida que o  $h$  aumenta. Nós vimos que esta propriedade é característica da variância do erro de previsão para todos os processos ARIMA não-estacionários.

## 9.4 Modelo ARMA Estacionário - Caso Geral

Para o caso geral do modelo ARMA( $p, q$ ) estacionário e invertível, a *forma da equação diferença* para o cálculo da previsão é dada por

$$\begin{aligned}\hat{Z}_t(h) = & \phi_1 \hat{Z}_t(h-1) + \phi_2 \hat{Z}_t(h-2) + \dots + \phi_p \hat{Z}_t(h-p) + \theta_0 \\ & - \theta_1 E(a_{t+h-1}/Z_t, Z_{t-1}, \dots, Z_1) - \theta_2 E(a_{t+h-2}/Z_t, Z_{t-1}, \dots, Z_1) \\ & - \dots - \theta E(a_{t+h-q}/Z_t, Z_{t-1}, \dots, Z_1)\end{aligned}\quad (9.34)$$

onde

$$E(a_{t+j}/Z_t, Z_{t-1}, \dots, Z_1) = \begin{cases} 0, & \text{para } j \geq 1 \\ a_{t+j}, & \text{para } j \leq 0 \end{cases} \quad (9.35)$$

Notamos que  $\hat{Z}_t(j)$  é uma previsão verdadeira para  $j > 0$ , mas

$$\hat{Z}_t(j) = Z_{t+j}, \text{ para } -(p-1) \leq j \leq 0 \quad (9.36)$$

### 9.4.1 Modelo ARMA(1,1)

Considere o modelo ARMA(1,1), temos

$$\hat{Z}_t(1) = \phi Z_t + \theta_0 - \theta a_t \quad (9.37)$$

com

$$\hat{Z}_t(2) = \phi \hat{Z}_t(1) + \theta_0$$

e para caso mais geral

$$\hat{Z}_t(h) = \phi \hat{Z}_t(h-1) + \theta_0 \text{ para } h \geq 2 \quad (9.38)$$

As equações (8.37) e (8.38) podem ser resolvidas recursivamente para obter-se a expressão

$$\hat{Z}_t(h) = \mu + \phi^h(Z_t - \mu) - \phi^{h-1}\theta a_t, \text{ para } h \geq 1 \quad (9.39)$$

Como as equações (8.34) e (8.35) indicam, os termos do ruído  $a_t, a_{t-1}, \dots, a_{t-(q-1)}$  aparecem diretamente no cálculo das previsões para  $h = 1, 2, \dots, q$ . Entretanto, para  $h > q$  tomamos a parte autoregressiva da equação diferença, ou seja,

$$\hat{Z}_t(h) = \phi_1 \hat{Z}_t(h-1) + \phi_2 \hat{Z}_t(h-2) + \dots + \phi_p \hat{Z}_t(h-p) + \theta_0 \text{ para } h > q \quad (9.40)$$

Se considerarmos  $\theta_0 = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ , podemos escrever (8.40) como

$$\hat{Z}_t(h) - \mu = \phi_1 [\hat{Z}_t(h-1) - \mu] + \phi_2 [\hat{Z}_t(h-2) - \mu] + \dots + \phi_p [\hat{Z}_t(h-p) - \mu] \text{ para } h > q \quad (9.41)$$

No processo estacionário ARMA,  $\hat{Z}_t(h) - \mu$  decai para zero quando  $h$  aumenta, e a longo prazo a previsão é simplesmente a média do processo, como visto anteriormente.

## 9.5 Modelos Não-Estacionários ARIMA

Como foi mostrado no caminho aleatório, a previsão dos modelos ARIMA é semelhante a previsões para o modelo estacionários ARMA, mas existe algumas diferenças. Vamos considerar um modelo ARIMA( $p, 1, q$ ), que pode ser escrito como um ARMA( $p+1, q$ ) não-estacionário, ou seja,

$$\begin{aligned}Z_t = & \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \dots + \varphi_{p+1} Z_{t-p-1} + \theta_0 \\ & + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}\end{aligned}$$

para  $t > -m$  (9.42)

onde

$$\begin{aligned}\varphi_1 &= 1 + \phi_1 \\ \varphi_j &= \phi_j - \phi_{j-1}, \quad j = 2, 3, \dots, p \\ \varphi_{p+1} &= -\phi_p\end{aligned}\tag{9.43}$$

para a ordem geral de  $d$  diferenças, substituímos  $p + d$  por  $p + 1$ , e vamos ter  $p + d$  coeficientes  $\varphi$ .

### 9.5.1 Modelo ARIMA(1,1,1)

Considere o modelo

$$Z_t - Z_{t-1} = \phi(Z_{t-1} - Z_{t-2}) + \theta_0 + a_t - \theta a_{t-1}\tag{9.44}$$

ou

$$Z_t = (1 + \phi)Z_{t-1} - \phi Z_{t-2} + \theta_0 + a_t - \theta a_{t-1}\tag{9.45}$$

A previsão será dada por

$$\begin{aligned}\widehat{Z}_t(1) &= (1 + \phi)Z_t - \phi Z_{t-1} + \theta_0 - \theta a_t \\ \widehat{Z}_t(2) &= (1 + \phi)\widehat{Z}_t(1) - \phi Z_t + \theta_0 \\ &\vdots \\ \widehat{Z}_t(h) &= (1 + \phi)\widehat{Z}_t(h-1) - \phi\widehat{Z}_t(h-2) + \theta_0 \text{ para } h > 2\end{aligned}\tag{9.46}$$

### 9.5.2 Erro de Previsão

O erro de previsão será dado por

$$e_t(h) = \sum_{j=0}^{h-1} \psi_j a_{t+h-j}, \quad h \geq 1\tag{9.47}$$

com

$$E[e_t(h)] = 0, \quad h \geq 1\tag{9.48}$$

e

$$\text{Var}[e_t(h)] = \sigma_a^2 \sum_{j=0}^{h-1} \psi_j^2, \quad h \geq 1\tag{9.49}$$

Entretanto, para os modelos não-estacionários os pesos  $\psi_j$  não decaem para zero quando  $j$  aumenta. Além disso, para alguns modelos não-estacionários, a equação (8.49) mostra que a variância do erro de previsão crescerá sem limite quando  $h$  aumenta. Este fato não será também uma surpresa, desde que com séries não-estacionárias o futuro distante é quase incerto.

## 9.6 Atualização das Previsões

Vamos calcular as previsões de  $Z_{t+h+1}$  feitas a partir de duas origens:

$$t+1 : \widehat{Z}_{t+1}(h) = \psi_h a_{t+1} + \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots\tag{9.50}$$

$$t : \widehat{Z}_t(h+1) = \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots\tag{9.51}$$

Subtraindo (8.51) de (8.50) temos que:

$$\widehat{Z}_{t+1}(h) = \widehat{Z}_t(h+1) + \psi_h a_{t+1}\tag{9.52}$$

Assim, a previsão de  $Z_{t+h+1}$ , feita no instante  $t$ , pode ser atualizada quando um novo dado,  $Z_{t+1}$ , é observado. Deste modo, faremos a previsão de  $Z_{t+h+1}$ , na origem  $t+1$ , adicionando-se à  $\hat{Z}_t(h+1)$  um múltiplo do erro de previsão  $e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_{t+1}$ .

## 9.7 Intervalo de Confiança para Previsão

Para um modelo com tendência determinística com ruído branco  $a_t$ , nós temos que

$$\hat{Z}_t(h) = \mu_{t+l}$$

e

$$\text{Var}[e_t(h)] = \text{Var}(a_{t+h}) = \gamma_0$$

Se  $a_t \sim N(0, \sigma^2)$ , então o erro de previsão  $e_t(h) = Z_{t+h} - \hat{Z}_t(h) = a_{t+h}$  é normalmente distribuído.

Além disso, para um dado nível de confiança  $1 - \alpha$ , podemos utilizar a tabela da normal padronizada para encontrar um intervalo de  $(1 - \alpha/2) 100\%$  de confiança, ou seja

$$P \left[ -z(1 - \alpha/2) < \frac{Z_{t+h} - \hat{Z}_t(h)}{\sqrt{\text{Var}[e_t(h)]}} < z(1 - \alpha/2) \right] = 1 - \alpha \quad (9.53)$$

ou

$$P \left[ \hat{Z}_t(h) - z(1 - \alpha/2) \sqrt{\text{Var}[e_t(h)]} < Z_{t+h} < \hat{Z}_t(h) + z(1 - \alpha/2) \sqrt{\text{Var}[e_t(h)]} \right] \quad (9.54)$$

temos que  $(1 - \alpha/2) 100\%$  de confiança que as observações futuras estarão contida dentro dos limites

$$\hat{Z}_t(h) \pm z(1 - \alpha/2) \sqrt{\text{Var}[e_t(h)]} \quad (9.55)$$

## 9.8 Transformações e Previsões

Se  $Z_t$  é a série original, e  $Y_t = g(Z_t)$  é uma transformação (instantânea, que não envolve  $Z_{t-j}, j \geq 1$ ) de  $Z_t$ . Uma das principais razões de se fazer uma transformação é que  $Y_t$  pode ser Gaussiana e, neste caso, a previsão ótima (no sentido de mínimos quadrados) é uma função linear das observações.

Em Econmia, é comum termos séries com tendência na média, de modo que tomando-se diferenças obtêm-se séries estacionárias. Mas se a variância aumenta com o tempo, só tomar diferenças pode não ser suficiente e uma transformação dos dados deverá ser tentada. O usual, para séries econômicas, é tomar uma diferença do logaritmo da série original. Para que a transformação logarítmica seja apropriada, a média e o desvio-padrão (ou outra medida de variabilidade) deverão ser proporcionais.

O problema que se apresenta é o de obter previsão para  $Z_{t+h}$ , dado que temos um modelo para  $Y_t$  e temos previsões para:

$$Y_{t+h} = g(Z_{t+h}) \quad (9.56)$$

Uma maneira “ingênua” de proceder é considerar a equação (8.56) e substituir previsões por valores futuros:

$$\hat{Y}_t(h) = g(\hat{Z}_t(h)). \quad (9.57)$$

Depois, tentamos obter  $\widehat{Z}_t(h)$  em função de  $\widehat{Y}_t(h)$  a partir de (8.57); em particular, se  $g$  admite inversa, temos que

$$\widehat{Z}_t(h) = g^{-1} \left( \widehat{Y}_t(h) \right).$$

Por exemplo, se

$$Y_t = \ln Z_t \quad (9.58)$$

então  $Z_t = \exp(Y_t)$ , e uma previsão para  $Z_{t+h}$  será

$$\widehat{Z}_t(h) = \exp \left( \widehat{Y}_t(h) \right) \quad (9.59)$$

Contudo, pode-se demonstrar que, no caso de  $Y_t$  ser gaussiana, a previsão ótima é:

$$\exp \left( \widehat{Y}_t(h) + \frac{1}{2} \text{Var}[e_t(h)] \right) \quad (9.60)$$

Vemos, então, que o procedimento (8.59) conduz a previsões viciadas e, como consequência, o EQM de previsão aumentará.

Se  $Y_t = \ln Z_t$  segue um modelo ARIMA, então sabemos que a distribuição condicional de  $Y_{t+h}$ , dado o passado, é  $N \left( \widehat{Y}_t(h), \text{Var}[e_t(h)] \right)$ , e um intervalo de confiança para  $Y_{t+h}$ , com coeficiente de 95% de confiança, será

$$\widehat{Y}_t(h) \pm 1.96 \left( \widehat{\text{Var}}[e_t(h)] \right)^{0.5} \quad (9.61)$$

Daqui, segue-se que um intervalo de confiança para  $Z_{t+h}$ , com coeficiente de 95% de confiança, será

$$\exp \left( \widehat{Y}_t(h) \pm 1.96 \left( \widehat{\text{Var}}[e_t(h)] \right)^{0.5} \right) \quad (9.62)$$

Lembremos que  $\widehat{\text{Var}}[e_t(h)]$  é a estimativa de  $\text{Var}[e_t(h)]$ , com  $\sigma_a^2$  substituído por  $\widehat{\sigma}_a^2$ , no ajuste do modelo  $Y_t$ .

## 9.9 Resumo de Previsões para alguns Modelos ARIMA

### 9.9.1 AR(1): $Z_t - \mu = \phi(Z_{t-1} - \mu) + a_t$

$$\begin{aligned} \widehat{Z}_t(h) &= \mu + \phi \left( \widehat{Z}_t(h-1) - \mu \right) \\ &= \mu + \phi^h (Z_t - \mu) \\ &\approx \mu \text{ para } h \text{ grande} \\ \text{Var}[e_t(h)] &= \sigma_a^2 \frac{1 - \phi^{2h}}{1 - \phi^2} \\ &\approx \frac{\sigma_a^2}{1 - \phi^2} \text{ para } h \text{ grande} \\ \psi_j &= \phi^j \text{ para } j > 0 \end{aligned}$$

### 9.9.2 MA(1): $Z_t = \mu + a_t - \theta a_{t-1}$

$$\begin{aligned}\widehat{Z}_t(h) &= \begin{cases} \mu - \theta a_t & \text{para } h = 1 \\ \mu & \text{para } h > 1 \end{cases} \\ \text{Var}[e_t(h)] &= \begin{cases} \sigma_a^2 & \text{para } h = 1 \\ \sigma_a^2(1 + \theta^2) & \text{para } h > 1 \end{cases} \\ \psi_j &= \begin{cases} -\theta & \text{para } h = 1 \\ 0 & \text{para } h > 1 \end{cases}\end{aligned}$$

### 9.9.3 IMA(1,1) com termo constante: $Z_t = Z_{t-1} + \theta_0 + a_t - \theta a_{t-1}$

$$\begin{aligned}\widehat{Z}_t(1) &= Z_t + \theta_0 - \theta a_t \\ \widehat{Z}_t(h) &= Z_t + h\theta_0 - \theta a_t \\ &= h\theta_0 + (1 - \theta) \sum_{j=0}^{\infty} \theta^j Z_{t-j} \\ \text{Var}[e_t(h)] &= \sigma_a^2 [1 + (h-1)(1 - \theta^2)] \\ \psi_j &= 1 - \theta \text{ para } j > 0\end{aligned}$$

Note que se  $\theta_0 \neq 0$ , a previsão segue uma reta com inclinação  $\theta_0$ , mas se  $\theta_0 = 0$ , que é o caso usual, então  $\widehat{Z}_t(h)$  é o mesmo para todo  $h$ .

### 9.9.4 IMA(2,2): $Z_t = 2Z_{t-1} - Z_{t-2} + \theta_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$

$$\begin{aligned}\widehat{Z}_t(1) &= 2Z_t - Z_{t-1} + \theta_0 - \theta_1 a_t - \theta_2 a_{t-1} \\ \widehat{Z}_t(2) &= 2\widehat{Z}_t(1) - Z_t + \theta_0 - \theta_2 a_t \\ \widehat{Z}_t(h) &= 2\widehat{Z}_t(h-1) - \widehat{Z}_t(h-2) + \theta_0, \text{ para } h > 2 \\ &= A + Bh + \frac{\theta_0}{2} h^2, \text{ para } h \geq 1\end{aligned} \tag{9.63}$$

onde

$$A = 2\widehat{Z}_t(1) - \widehat{Z}_t(2) + \theta_0 \tag{9.64}$$

e

$$B = \widehat{Z}_t(2) - \widehat{Z}_t(1) - \frac{3}{2}\theta_0 \tag{9.65}$$

Se  $\theta_0 \neq 0$ , então a previsão segue uma curva quadrática em  $h$ , mas se  $\theta_0 = 0$ , a previsão forma uma linha reta com inclinação  $\widehat{Z}_t(2) - \widehat{Z}_t(1)$  e passará através das duas previsões iniciais  $\widehat{Z}_t(1)$  e  $\widehat{Z}_t(2)$ . Podemos mostrar que a  $\text{Var}[e_t(h)]$  é uma função cúbica de  $h$  (Box e Jenkins, 1976). Nós temos também

$$\psi_j = 1 + \theta_2 + (1 - \theta_1 - \theta_2)j \text{ para } j > 0 \tag{9.66}$$

## 9.10 Exercícios

1) Para um modelo AR(1) com  $Z_t = 12.2$ ,  $\phi = -0.5$ , e  $\mu = 10.8$ , encontrar  $\widehat{Z}_t(1)$ .

2) Suponha que as vendas anuais (em milhões de reais) de uma empresa, segue o modelo AR(2):

$$Z_t = 5 + 1.1Z_{t-1} - 0.5Z_{t-2} + a_t, \text{ com } \sigma_a^2 = 2$$

a) Se as vendas para 1985, 1984, e 1983 foram R\$10 milhões, R\$11 milhões, e R\$9 milhões, respectivamente, prever as vendas para 1986 e 1987.

b) Mostrar que  $\psi_1 = 1.1$  para este modelo.

c) Calcule um intervalo de 95% de confiança para os anos de 1986 e 1987.

d) Se as vendas em 1986 são de R\$12 milhões, qual será a previsão para 1987.

3) Considere o modelo

$$Z_t = \beta_0 + \beta_1 t + X_t \text{ com } X_t = \phi X_{t-1} + a_t$$

Assumindo que  $\beta_0, \beta_1$ , e  $\phi$  são conhecidos. Mostrar que o *eqm* mínimo de previsão a  $h$  passos a frente pode ser escrito como

$$\hat{Z}_t(h) = \beta_0 + \beta_1(t+h) + \phi^h(Z_t - \beta_0 - \beta_1 t)$$

4) Verifique a equação (8.22)

5) Verifique a equação (8.39)

6) Obtenha a função de previsão  $\hat{Z}_t(h)$ ,  $h = 1, 2, \dots$ , para os seguintes modelos ARIMA:

a)  $(1 - X - 0.25X^2)Z_t = (1 + 0.5X)a_t$

b)  $Z_t(1 - 0.5X + 0.5X^2) = a_t(1 - 0.5X - 0.25X^2)$

7) Considere o modelo  $Z_t - \mu = 0.8(Z_t - \mu) + a_t$ ,  $a_t \sim N(0, 1)$ . Suponha os seguintes dados:

$t$	1	2	3	4	5	6	7
$Z_t$	6	5	4	6	4	7	5

Obtenha: a)  $\hat{Z}_t(h)$  para  $h = 1, 2, 3$  e 25.

b) A  $Var[e_t(h)]$  para  $h = 1, 2, 3$  e 25.

c) Os intervalos com 95% de confiança para  $Z_8, Z_9$  e  $Z_{10}$ .

8) As seguintes observações  $Z_{91}, Z_{92}, \dots, Z_{100}$  representam os valores de uma série temporal ajustada pelo modelo:

$$\begin{aligned} Z_t - Z_{t-1} &= a_t - 1.1a_{t-1} + 0.3a_{t-2}, \text{ onde os valores são} \\ \mathbf{Z} &= [166; 172; 172; 169; 164; 168; 171; 167; 168; 172] \end{aligned}$$

a) Calcule as previsões  $\hat{Z}_{100}(h)$ ,  $h = 1, 2, \dots, 10$  (utilize  $\hat{a}_{99} = \hat{a}_{100} = 0$ ).

b) Sabendo que  $\sigma_a^2 = 1.1$ , calcule a  $Var[e_t(h)]$ ,  $h = 1, 2, \dots, 10$  e construa intervalos de confiança para os valores  $Z_{t+h}$ .

c) Determine os coeficientes  $\psi_j$  e, utilizando a nova observação  $Z_{101} = 174$ , calcule as previsões atualizadas  $\hat{Z}_{101}(h)$ ,  $h = 1, \dots, 9$ .

9) Considere o modelo  $Z_t = 0.8Z_t + a_t$ , com  $a_t \sim N(0, 1)$ .

a) Obtenha  $\hat{Z}_t(h)$  para  $h = 1, 2, 3, 4$ .

b) Obtenha a  $Var[e_t(h)]$ ,  $h = 1, 2, 3, 4$ .

c) Suponha os dados:

$t$	1	2	3	4	5	6	7
$Z_t$	0.66	0.57	0.66	-1.47	-1.38	-1.9	-0.7

c.1) Calcule  $\hat{Z}_t(h)$  para  $h = 1, 2, 3, 4$ .

c.2) Obtenha um intervalo de 95% de confiança para  $Z_8$  e  $Z_9$ .

10) a) Mostre como se faz previsões no modelo ARIMA(1,1,1) utilizando a equação diferença. Utilize o fato que  $Z_{t+1} - \hat{Z}_t(1) = a_{t+1}$  para que as previsões fiquem apenas em função das observações.

b) Mostre como previsões podem ser atualizadas.

c) Utilizando o modelo do item a), suponha que sua série é  $Z_1, Z_2, \dots, Z_{100}$ . Quando o valor real  $Z_{101}$  for conhecido como você atualizaria sua previsão?



## Capítulo 10

# Modelos Sazonais

No capítulo 3 vimos um pouco sobre tendência sazonal, neste capítulo vamos estudar os modelos sazonais determinísticos e estocásticos, procurando dar mais ênfase aos modelos estocásticos, ou seja, os modelos de Box-Jenkins, para séries sazonais, chamados SARIMA.

É possível que, mesmo após eliminar a componente sazonal determinística, ainda reste correlação significativa em:

(i) "lags" de baixa ordem, indicando que os resíduos ainda são correlacionados, podendo-se ajustá-los através de um modelo ARIMA;

(ii) "lags" sazonais, isto é, múltiplos de período  $s$ . Isto significa que há necessidade de se considerar uma sazonalidade estocástica, ou seja, ajustar à série original um modelo ARIMA sazonal (SARIMA).

Consideraremos, por simplicidade, dados observados mensalmente e sazonalidade de período  $s = 12$ . Trataremos, separadamente, os dois tipos de sazonalidade.

### 10.1 Sazonalidade Determinística

Quando  $Z_t$  exibe um comportamento sazonal determinístico com período 12, um modelo que pode ser útil é:

$$Z_t = \mu_t + N_t \quad (10.1)$$

onde  $\mu_t$  é uma função determinística periódica, satisfazendo  $\mu_t - \mu_{t-12} = 0$ , ou

$$(1 - X^{12}) \mu_t = 0 \quad (10.2)$$

e  $N_t$  é um processo estacionário que pode ser modelado por um ARMA( $p, q$ ).

Dessa maneira,  $N_t$  satisfaz à equação

$$\phi(X) N_t = \theta(X) a_t \quad (10.3)$$

onde  $a_t$  é ruído branco e  $\mu_t$  tem solução geral dada por:

$$\mu_t = \mu + \sum_{j=1}^6 \left[ \alpha_j \cos \frac{(2\pi jt)}{12} + \beta_j \sin \frac{(2\pi jt)}{12} \right] \quad (10.4)$$

onde  $\mu, \alpha_j, \beta_j, j = 1, \dots, 6$ , são constantes desconhecidas.

Para um modelo sazonal determinístico, aplicando a diferença sazonal  $(1 - X^{12})$  à expressão (9.1), temos:

$$(1 - X^{12}) Z_t = (1 - X^{12}) \mu_t + (1 - X^{12}) N_t \quad (10.5)$$

de acordo com (9.2), temos:

$$(1 - X^{12}) Z_t = (1 - X^{12}) N_t \quad (10.6)$$

Substituindo (9.3) em (9.5), temos:

$$\begin{aligned} (1 - X^{12}) Z_t &= (1 - X^{12}) \frac{\theta(X)}{\phi(X)} a_t \\ \phi(X) (1 - X^{12}) Z_t &= (1 - X^{12}) \theta(X) a_t \\ \phi(X) W_t &= (1 - X^{12}) \theta(X) a_t \end{aligned} \quad (10.7)$$

onde  $W_t = (1 - X^{12}) Z_t$ .

## 10.2 Identificação

A identificação de modelos da forma (9.7) é feita em 2 passos:

I. Obtemos as estimativas preliminares  $\tilde{\mu}, \tilde{\alpha}_j, \tilde{\beta}_j$  de  $\mu, \alpha_j$  e  $\beta_j$ ,  $j = 1, \dots, 6$ , em (9.4), através de uma análise de regressão de  $Z_t$  sobre 1,  $\sin \frac{(2\pi jt)}{12}$  e  $\cos \frac{(2\pi jt)}{12}$ ,  $j = 1, \dots, 6$ .

II. Calculamos os resíduos:

$$\tilde{N}_t = Z_t - \tilde{\mu} - \sum_{j=1}^6 \left[ \tilde{\alpha}_j \cos \frac{(2\pi jt)}{12} + \tilde{\beta}_j \sin \frac{(2\pi jt)}{12} \right]$$

e examinamos as funções de autocorrelação e autocorrelação parcial para identificar um modelo ARMA( $p, q$ ) para  $N_t$ .

## 10.3 Estimação

A estimação de máxima verossimilhança dos parâmetros  $\mu, \alpha_j$  e  $\beta_j$ ,  $j = 1, \dots, 6$ , é obtida por métodos similares aqueles apresentados na estimação dos parâmetros de um modelo ARMA.

## 10.4 Previsão

As previsões de valores futuros  $Z_{t+h}$ , dados  $Z_1, \dots, Z_t$ , são obtidos observando que:

$$\begin{aligned} \hat{Z}_t(h) &= E(Z_{t+h}/Z_t, Z_{t-1}, \dots, Z_1) \\ &= E(\mu_{t+h} + N_{t+h}/Y_t, Y_{t-1}, \dots, Y_1) \\ &= \mu_{t+h} + E(N_{t+h}/Y_t, Y_{t-1}, \dots, Y_1) \end{aligned}$$

vamos ter

$$\hat{Z}_t(h) = \mu_{t+h} + \hat{N}_t(h) \quad (10.8)$$

onde  $\mu_{t+h}$  e  $\hat{N}_t(h)$  são calculados utilizando os modelos (9.4) e (9.3), respectivamente. O  $\tilde{\mu}_t$  é obtido através de (9.4), substituindo cada um dos parâmetros pelos seus estimadores de mínimos quadrados de  $\mu, \alpha_j$  e  $\beta_j$  e substituindo  $N_t$  por  $\tilde{N}_t = Z_t - \tilde{\mu}_t$ .

## 10.5 Sazonalidade Estocástica

Pode ser adequado considerar  $\mu_t$  em (9.1) como um processo estocástico satisfazendo:

$$(1 - X^{12}) \mu_t = Y_t \quad (10.9)$$

onde  $Y_t$  é um processo estacionário. Aplicando, agora, o operador  $(1 - X^{12})$  à equação (9.1), obtemos:

$$(1 - X^{12}) Z_t = (1 - X^{12}) \mu_t + (1 - X^{12}) N_t$$

De acordo com (9.9), temos:

$$(1 - X^{12}) Z_t = Y_t + (1 - X^{12}) N_t \quad (10.10)$$

com  $\phi_Y(X) Y_t = \theta_Y(X) a_t$  e  $\phi_N(X) N_t = \theta_N(X) \varepsilon_t$ , onde  $a_t$  e  $\varepsilon_t$  são ruídos brancos independentes.

Podemos demonstrar que a expressão (9.10) é equivalente a:

$$(1 - \Phi_1 X^{12} - \dots - \Phi_P X^{12P}) (1 - X^{12})^D Z_t = (1 - \Theta_1 X^{12} - \dots - \Theta_Q X^{12Q}) \alpha_t \quad (10.11)$$

ou

$$\Phi(X^{12}) \Delta_{12}^D Z_t = \Theta(X^{12}) \alpha_t \quad (10.12)$$

onde  $\Phi(X) = (1 - \Phi_1 X^{12} - \dots - \Phi_P X^{12P})$  é o operador AR-Sazonal de ordem  $P$ , estacionário;  $\Theta(X) = (1 - \Theta_1 X^{12} - \dots - \Theta_Q X^{12Q})$  é o operador MA-Sazonal de ordem  $Q$ , invertível;  $\Delta_{12} = (1 - X^{12})$  é o operador diferença Sazonal;  $\Delta_{12}^D = (1 - X^{12})^D$ , indicando o número de "diferenças sazonais", onde  $\alpha_t$  pode ser, eventualmente, ruído branco; neste caso, a *fac* do processo  $Z_t$  é zero para todos os "lags" não-sazonais e o modelo (9.11) é denominado modelo sazonal puro.

Supondo, que o processo  $\alpha_t$  em (9.12) satisfaz um modelo ARIMA( $p, d, q$ ),

$$\phi(X) \alpha_t = \theta(X) a_t \quad (10.13)$$

onde  $a_t$  é um processo de ruído branco. Então, demonstra-se que  $Z_t$  satisfaz o modelo

$$\phi(X) \Phi(X^{12}) \Delta_{12}^D Z_t = \theta(X) \Theta(X^{12}) a_t \quad (10.14)$$

onde  $\phi(X) = (1 - \phi_1 X - \dots - \phi_p X^p)$  e  $\theta(X) = (1 - \theta_1 X - \dots - \theta_q X^q)$ , e os demais polinômios definidos em (9.12).

O modelo (9.14) é denominado ARIMA sazonal multiplicativo (SARIMA) de ordem  $(p, d, q) \times (P, D, Q)_{12}$ .

## 10.6 Identificação, Estimação e Verificação

Em princípio não teremos nenhuma dificuldade adicional nestas três etapas para os modelos sazonais. A diferença é que temos que *diferençar* a série com respeito a  $\Delta$  e  $\Delta_{12}$  (se  $s = 12$ ), a fim de produzir estacionariedade, obtendo valores para  $d$  e  $D$ .

Inspecionamos as *fac* e *facp* amostrais da série adequadamente *diferenciada* nos "lags" 1, 2, 3, ... para obter valores de  $p$  e  $q$  nos "lags" 12, 24, 36, ... para obter valores de  $P$  e  $Q$ , selecionando-se, desse modo, um modelo tentativo.

Estimam-se os valores dos parâmetros identificados, utilizando estimadores de máxima verossimilhança.

Finalmente, para verificar se o modelo proposto é adequado, utilizam-se testes de autocorrelação residual, Box-Pierce, etc.

## 10.7 Exemplos de Modelos Sazonais

### 10.7.1 Exemplo 1

Modelo Sazonal MA( $Q$ ) $s$  de ordem  $Q$ :

$$Z_t = a_t - \Theta_1 a_{t-s} - \dots - \Theta_Q a_{t-Qs} \quad (10.15)$$

com polinômio característico sazonal MA

$$\Theta(X) = (1 - \Theta_1 X^s - \dots - \Theta_Q X^{sQ}) \quad (10.16)$$

podemos verificar que esta série é estacionária e que a fac será diferente de zero somente nos "lags" sazonais de  $s, 2s, 3s, \dots, Qs$ . Temos um caso especial quando  $q = Qs$ , com todos os  $\theta'_s = 0$ , exceto para  $s, 2s, \dots, Qs$ . Particularmente,

$$\rho_{ks} = \frac{-\Theta_k + \Theta_1 \Theta_{k+1} + \dots + \Theta_{Q-k} \Theta_Q}{1 + \Theta_1^2 + \dots + \Theta_Q^2}, k = 1, \dots, Q \quad (10.17)$$

o modelo será invertível se as raízes de  $\Theta(X) = 0$  excede a unidade em valor absoluto.

### 10.7.2 Exemplo 2

Modelo Sazonal AR( $P$ ) $s$  de ordem  $P$ :

$$Z_t = \Phi_1 Z_{t-s} + \dots + \Phi_P Z_{t-Ps} + a_t \quad (10.18)$$

com polinômio característico sazonal AR

$$\Phi(X) = (1 - \Phi_1 X^{12} - \dots - \Phi_P X^{12P}) \quad (10.19)$$

Estamos supondo que  $a_t$  é independente de  $Z_{t-1}, Z_{t-2}, \dots$  e, para a série ser estacionária as raízes de  $\Phi(X) = 0$  são maiores que 1 em valor absoluto. Novamete um caso especial é quando  $p = Ps$ , com todos os  $\phi'_s = 0$ , exceto nos "lags"  $s, 2s, \dots, Ps$ .

### 10.7.3 Exemplo 3

Modelo Sazonal Multiplicativo ARMA( $p, q$ ) $\times$ ( $P, Q$ ) $s$ :

$$\phi(X) \Phi(X^{12}) Z_t = \theta(X) \Theta(X^{12}) a_t \quad (10.20)$$

com polinômios característicos AR's  $\phi(X)$  e  $\Phi(X^{12})$ , e polinômios MA's  $\theta(X)$  e  $\Theta(X^{12})$ .

### 10.7.4 Exemplo 4

Modelo ARIMA( $0, 1, 1$ ) $\times$ ( $1, 0, 1$ ) $12$

$$Z_t - Z_{t-1} = \Phi(Z_{t-12} - Z_{t-13}) + a_t - \theta a_{t-1} - \Theta a_{t-12} + \theta \Theta a_{t-13} \quad (10.21)$$

A função de previsão a um passo é dada por:

$$\begin{aligned} \hat{Z}_t(1) &= Z_t + \Phi Z_{t-11} - \Phi Z_{t-12} - \theta a_t - \Theta a_{t-12} + \theta \Theta a_{t-13} \\ \hat{Z}_t(2) &= \hat{Z}_t(1) + \Phi Z_{t-10} - \Phi Z_{t-11} - \Theta a_{t-10} + \theta \Theta a_{t-11} \end{aligned} \quad (10.22)$$

Os termos do ruído  $a_{t-13}, a_{t-12}, \dots, a_t$  entram na previsão para  $h = 1, 2, \dots, 13$ , mas para  $h > 13$  tomamos a parte autoregressiva do modelo, e temos

$$\hat{Z}_t(2) = \hat{Z}_t(h-1) + \Phi Z_t(h-12) - \Phi Z_t(h-13), \text{ para } h > 13 \quad (10.23)$$

Para entender melhor os modelos de previsão, vamos considerar alguns casos especiais a seguir.

### 10.7.5 Exemplo 5

O modelo sazonal AR(1)12

$$Z_t = \Phi Z_{t-12} + a_t \quad (10.24)$$

Vamos ter então

$$\hat{Z}_t(h) = \Phi \hat{Z}_t(h-12) \quad (10.25)$$

Entretanto podemos escrevê-lo como

$$\hat{Z}_t(h) = \Phi^{k+1} Z_{t+r-11} \quad (10.26)$$

onde  $k$  e  $r$  são definidos por  $h = 12k + r + 1$  com  $0 \leq r < 12$  e  $k = 0, 1, 2, \dots$ . Em outras palavras,  $k$  é a parte inteira de  $(h-1)/12$ , e  $r/12$  é a parte fracional de  $(h-1)/12$ .

Vamos determinar os pesos  $\psi'_s$  para esse modelo, sabendo que, somente para os múltiplos de 12, eles são diferente de zero, ou seja

$$\psi_j = \begin{cases} \Phi^{j/12}, j = 0, 12, 24, \dots \\ 0, \text{ caso contrário} \end{cases} \quad (10.27)$$

temos então que a variância do erro de previsão pode ser escrita como

$$Var[e_t(h)] = \frac{1 - \Phi^{2k+2}}{1 - \Phi^2} \sigma_a^2 \quad (10.28)$$

onde  $h = 12k + r + 1$ .

### 10.7.6 Exemplo 6

Modelo sazonal MA(1)12, temos

$$Z_t = a_t - \Theta a_{t-12} + \theta_0 \quad (10.29)$$

Neste caso temos a seguinte função de previsão

$$\begin{aligned} \hat{Z}(1) &= -\Theta a_{t-11} + \theta_0 \\ \hat{Z}(2) &= -\Theta a_{t-10} + \theta_0 \\ &\vdots \\ \hat{Z}(h) &= \theta_0, \text{ para } h > 12 \end{aligned} \quad (10.30)$$

Aqui obtemos diferentes previsões para os meses do primeiro ano, mas para os anos seguintes as previsões são dadas pela média do processo.

Para este modelo  $\psi_0 = 1$ ,  $\psi_{12} = -\Theta$ , e  $\psi_j = 0$  caso contrário. Além disso, teremos a seguinte variância para o erro de previsão

$$Var[e_t(h)] = \begin{cases} \sigma_a^2, & 1 \leq h \leq 12 \\ (1 + \Theta^2) \sigma_a^2, & h > 12 \end{cases} \quad (10.31)$$

### 10.7.7 Exemplo 7

Modelo ARIMA(0, 0, 0)  $\times$  (0, 1, 1) 12, temos

$$Z_t - Z_{t-12} = a_t - \Theta a_{t-12} \quad (10.32)$$

Com função de previsão dada por:

$$\begin{aligned}\widehat{Z}(1) &= Z_{t-11} - \Theta a_{t-11} \\ \widehat{Z}(2) &= Z_{t-10} - \Theta a_{t-10} \\ &\vdots \\ \widehat{Z}_t(12) &= \widehat{Z}_t(h-12), \quad h > 12\end{aligned}\tag{10.33}$$

Observamos que todos os meses de janeiro, fevereiro,..., dezembro, têm previsões idênticas entre si, e assim por diante.

Neste caso temos  $\psi_j = 1 - \Theta$ , para  $j = 12, 24, \dots$ , e zero caso contrário. A variância do erro de previsão é dada, então, por

$$\text{Var}[e_t(h)] = \sigma_a^2 \left[ 1 + k(1 - \Theta)^2 \right],\tag{10.34}$$

onde  $h = 12k + r + 1$ ,  $k = 0, 1, \dots$ , e  $0 \leq r < 12$ .

### 10.7.8 Exemplo 8

Modelo  $\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$ , temos

$$Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} + a_t - \theta a_{t-1} - \Theta a_{t-12} + \theta \Theta a_{t-13}\tag{10.35}$$

A função de previsão satisfaz

$$\begin{aligned}\widehat{Z}_t(1) &= Z_t + Z_{t-11} - Z_{t-12} - \theta a_{t-1} - \Theta a_{t-11} + \theta \Theta a_{t-12} \\ \widehat{Z}_t(2) &= \widehat{Z}_t(1) + Z_{t-10} - Z_{t-11} - \Theta a_{t-10} + \theta \Theta a_{t-11} \\ &\vdots \\ \widehat{Z}_t(12) &= \widehat{Z}_t(11) + Z_t - Z_{t-1} - \Theta a_t + \theta \Theta a_{t-1} \\ \widehat{Z}_t(13) &= \widehat{Z}_t(12) + \widehat{Z}_t(1) - Z_t + \theta \Theta a_t\end{aligned}\tag{10.36}$$

e

$$\widehat{Z}_t(h) = \widehat{Z}_t(h-1) + \widehat{Z}_t(h-12) - \widehat{Z}_t(h-13), \quad \text{para } h > 13$$

Para entender o padrão geral dessas previsões, podemos utilizar a representação

$$\widehat{Z}_t(h) = A_1 + A_2 h + \sum_{j=0}^6 \left[ B_{1j} \cos \frac{2\pi j h}{12} + B_{2j} \sin \frac{2\pi j h}{12} \right], \quad h > 0\tag{10.37}$$

onde  $A_j$  e  $B_j$  são independentes de  $Z_t, Z_{t-1}, \dots$ , ou, alternativamente, determinados das previsões iniciais  $\widehat{Z}_t(1), \widehat{Z}_t(2), \dots, \widehat{Z}_t(13)$ .

Este resultado segue-se da teoria geral das equações diferenças e envolve as raízes de  $(1 - X)(1 - X^{12}) = 0$ . (ver Abraham & Box, 1978).

## 10.8 Exercícios

1) Baseado em dados trimestrais, um modelo sazonal da forma

$$Z_t = Z_{t-4} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

foi ajustado para série temporal.

a) Encontrar os 4 primeiros pesos  $\psi'_s$  deste modelo.

b) Suponha que  $\theta_1 = 0.5$ ,  $\theta_2 = -0.25$ , e  $\sigma_a^2 = 1$  e que os dados para os últimos 4 trimestres são:

	I	II	III	IV
Série	25	20	25	40
Resíduo	2	1	2	3

Encontre a previsão para os próximos 4 trimestres.

c) Encontre um intervalo de 95% de confiança para as previsões em b.

**2)** Para o modelo sazonal  $Z_t = \Phi Z_{t-4} + a_t - \theta a_{t-1}$ , com  $|\Phi| < 1$ , encontrar  $\gamma_0$  e  $\rho_k$  para  $k > 0$ .

**3)** Identifique os seguintes modelos sazonais multiplicativos ARIMA:

a)  $Z_t = 0.5Z_{t-1} + Z_{t-4} - 0.5Z_{t-5} + a_t - 0.3a_{t-1}$ .

b)  $Z_t = Z_{t-1} + Z_{t-12} - Z_{t-13} + a_t - 0.5a_{t-1} - 0.5a_{t-12} + 0.25a_{t-13}$ .

**4)** Suponha que o processo  $\{Z_t\}$  se desenvolve de acordo com

$$Z_t = Z_{t-4} + a_t, \text{ para } t > 4$$

com  $Z_t = a_t$  para  $t = 1, 2, 3, 4$ . Encontrar a função variância e a função de autocorrelação para  $\{Z_t\}$ .

## Apêndice A

# Esperança Condicional

Se  $X$  e  $Y$  tem f.d.p. conjunta  $f(x, y)$  e seja  $f(x)$  a f.d.p. marginal de  $X$ , então a f.d.p. condicional de  $Y|X = x$  é dada por

$$f(y|x) = \frac{f(x, y)}{f(x)}.$$

A esperança condicional de  $Y|X = x$  é então definida como

$$E(Y|X = x) = \int_{-\infty}^{\infty} y f(y|x) dy.$$

Observe que é a média da distribuição condicional, e além disso, tem todas as propriedades de médias comum e valores esperados. Por exemplo,

$$E(aY + bZ|X = x) = aE(Y|X = x) + bE(Z|X = x)$$

e

$$E[h(Y)|X = x] = \int_{-\infty}^{\infty} h(y) f(y|x) dy.$$

Pode-se demonstrar a seguinte propriedade

$$E[h(X)|X = x] = h(X)$$

a variável aleatória  $h(X)$  pode ser tratada como uma constante. Contudo geralmente,

$$E[h(X, Y)|X = x] = E[h(x, Y)|X = x].$$

Se fizer  $E[h(X)|X = x] = g(x)$ , então  $g(X) = Y$  é uma v.a., e

$$E[g(X)] = E(Y).$$

Que pode ser escrito como  $E[E(Y|X)] = E(Y)$ . Se  $Y$  é independente de  $X$ , então  $E(Y|X) = E(Y)$ .



## Apêndice B

# Predição de Erro Quadrático Médio Mínimo

Suponha  $Y$  uma v.a. com média  $\mu_y$  e variância  $\sigma_y^2$ . Se o objetivo é fazer a predição de  $Y$  utilizando somente uma constante  $c$ , qual é a melhor escolha para  $c$ ? Um bom critério é escolher um valor de  $c$  que minimiza o erro quadrático médio (EQM) de predição, ou seja,

$$\min g(c) = E(Y - c)^2.$$

Expandindo  $g(c)$ , temos

$$g(c) = E(Y^2) - 2cE(Y) + c^2$$

como  $g(c)$  é uma função quadrática em  $c$ , resolvendo a equação  $g'(c) = 0$ , encontra-se o ponto de mínimo da função, tem-se então

$$g'(c) = -2E(Y) + 2c = 0$$

encontrando

$$c = E(Y) = \mu_y \quad (\text{B.1})$$

nota-se que

$$\min g(c) = E(Y - \mu_y)^2 = \sigma_y^2. \quad (\text{B.2})$$

Agora considere a situação em que uma segunda v.a.  $X$  é utilizada, e deseja-se utilizar o valor observado de  $X$  para prever  $Y$ . Seja  $\rho = \text{Corr}(X, Y)$ .

Supondo que, somente funções lineares  $a + bX$  podem ser utilizadas para predição. O EQM é então dado por

$$g(a, b) = E(Y - a - bX)^2$$

expandindo a função  $g(a, b)$  tem-se

$$g(a, b) = E(Y^2) + a^2 + b^2E(X^2) - 2aE(Y) + 2abE(X) - 2bE(XY).$$

A função é quadrática em  $a$  e  $b$ . Além disso, pode-se encontrar o ponto de mínimo resolvendo simultaneamente as equações lineares

$$\frac{\partial g(a, b)}{\partial a} = 0 \text{ e } \frac{\partial g(a, b)}{\partial b} = 0.$$

tem-se

$$\frac{\partial g(a, b)}{\partial a} = 2a - 2E(Y) + 2bE(X) = 0$$

e

$$\frac{\partial g(a, b)}{\partial b} = 2bE(X^2) + 2aE(X) - 2E(XY) = 0$$

as quais podemos escrever como

$$a + E(X)b = E(Y) \quad (\text{B.3})$$

$$E(X)a + E(X^2)b = E(XY) \quad (\text{B.4})$$

multiplicando (B.3) por  $E(X)$  e subtraindo de (B.4) obtém-se:

$$E(X^2)b - E^2(X)b = E(XY) - E(X)E(Y)$$

obtendo

$$b = \frac{E(XY) - E(X)E(Y)}{E(X^2) - E^2(X)} = \frac{Cov(X, Y)}{Var(X)}$$

ou

$$b = Corr(X, Y) \sqrt{\frac{Var(Y)}{Var(X)}} = \rho \frac{\sigma_y}{\sigma_x}$$

então substituindo em (B.3) o valor de  $b$  obtém-se:

$$a = E(Y) - \rho \frac{\sigma_y}{\sigma_x} E(X)$$

$$a = \mu_y - \rho \frac{\sigma_y}{\sigma_x} \mu_x$$

Se  $\hat{Y}$  é a predição de EQM mínimo de  $Y$ , baseado numa função linear de  $X$ , então pode-se escrever

$$\hat{Y} = \mu_y - \rho \frac{\sigma_y}{\sigma_x} \mu_x + \rho \frac{\sigma_y}{\sigma_x} X \quad (\text{B.5})$$

ou

$$\frac{\hat{Y} - \mu_y}{\sigma_y} = \rho \frac{X - \mu_x}{\sigma_x}$$

em termos de v.a. padronizadas  $Y'$  e  $X'$ , tem-se  $Y' = \rho X'$ .

Pode-se verificar também que

$$\min g(\hat{a}, \hat{b}) = E \left[ Y - \rho \frac{\sigma_y}{\sigma_x} \mu_x + \rho \frac{\sigma_y}{\sigma_x} X \right]^2 = \sigma_y^2 (1 - \rho^2).$$

*Prova*

$$\begin{aligned} \min g(\hat{a}, \hat{b}) &= E \left[ Y - \rho \frac{\sigma_y}{\sigma_x} \mu_x + \rho \frac{\sigma_y}{\sigma_x} X \right]^2 \\ &= \sigma_y^2 E \left( \frac{\hat{Y} - \mu_y}{\sigma_y} - \rho \frac{X - \mu_x}{\sigma_x} \right)^2 \\ &= \sigma_y^2 [Var(Y') + \rho^2 Var(X') - 2\rho Cov(Y', X')] \\ &= \sigma_y^2 (1 + \rho^2 - 2\rho\rho) = \sigma_y^2 (1 - \rho^2). \end{aligned}$$

Considere agora o caso mais geral do problema de predição de  $Y$  com uma função arbitrária de  $X$ . Desde que o critério escolhido seja minimizar o EQM de predição, precisa-se escolher uma função  $h(X)$  que minimize

$$E[Y - h(X)]^2 \quad (\text{B.6})$$

pode-se escrever esta função como:

$$E [Y - h (X)]^2 = E \left\{ E \left( [Y - h (X)]^2 | X \right) \right\} \quad (\text{B.7})$$

para  $X = x$  obtém-se

$$E \left( [Y - h (X)]^2 | X = x \right) = E \left( [Y - h (x)]^2 | X = x \right) \quad (\text{B.8})$$

para cada valor de  $x$ ,  $h (x)$  é uma constante, e pode-se aplicar o resultado  $E (Y) = \mu_y$  para a distribuição condicional  $Y|X = x$ . Além disso, para cada  $x$ , a melhor escolha de  $h (x)$  é

$$h (x) = E (Y|X = x) .$$

Desde que a escolha de  $h (x)$  minimize a esperança interna na equação (B.7), a função  $h (x)$ , também fornece o mínimo da equação (B.5), e

$$h (X) = E (Y|X) \quad (\text{B.9})$$

é o melhor preditor de  $Y$  de todas as funções de  $X$ .

Se  $X$  e  $Y$  tem distribuição normal bivariada, mostra-se que

$$E (Y|X) = \mu_y + \rho \frac{\sigma_y}{\sigma_x} (X - \mu_x) \quad (\text{B.10})$$

observa-se que as soluções (B.10) e (B.5) coincidem. Neste caso, a melhor de todas as funções é linear.

Agora se a predição de  $Y$  for baseada numa função de variáveis aleatórias  $X_1, X_2, \dots, X_n$ , então pode-se argumentar facilmente que o preditor de EQM mínimo é dado por

$$E (Y|X_1, X_2, \dots, X_n)$$

# Referências Bibliográficas

- [1] BOX, G.E.P. & JENKINS, G.W. *Time Series Analysis, Forecasting and Control*. Holden-Day, 1976.
- [2] BOX, G.E.P. & JENKINS, G.W. and REINSEL, G.C. *Time Series Analysis: Forecasting and Control*. Third Edition. Prentice-Hall, Englewood Cliffs, N.J., 1994
- [3] MORETTIN, P.A. & TOLOI, C.M. *Previsão de Séries Temporais*. São Paulo, Atual, 1985.
- [4] MORETTIN, P.A. & TOLOI, C.M. *Análise de Séries Temporais*. Editora Edgard Blücher, São Paulo, 2004.
- [5] CRYER, J.D. *Time Series Analysis*. PWA- Dent Publishing Company-Boston, 1986.
- [6] CHATFIELD, C. *The Analysis of Time Series: An Introduction*. Chapman-Hall, 1984.
- [7] MONTGOMERY, D.C. & JONHSON, L.A. *Forecasting and Time Series Analysis* New York, Mcgraw-Hill, 1976.